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.

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.
- 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.

Linear Mixed Effect Model

- 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 β, random effects b_i occur linearly in the model.

Non-linear Mixed Effect Model

- The non-linear mixed effect model extends from the linearity assumption to allow for more flexible function forms.

Non-linear Mixed Effect Model

- 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 ϕ_i is modeled as,

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

where $b_i \sim N(0, \Psi)$,β is a p-dimensional vector of fixed effects and b_i is a q-dimensional random effects vector associated with the i-th group.

Carbon Dioxide Uptake

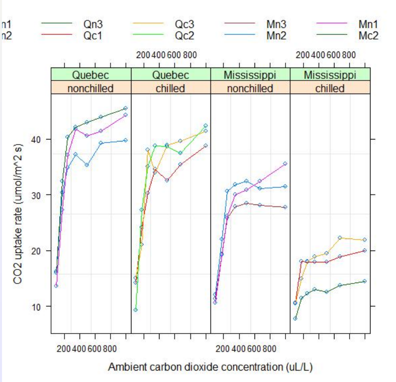
- Data is contained in the NLME library, from a study of the cold tolerance of a C4 grass species. A total of 12 fourweek-old plants, 6 from Quebec and 6 from Mississippi, were divided into two groups: control plants that were kept at 26 degree and chilled plants that were subject to 14 h of chilling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

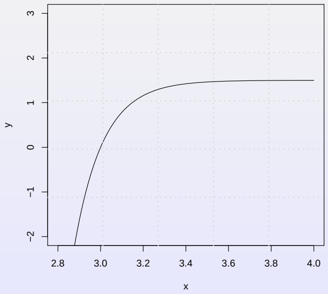
Figure: CO2 uptake versus ambient CO2 by treatment and type f, 6 from Quebec and 6 from Mississippi. Half the plants of each type were chilled overnight before the measurements were taken.



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:

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

```
> 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)
```



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?

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 $\times 3$.

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 $\times 3$.
- 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.

nlsList()

- A typical call to *nlsList* is $nlsList(model, data)$.

nlsList()

- A typical call to *nlsList* is $nlsList(model, data)$.
- Note that *nlsList*() requires initial value for the model.

nlsList()

- 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).

nlsList()

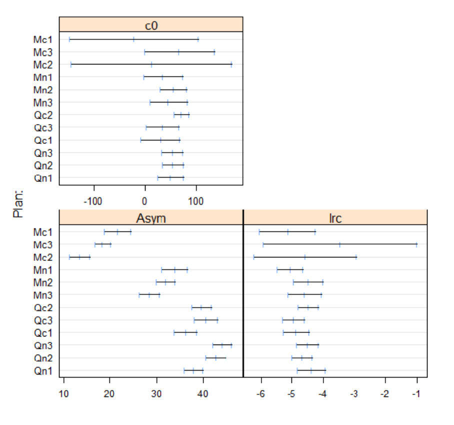
- A typical call to *nlsList* is *nlsList*(*model*, *data*).
- Note that *nlsList*() requires initial value for the model.
- A *selfStart*() function, which specify the function formula and can automatically generate individual initial estimates for each group, is recommended.(See details in the book <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:
        Asym       lrc        c0
Qn1 38.13978 -4.380647  51.22324
Qn2 42.87169 -4.665728  55.85816
Qn3 44.22800 -4.486118  54.64958
Qc1 36.42873 -4.861741  31.07538
Qc3 40.68370 -4.945218  35.08889
Qc2 39.81950 -4.463838  72.09422
Mn3 28.48285 -4.591566  46.97188
Mn2 32.12827 -4.466157  56.03863
Mn1 34.08481 -5.064579  36.40805
Mc2 13.55520 -4.560851  13.05675
Mc3 18.53506 -3.465158  67.84877
Mc1 21.78723 -5.142256 -20.39998

Degrees of freedom: 84 total; 48 residual
Residual standard error: 1.79822
```

intervals() function returns the information about confidence intervals of estimates.

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

Figure: Confidence Intervals for parameters Asym, lrc.



nlsList

- The plot of the individual confidence intervals from fm1CO2.lis indicates that there is substantial between-plant variation in the parameter ϕ_1 = Asym, and only moderate variation among ϕ_2 = lrc, and ϕ_3 = c0.

nlsList

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

nlsList

- 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 approriate.
- 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 *nlsList*() CO2.list can be used to implement a full mixed effect model.

```
> 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
```

A strong correlation between c0 and lrc, suggests that only one of the random effects is needed.

> CO2.nlme2 <- update(CO2.nlme, random = Asym + lrc ~ 1)

```
Nonlinear mixed-effects model fit by maximum likelihood
  Model: uptake ~ SSasympOff(conc, Asym, lrc, c0)
  Data: CO2
  Log-likelihood: -202.7583
  Fixed: list(Asym ~ 1, lrc ~ 1, c0 ~ 1)
     Asym       lrc        c0
32.411764 -4.560265 49.343573

Random effects:
 Formula: list(Asym ~ 1, lrc ~ 1)
 Level: Plant
 Structure: General positive-definite, Log-Cholesky parametrization
         StdDev    Corr
Asym     9.6593926 Asym
lrc      0.1995124 -0.777
Residual 1.8079224
```

```
> anova(CO2.nlme, CO2.nlme2)

