MIXED-EFFECTS MODELS METHODS AND CLASSES FOR S AND S PLUS

Jose C. Pinheiro, PhD
Department of Biostatistics

Douglas M. Bates, PhD
Department of Statistics

MADISON, WISCONSIN
Mixed-Effects Models Methods and Classes for S and S-plus

March 10, 1995

José C. Pinheiro and Douglas M. Bates
Department of Biostatistics Department of Statistics
University of Wisconsin – Madison

Abstract

Mixed-effects models provide a powerful and flexible tool for analyzing clustered data, such as repeated measures data and nested designs. We describe a set of S functions, classes, and methods for the analysis of both linear and nonlinear mixed-effects models. These extend the linear and nonlinear modeling facilities available in release 3 of S and S-plus. The source code, written in S and C is available in the S collection at StatLib. Help files for all functions and methods described here are included in Appendix A.
Contents

1 Introduction 1

2 The lme class and related methods 1
  2.1 The lme function ........................................... 2
  2.2 The print, summary, and anova methods. .................... 2
  2.3 The plot method ............................................. 4
  2.4 Other methods .............................................. 6
  2.5 Structured variance-covariance matrices and random effects blocks .......... 6

3 The nlme class and related methods 10
  3.1 The nlme function ............................................ 11
  3.2 Methods for nlme objects .................................... 12
  3.3 Structured variance-covariance matrices and random effects blocks .......... 16

4 Future Developments 16

Appendix A: Help Files 18
1 Introduction

Mixed-effects models provide a powerful and flexible tool for analyzing clustered data, such as repeated measures data and nested designs. We describe a set of S functions, classes, and methods for the analysis of linear or nonlinear mixed-effects models. These extend the linear and nonlinear modeling facilities available in release 3 of S (Chambers and Hastie, 1992) and S-plus.

The source code, written in S and C, is available in the S collection at StatLib. It can be obtained either through electronic mail, by sending the one-line message send nlme from S to the address statlib@lib.cmu.edu, or via anonymous ftp from ftp://lib.stat.cmu.edu/S/nlme or ftp://ftp.stat.wisc.edu/src/NLME/Unix. The StatLib version of the code is intended for Unix based machines. A MicroSoft Windows version for S-plus version 3.2 is available for anonymous ftp from ftp://ftp.stat.wisc.edu/src/NLME/Windows. Help files for all functions and methods described here can be found in Appendix A.

Section 2 presents the functions and methods for fitting and analyzing linear mixed-effects models. The nonlinear mixed-effects functions and methods are described in section 3. Section 4 presents some future directions for the code development.

2 The lme class and related methods

Fitting and analyzing linear mixed-effects models will be described here. First we consider the data from a dental study presented in Potthoff and Roy (1964). The data, displayed in Figure 1, consist of four measurements of the distance (in millimeters) from the center of the pituitary to the pterygomaxillary fissure made at ages 8, 10, 12, and 14 years on 16 boys and 11 girls. A linear model seems adequate to explain the distance as a function of age, but the intercept and the slope seem to vary with the individual. The corresponding linear mixed-effects model is

\[ d_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i}) \text{age}_j + \epsilon_{ij}, \]  

(1)

where \( d_{ij} \) represents the distance for the ith individual at age \( j \), \( \beta_0 \) and \( \beta_1 \) are the population average intercept and the population average slope, \( b_{0i} \) and \( b_{1i} \) are the effects in intercept and slope associated with the ith individual, and \( \epsilon_{ij} \) is the within-subject error term. It is assumed that the \( b_i = (b_{0i}, b_{1i})^T \) are independent and identically distributed with a \( N(0, \sigma^2 D) \) distribution and the \( \epsilon_{ij} \) are independent and identically distributed with a \( N(0, \sigma^2) \) distribution, independent of the \( b_i \).

One of the questions of interest for these data is whether these curves show significant differences between boys and girls. Model (1) can be modified as

\[ d_{ij} = (\beta_{00} + \beta_{01} \text{sex}_i + b_{0i}) + (\beta_{10} + \beta_{11} \text{sex}_i + b_{1i}) \text{age}_j + \epsilon_{ij} \]  

(2)

to test for sex related differences in intercept and slope. In model (2), \( \text{sex}_i \) is an indicator variable assuming value zero if the ith individual is a boy and one if she is a girl. \( \beta_{00} \) and \( \beta_{10} \) represent the population average intercept and slope for the boys and \( \beta_{01} \) and \( \beta_{11} \) are the changes in population average intercept and slope for girls. Differences between boys and girls can be evaluated by testing whether \( \beta_{01} \) and \( \beta_{11} \) are significantly different from zero. The remaining terms in (2) are defined as in (1). It will be assumed here that the data are available in a data frame called dental, with columns distance, age, subject, and sex as below.
Mixed effects methods and classes for S and S-plus

> dental
  distance age subject sex
  1 26.0  8 1 0
  2 25.0 10 1 0
  3 29.0 12 1 0
  4 31.0 14 1 0
  105 24.5  8 27 1
  106 25.0 10 27 1
  107 28.0 12 27 1
  108 28.0 14 27 1

2.1 The lme function

The lme function is used to fit a linear mixed-effects model, as described in Laird and Ware (1982), using either maximum likelihood or restricted maximum likelihood. It produces an object of class lme. Several optional arguments can be used with this function, but the typical call is

lme(fixed, random, cluster, data)

The first three arguments are required. The arguments fixed and random are formulas as shown below. Any linear model formula (Chambers and Hastie, 1992, chapter 3) is allowed, giving the model formulation considerable flexibility. For the dental data these formulas would be written as

fixed = distance ~ age, random = ~ age

for model (1) and

fixed = distance ~ age * sex, random = ~ age

for model (2). Note that the response variable is given only in the formula for the fixed argument.

The cluster argument is a formula, or expression, defining the labels for the different subjects in the data. For the dental data we would use

cluster = ~ subject

for both models (1) and (2). The optional argument data specifies the data frame in which the variables used in the model are available. A simple call to lme to fit model (1) would be

> dental.fit1 <- lme(fixed = distance ~ age, random = ~ age,
  +  cluster = ~ subject, data = dental)

To fit model (2) we would use

> dental.fit2 <- lme(fixed = distance ~ age * sex, random = ~ age,
  +  cluster = ~ subject, data = dental)

There are several methods available for the fitted objects of class lme, including those for the generic functions print, summary, and plot.

2.2 The print, summary, and anova methods.

A brief description of the estimation results is returned by the print method. It gives estimates of the standard errors and correlations of the random effects, the cluster variance, and the fixed effects. For the dental.fit1 object we get

> dental.fit1
Call:
  Fixed: distance ~ age
  Random: ~ age
  Cluster: ~ subject
  Data: dental

For example, the output might include:

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>distance</td>
<td>0.2350</td>
<td>0.0120</td>
<td>19.55</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th></th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>0.2090</td>
</tr>
</tbody>
</table>

Covariance parameters:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>age:age</td>
<td>0.0210</td>
</tr>
</tbody>
</table>

Correlation structure:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>age:age</td>
<td>0.820</td>
</tr>
</tbody>
</table>

Estimation method: REML

The summary method provides a more detailed summary of the fitted model, including information about the fixed and random effects, the residual variance, and the fitted values.

> summary(dental.fit1)

For example, the output might include:

Model: distance ~ age + (1 + age | subject)

Number of observations: 112
Number of groups: 28

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>distance</td>
<td>0.2350</td>
<td>0.0120</td>
<td>19.55</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Random effects:

<table>
<thead>
<tr>
<th></th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>0.2090</td>
</tr>
</tbody>
</table>

Covariance parameters:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>age:age</td>
<td>0.0210</td>
</tr>
</tbody>
</table>

Correlation structure:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>age:age</td>
<td>0.820</td>
</tr>
</tbody>
</table>

Residual standard error: 0.209

Estimation method: REML
Variance/Covariance Components Estimates:

Structure: unstructured
Parametrization: logcholesky
Standard Deviation(s) of Random Effect(s)
(Intercept)   age
    2.194103 0.2149245
Correlation of Random Effects
(Intercept)
    age -0.5814881

Cluster Residual Variance: 1.716204

Fixed Effects Estimates:
(Intercept)   age
    16.76111 0.6601852

Number of Observations: 108
Number of Clusters: 27

A more complete description of the estimation results is returned by summary.