          Model df      AIC      BIC    logLik   Test L.Ratio p-value
CO2.nlme      1 10 422.6212 446.9293 -201.3106
CO2.nlme2     2  7 419.5167 436.5324 -202.7583 1 vs 2 2.89549   0.408
```

Add covariates into mixed effect model

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

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 ϕ_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 ϕ_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

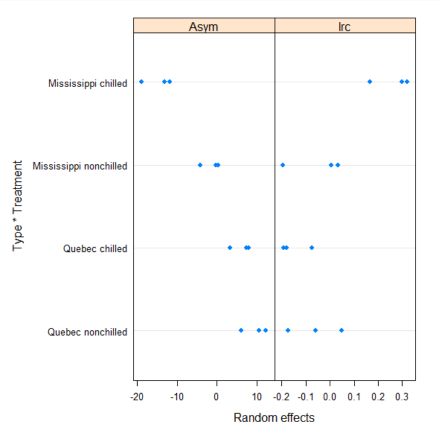
- Now the random effects accommodate individual deviations from the fixed effects.
- The primary question of interest for the CO2 data is the effect of plant type and chilling treatment on the individual model parameters ϕ_i.
- So, we want to know whether the variation between groups can be attributed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

           Asym          lrc        Type  Treatment conc   uptake
Qn1   6.1715987  0.048361985      Quebec nonchilled  435 33.22857
Qn2  10.5325882 -0.172842971      Quebec nonchilled  435 35.15714
Qn3  12.2180908 -0.057987100      Quebec nonchilled  435 37.61429
Qc1   3.3521234 -0.075586358      Quebec    chilled  435 29.97143
Qc3   7.4743083 -0.192416381      Quebec    chilled  435 32.58571
Qc2   7.9284657 -0.180323624      Quebec    chilled  435 32.70000
Mn3  -4.0733486  0.033449394 Mississippi nonchilled  435 24.11429
Mn2  -0.1419773  0.005645756 Mississippi nonchilled  435 27.34286
Mn1   0.2406596 -0.193859245 Mississippi nonchilled  435 26.40000
Mc2 -18.7991627  0.319367709 Mississippi    chilled  435 12.14286
Mc3 -13.1168244  0.299428913 Mississippi    chilled  435 17.30000
Mc1 -11.7865217  0.166761922 Mississippi    chilled  435 18.00000
```

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 Virsus Covariates



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.

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

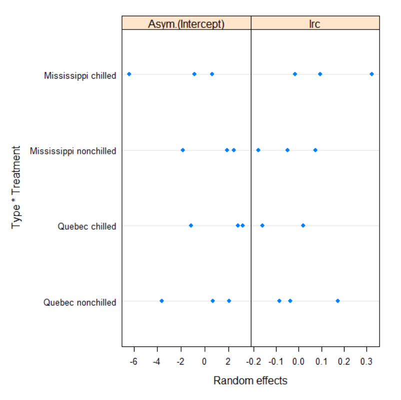
- The figure shows a strong relationship between the parameter asym and the covariates: Asym decreases when the plants are chilled and is higher among Quebec plants than Mississippi plants.
- The increase in Asym from chilled to nonchilled plants is larger among Mississippi plants than Quebec plants, 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
```

```
Fixed effects: list(Asym ~ Type * Treatment, lrc + c0 ~ 1)
                                        Value Std.Error DF   t-value
Asym.(Intercept)                     42.17337  1.345946 67  31.33364
Asym.TypeMississippi                -11.82431  1.537304 67  -7.69159
Asym.Treatmentchilled                -5.23756  1.464784 67  -3.57565
Asym.TypeMississippi:Treatmentchilled -4.78084  2.353801 67  -2.03111
lrc                                  -4.58925  0.084821 67 -54.10511
c0                                   49.48163  4.456630 67  11.10293
                                    p-value
Asym.(Intercept)                     0.0000
Asym.TypeMississippi                 0.0000
Asym.Treatmentchilled                0.0007
Asym.TypeMississippi:Treatmentchilled 0.0462
lrc                                  0.0000
c0                                   0.0000
                                    As.(I) Asy.TM Asym.T A.TM:T lrc
Asym.TypeMississippi                -0.559
Asym.Treatmentchilled               -0.540  0.471
Asym.TypeMississippi:Treatmentchilled 0.252 -0.640 -0.622
lrc                                 -0.540  0.056 -0.016  0.132
c0                                  -0.086 -0.001 -0.036  0.063  0.65
```

```
> CO2.nlmeRE3 <- ranef( CO2.nlme3, aug = T )
> plot( CO2.nlmeRE3, form = ~ Type * Treatment )
```

Figure: Predicted random effect Virsus Covariate



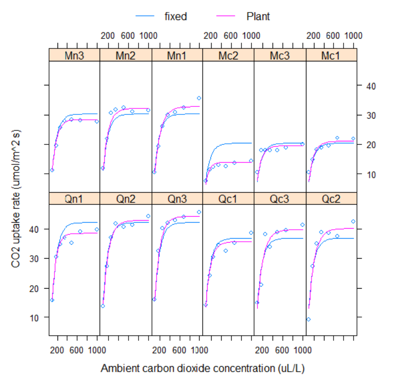
- After covariates have been introduced in the model to account for intergroup variation, a natural question is which random effects are still needed.

- 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.

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

Figure: Predicted value of uptake by groups



Thank you for your listening!