> summary(dental.fit2)
...
Loglikelihood: -114.6576
AIC: 245.3152

Variance/Covariance Components Estimates:
Structure: unstructured
Parametrization: logcholesky
Standard Deviation(s) of Random Effect(s)
(Intercept)   age
    2.134464 0.1541247
Correlation of Random Effects
(Intercept)
    age -0.6024329

Cluster Residual Variance: 1.716232

Fixed Effects Estimates:

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Std.Error</th>
<th>z ratio</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>16.3406</td>
<td>0.980057</td>
<td>16.6731</td>
<td>0.00</td>
</tr>
<tr>
<td>age</td>
<td>0.7843</td>
<td>0.082751</td>
<td>9.4786</td>
<td>0.00</td>
</tr>
<tr>
<td>sex</td>
<td>1.0321</td>
<td>1.5354</td>
<td>-0.6721</td>
<td>0.00</td>
</tr>
<tr>
<td>age:sex</td>
<td>-0.3048</td>
<td>0.129647</td>
<td>-2.3512</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Conditional Correlations of Fixed Effects Estimates

<table>
<thead>
<tr>
<th></th>
<th>age</th>
<th>sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.88</td>
<td>-0.63</td>
</tr>
<tr>
<td>age</td>
<td>0.56</td>
<td>-0.63</td>
</tr>
<tr>
<td>sex</td>
<td>0.56</td>
<td>-0.88</td>
</tr>
</tbody>
</table>

The approximate standard errors for the fixed effects are derived using the asymptotic theory described in Pinheiro (1994). The results above indicate that the measurement increases faster in boys than in girls (significant, negative age:sex fixed effect), but the average intercept is common to boys and girls (non-significant sex fixed effect).

A likelihood ratio test to evaluate the hypothesis of no sex differences in distance development is given by the anova method.
> anova(dental.fit1, dental.fit2)

<table>
<thead>
<tr>
<th>Model</th>
<th>Df</th>
<th>AIC</th>
<th>LogLik</th>
<th>Test Lik.Ratio</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>dental.fit1</td>
<td>1 6</td>
<td>252.72</td>
<td>-120.36</td>
<td>1 vs. 2</td>
<td>11.406 0.0033365</td>
</tr>
<tr>
<td>dental.fit2</td>
<td>2 8</td>
<td>245.32</td>
<td>-114.66</td>
<td>1 vs. 2</td>
<td>0.44806 0.50326</td>
</tr>
</tbody>
</table>

The likelihood ratio test strongly rejects the null hypothesis of no sex differences. To use a likelihood test to test if the growth rate only is dependent on sex we fit:

> dental.fit3 <- lme(fixed = distance ~ age + age:sex, random = ~ age, + cluster = ~ subject, data = dental)

and use the anova method again.

> anova(dental.fit2, dental.fit3)

<table>
<thead>
<tr>
<th>Model</th>
<th>Df</th>
<th>AIC</th>
<th>LogLik</th>
<th>Test Lik.Ratio</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>dental.fit2</td>
<td>1 8</td>
<td>245.32</td>
<td>-114.66</td>
<td>1 vs. 2</td>
<td>0.44806 0.50326</td>
</tr>
<tr>
<td>dental.fit3</td>
<td>2 7</td>
<td>243.76</td>
<td>-114.88</td>
<td>1 vs. 2</td>
<td>0.44806 0.50326</td>
</tr>
</tbody>
</table>

As expected, the likelihood ratio test indicates that the initial distances do not depend on sex.

### 2.3 The plot method

Plots of random effects estimates, residuals, and fitted values can be obtained using the plot method for class lme. The following call will produce a scatter plot of the estimated random effects for intercept and slope in model (2), as shown in Figure 2.

> plot(dental.fit2, pch = "o")

The point at the upper left corner of Figure 2 appears to be an outlying value that could have a great impact on the correlation and variance estimates.

Residual plots may be obtained by setting the argument option in the plot method to "r".

> par(mfrow = c(2, 2))
> plot(dental.fit3, option = "r", pch = "o")

The resulting plots are included in Figure 3. The first plot, observed versus fitted values, indicates that the linear model does a reasonable job of explaining the distance growth. The points fall relatively close to the $y = x$ line, indicating a reasonable agreement between the fitted and observed values. The second plot, residuals versus fitted values, suggests the presence of three outliers in the data. The remaining residuals appear to be homogeneously scattered around the $y = 0$ line. The final plot, featuring the boxplot of the residuals by subject, suggests that the outliers occurred for subjects 9 and 13. There seems to be considerable variation in the within-subject variability, but it must be remembered that each of the boxplots represent only four residuals.

Figure 2: Scatter plot of the conditional modes of the intercept and slope random effects in model (2).
Figure 3: Residuals and fitted values plots.

Figure 4 reproduces the original data plot of Figure 1, the random effects estimates scatter plot of Figure 2, and the residual versus fitted values plot of Figure 3, highlighting subjects 9 and 13. We can see from these plots that the two individuals have a large influence on the fit. We also see that the data for subject 9 is probably in error because the measurement decreases substantially between ages 8 and 10 and between ages 12 and 14.

Figure 4: Plot of original data, scatter plot of random effects estimates, and residual versus fitted values plot, with subjects 9 and 13 highlighted.
2.4 Other methods

Standard S methods for extracting components of fitted objects, such as residuals, fitted, and coefficients, can be also be used on lme objects. The first two methods return data frames with two columns, population and cluster, while the last method returns a list with two components, the fixed and the random effects estimates. A more detailed description of these objects is available in the help files, included in Appendix A.

Estimates of the individual parameters are obtained using the `cluster.coef` method.

```r
> cluster.coef(dental.fit3)
(Intercept) age age:sex
1 18.27436 0.8322535 -0.2281247
2 15.46918 0.7357740 -0.2281247
3 16.18725 0.7418871 -0.2281247
   27  19.21336 0.8370006 -0.2281247
```

Predicted values are returned by the `predict` method. For example, if we are interested in predicting the average measurement for both boys and girls at ages 14, 15, and 16, as well as for subjects 1 and 20 at age 13, we should create a new data frame, say `dental.new`, as follows,

```r
> dental.new <-
  + data.frame(sex = c(1, 1, 1, 0, 0, 0, 1, 0),
  + age = c(14, 15, 16, 14, 15, 16, 13, 13),
  + subject = c(NA, NA, NA, NA, NA, NA, 1, 20))
```

and then use

```r
> predict(dental.fit3, ~ subject, dental.new)

   cluster fit.cluster fit.population
1        NA         24.1111
2        NA         24.6361
3        NA         25.1611
4        NA         27.3048
5        NA         28.0579
6        NA         28.8111
7        1 26.12804 23.58611
8       20 25.05424 26.55173
```

to get the cluster and population predictions.

2.5 Structured variance-covariance matrices and random effects blocks

Structured variance-covariance matrices for the random effects can be specified using the optional argument `re.structure` in the call to `lme`. If there are q random effects, the variance-covariance matrix is a q x q symmetric matrix. Predefined structures available include `unstructured` for a general variance-covariance matrix (requiring q(q+1)/2 distinct parameters); `diagonal` for independent random effects with possibly different variances (q parameters); `identity` for independent random effects with the same variance (1 parameter); `compsym` for a compound symmetry structure where all random effects have the same variance and the same correlation (2 parameters); and `ar1` for a common variance and AR(1) correlation structure (Box, Jenkins and Reinsel, 1994, chapter 2) (2 parameters).
The random effects can also be grouped into separate blocks, using the optional argument `re.block`. Random effects belonging to different blocks are assumed to be independent. Different variance-covariance structures can be specified for each block.

We illustrate the use of `re.structure` and `re.block` with an example from radiology. The data consist of repeated measures of mean pixel values from CT scans of the right and the left lymphnodes in the axillary region of 10 dogs over a period of 14 days after application of a contrast. The purpose of the experiment was to model the mean pixel value as a function of time, so as to estimate the time where the maximum mean pixel value was attained. The data are presented in Figure 5.

A second order polynomial seems adequate for these data. Preliminary analyses indicated that the intercept varies with side within dog and the linear term varies with dog, but not with side.

The corresponding linear mixed-effects model is

$$\text{Pixel}_{ijk} = (\beta_0 + b_{0ij}) + (\beta_1 + b_{1i}) \text{Time}_{ijk} + \beta_2 \text{Time}_{ijk}^2 + \varepsilon_{ijk},$$  

where $i$ refers to the dog number (1 through 10), $j$ to the lymphnode side (1 — right, 2 — left), and $k$ refers to time; $\beta_0$, $\beta_1$, and $\beta_2$ denote respectively the intercept, the linear term, and the quadratic term fixed effects; $b_{0ij}$ denotes the intercept random effect (side specific, nested within dog) and $b_{1i}$ denotes the linear term random effect (dog specific); $\varepsilon_{ijk}$ denotes the error term. We assume that the $b_i = [b_{0i1}, b_{0i2}, b_{1i}]^T$ are independent and identically distributed with common distribution $\mathcal{N}(0, \sigma^2 I)$ and the $\varepsilon_{ijk}$ are independent and identically distributed with common distribution $\mathcal{N}(0, \sigma^2)$ and independent of the $b_i$.

A common assumption about the structure of $D$ in model (3) is that the $b_{0ij}$ have equal variance and are independent of the $b_{1i}$. This can be rephrased by saying that the $b_{0ij}$ form a block of random effects with compound symmetry variance-covariance matrix and the $b_{1i}$ form another block with an unstructured variance-covariance matrix. Actually, because all the structured forms for a block are equivalent when the block contains only one random effect, any of the structures could be used. Assuming the data are available in a data frame called `pixel` with columns `dog`, `day`, `side`, and `Pixel` as below

```r
> pixel
dog day side Pixel
1  1  0  r 1045.81
2  1  1  r 1044.54
3  1  2  r 1042.93
4  1  4  r 1050.44
   .
99  9  4  l 1097.22
101  9  8  l 1099.52
102 10  4  l 1132.26
103 10  6  l 1154.51
104 10  8  l 1161.07
```

we can fit model (3) by

```r
> pixel.fit <- lme(fixed = Pixel ~ day + day^2, random = ~ side + day - 1,
                  cluster = ~ dog, data = pixel,
                  re.block = list(c(1,2),3),
                  re.structure = list("compsymm","unstructured"))
```

The `re.block` argument specifies a list which components are vectors or scalars defining the number of the random effects that belong to the block. The number assigned to a random effect is established by the way they are extracted
from the random formula. Alternatively, the names of the random effects can be used in \texttt{re.block}. For example, using \texttt{re.block = list(c("sider","sider"),"day")} is equivalent to the above expression.

The \texttt{re.structure} argument is also a list, with length equal to the number of random effects blocks defined in \texttt{re.block}, when different structures are used, or with length one, when the same structure applies to all random effects blocks. Each component of the \texttt{re.structure} list defines the variance-covariance structure to be used for the corresponding block of random effects in the \texttt{re.block} list. Partial matching is used on this argument, so that only the first letter of the structure’s name is required (e.g. we could have used \texttt{re.structure = list("c","u") above}).

The \texttt{print} method changes slightly when random effects blocks are used.

```r
> pixel.fit
Call:
  Fixed: Pixel ~ day + day^2
  Random: ~ side + day - 1
  Cluster: ~ dog
  Data: pixel

Variance/Covariance Components Estimates:

Block: 1
  Structure: compound symmetry
  Standard Deviation(s) of Random Effect(s)
    side 1 30.2031
    day 1 30.2031
  Correlation of Random Effects
    side 1
    0.6896607

Block: 2
  Structure: identity
  Standard Deviation(s) of Random Effect(s)
    day 1 1.667825

Cluster Residual Variance: 80.42112

Fixed Effects Estimates:
  (Intercept)      day 1(day^2)
  1073.593  6.228684 -0.367346

Number of Observations: 102
Number of Clusters: 10
```

Users can define their own variance-covariance structures if desired. \texttt{lme} uses two generic functions (\texttt{lmemktheta} and \texttt{lmemkFactor}) when estimating the variance-covariance components for the random effects. The first function should produce the vector of parameters that define a given variance-covariance matrix \( D \) and the second one should produce a matrix \( L \) (such that \( L^T L = D \)) from a vector of parameters. The structure defines a class for which methods for the generic functions \texttt{lmemktheta} and \texttt{lmemkFactor} are given. By writing custom methods for these generic functions, users can define special variance-covariance structures.

A technical point that should be observed when writing methods for \texttt{lmemktheta} and \texttt{lmemkFactor} is that the optimization algorithm in \texttt{lme} assumes an unrestricted parametrization for the variance-covariance matrices. This issue may be better illustrated with an example. Suppose that we do not wish to assume independence between \( b_{ij} \) and \( b_{ik} \) in model (3), but rather assume that they have a common, possibly non-zero, correlation. This structure is not among the available options, so we would write the special methods for \texttt{lmemktheta} and \texttt{lmemkFactor}. Letting \( D \) represent the variance-covariance matrix of \( b_i \), with the assumptions that \( D_{11} = D_{22} \) and \( D_{13} = D_{23} \), we note that the positive definiteness of \( D \) requires that \( |D_{12}| < D_{11} \) and \( |D_{13}| < \sqrt{D_{11}D_{33}} \).
These restrictions can be incorporated, in an unconstrained framework, by defining

\[
\begin{align*}
\theta_1 &= \log(D_{11}), \quad \theta_2 = \log\left(\frac{(D_{11} + D_{12})}{(D_{11} - D_{12})}\right) \\
\theta_3 &= \log(D_{33}), \quad \theta_4 = \log\left(\frac{\sqrt{D_{11}D_{33}} + D_{13}}{\sqrt{D_{11}D_{33}} - D_{13}}\right)
\end{align*}
\]

and setting

\[
\begin{align*}
D_{11} &= D_{22} = \exp(\theta_1), \quad D_{12} = \exp(\theta_1)(\exp(\theta_2) - 1)/\left(\exp(\theta_2) + 1\right) \\
D_{33} &= \exp(\theta_3), \quad D_{13} = D_{23} = \exp(\theta_3/2)(\exp(\theta_4) - 1)/\left(\exp(\theta_4) + 1\right)
\end{align*}
\]

Sample methods using this structure are:

```r
lmemktheta.mystuct <- function(D)
{
  ax1 <- log(D[1,1])
  ax2 <- log(D[1,3])
  ax3 <- sqrt(D[1,1] * D[3,3])
  c(ax1, log((D[1,1] + D[1, 2])/(D[1, 1] - D[1, 2])),
     ax2, log((D[1, 3] + ax3)/(ax3 - D[1, 3])))
}

lmemkFactor.mystuct <- function(theta, q)
{
  ax1 <- exp(theta[1])
  ax2 <- ax1 * (exp(theta[2]) - 1)/(exp(theta[2]) + 1)
  ax3 <- exp(theta[3])
  ax4 <- sqrt(ax1 * ax3) * (exp(theta[4]) - 1)/(exp(theta[4]) + 1)
  chol(array(c(ax1, ax2, ax4, ax2, ax1, ax4, ax4, ax3, c(3,3))))
}
```

We can then use them to fit the desired model as below.

```r
> pixel.fit2 <- lme(fixed = Pixel ~ day + day^2, random = ~ side + day - 1,
                   cluster = ~ dog, data = pixel, re.structure = "mystuct")
```

The resulting fit is

```r
> pixel.fit2
Call:
  Fixed: Pixel ~ day + day^2
  Random: ~ side + day - 1
  Cluster: ~ dog
  Data: pixel

Variance/Covariance Components Estimates:

  Structure: mystuct
  Standard Deviation(s) of Random Effect(s)
    side   side   day
    31.46914 31.46914 1.733844
  Correlation of Random Effects
    side   side
    side  0.7135529
    day -0.4723785 -0.4723785

Cluster Residual Variance: 79.71005
Fixed Effects Estimates:
  (Intercept)  day  I(day^2)
  1073.31  6.12583 -0.3664422

Number of Observations: 102
Number of Clusters: 10

We can compare pixel.fit and pixel.fit2 with the anova method as described in section 2.2.

> anova(pixel.fit, pixel.fit2)
Response: Pixel
Pixel
  fixed: (Intercept), day, I(day^2)
  random: side, side, day
  block: list(c(1, 2), 3)
  covariance structure: compound symmetry, identity
Pixel.fit1
  fixed: (Intercept), day, I(day^2)
  random: side, side, day
  block: list(1:3)
  covariance structure: mStruct
Model  Df  AIC  Loglik  Test Lik.Ratio  P value
pixel.fit  1  7  843.77 -414.89
pixel.fit1  2  8  843.34 -413.67  1 vs. 2  2.4275  0.11923

We conclude that the assumption of independence between the $b_{0ij}$ and the $b_{1i}$ is reasonable in this case (P value of 0.12).

3 The `nlme` class and related methods

We will illustrate the use of the functions and methods for the nonlinear mixed-effects model by analyzing some CO$_2$ uptake data shown in Figure 6 and described in Potvin and Lechowicz (1990). These data come from a study of the cold tolerance of a C$_4$ grass species, *Echinochloa crus-galli*. A total of twelve four-week-old plants, six from Québec and six from Mississippi, were divided into two groups: control plants that were kept at 26°C and chilled plants that were subject to 14 h of chilling at 7°C. After 10 h of recovery at 20°C, CO$_2$ uptake rates (in µmol/m$^2$s) were measured for each plant at seven concentrations of ambient CO$_2$ (100, 175, 250, 350, 500, 675, 1000µL/L). Each plant was subjected to the seven concentrations of CO$_2$ in increasing, consecutive order. The objective of the experiment was to evaluate the effect of plant type and chilling treatment on the CO$_2$ uptake. The model used in Potvin and Lechowicz (1990) is

$$U_{ij} = \phi_{1i} \{1 - \exp[-\phi_{2i}(C_j - \phi_{3i})]\} + \varepsilon_{ij},$$

where $U_{ij}$ denotes the CO$_2$ uptake rate of the $i$th plant at the $j$th CO$_2$ ambient concentration; $\phi_{1i}$, $\phi_{2i}$, and $\phi_{3i}$ denote respectively the asymptotic uptake rate, the uptake growth rate, and the maximum ambient CO$_2$ concentration at which no uptake is verified for the $i$th plant; $C_j$ denotes the $j$th ambient CO$_2$ level; and the $\varepsilon_{ij}$ are independent and identically distributed error terms with distribution $N(0, \sigma^2)$.

![Figure 6: CO$_2$ uptake rates (in µmol/m$^2$s) for Québec and Mississippi plants of *Echinochloa crus-galli*, control and chilled at different ambient CO$_2$ concentrations.](image-url)
We arrange the CO₂ uptake data in a data frame called \texttt{co2}, with columns \texttt{plant}, \texttt{type}, \texttt{trt}, \texttt{conc}, and \texttt{uptake} as below.

\begin{tabular}{lrrrr}
\hline
\texttt{plant} & \texttt{type} & \texttt{trt} & \texttt{conc} & \texttt{uptake} \\
\hline
1 & 1 & Quebec & nonchilled & 95 & 16.0 \\
2 & 1 & Quebec & nonchilled & 175 & 30.4 \\
3 & 1 & Quebec & nonchilled & 250 & 34.8 \\
\vdots & & \vdots & \vdots & \vdots & \vdots \\
83 & 12 & Mississippi & chilled & 675 & 18.9 \\
84 & 12 & Mississippi & chilled & 1000 & 19.9 \\
\hline
\end{tabular}

### 3.1 The \texttt{nlme} function

The \texttt{nlme} function is used to fit nonlinear mixed-effects models, as defined in Lindstrom and Bates (1990), using either maximum likelihood or restricted maximum likelihood. Several optional arguments can be used with this function, but a typical call is

\begin{verbatim}
 nlme(model, fixed, random, cluster, data, start)
\end{verbatim}

The \texttt{model} argument is required and consists of a formula specifying the nonlinear model to be fitted. Any S nonlinear formula can be used, giving the function considerable flexibility. From (4) we have that for the CO₂ uptake data this argument is declared as

\begin{equation}
\text{uptake} \sim A \ast (1 - \exp(-B \ast (\text{conc} - C)))
\end{equation}

where we have used the notation $A = \phi_1$, $B = \phi_2$, and $C = \phi_3$. Alternatively, we can define an S function, say \texttt{co2.uptake}, as

\begin{verbatim}
> co2.uptake <- function(A, B, C, conc) A * (1 - exp(-B*(conc - C)))
\end{verbatim}

then write the model argument as

\begin{verbatim}
> model <- formula(uptake ~ co2.uptake(A, B, C, conc))
\end{verbatim}

The advantage of this latter approach is that the analytic derivatives of the model function can be passed to the \texttt{nlme} function as the \texttt{gradient} attribute of the returned value from \texttt{co2.uptake} and used in the optimization algorithm. The S function \texttt{deriv} can be used to create expressions for the derivatives.

\begin{verbatim}
> deriv(uptake ~ A * (1 - exp(-B * (conc - C))), + LETTERS[1:3], function(A, B, C, conc))
\end{verbatim}

If the value returned by the model function does not have a \texttt{gradient} attribute, numerical derivatives are used in the optimization.

The required arguments \texttt{fixed} and \texttt{random} are lists of formulas that define the structures of the fixed and random effects in the model. In these formulas a . on the right hand side of a formula indicates that a single parameter is associated with the effect, but any linear formula in S could be used instead. This gives considerable flexibility to the model, as time-dependent parameters can be easily incorporated (e.g. when a formula in the \texttt{fixed} list involves a covariate that changes with time). Usually every parameter in the model will have an associated fixed effect, but it may, or may not, have an associated random effect. Since we assumed that all random effects have mean zero, the inclusion of a random effect without a corresponding fixed effect would be unusual. Note that the \texttt{fixed} and \texttt{random} formulas could be directly incorporated in the model declaration. The approach used in \texttt{nlme} allows for more efficient calculation of derivatives and will be useful for \texttt{update} methods that will be incorporated in the code in the future.

For the CO₂ uptake data, if we want to fit a model in which all parameters are random and no covariates are included we use

\begin{verbatim}
> fixed = list(A~, B~, C~); random = list(A~, B~, C~)
\end{verbatim}

If we want to estimate the effects of plant type and chilling treatment on the parameters in the model we can use

\begin{verbatim}
> fixed = list(A ~ type*trt, B ~ type*trt, C ~ type*trt), 
> random = list(A ~., B ~., C ~.)
\end{verbatim}
The cluster argument is required and defines the cluster label of each observation. An S expression or a formula with no left hand side can be used here. Data is an optional argument that names a data frame and start provides a list of starting values for the iterative algorithm. Only the fixed effects starting estimates are required. The default starting estimates for the random effects are zero. Starting estimates for the variance-covariance matrix of the random effects ($\mathbf{D}$) and the cluster variance ($\sigma^2$) are automatically generated using a formula from Laird, Lange and Stram (1987) if they are not supplied. Further information on the arguments of \texttt{nlme} is available in the help files in Appendix A.

A simple call to \texttt{nlme} to fit model (4), without any covariates and with all parameters as mixed effects is

```r
> co2.fit1 <-
+ nlme(model = uptake ~ co2.uptake(A, B, C, conc),
+       fixed = list(A ~ .., B ~ .., C ~ .),
+       random = list(A ~ .., B ~ .., C ~ .),
+       cluster = ~ plant, data = CO2,
+       start = list(fixed = c(30, 0.01, 50)))
```

The initial values for the fixed effects were obtained from Potvin and Lechowicz (1990).

### 3.2 Methods for nlme objects

Objects returned by the \texttt{nlme} function are of class \texttt{nlme} which inherits from \texttt{lme}. All methods described in section 2 are available for the \texttt{nlme} class. In fact, with the exception of the \texttt{predict} method, all methods are common to both classes. We illustrate their use here with the CO$_2$ uptake data.

The \texttt{print} method provides a brief description of the estimation results. It gives estimates of the standard errors and correlations of the random effects, of the cluster variance, and of the fixed effects.

```r
> co2.fit1
Call:
  Model: uptake ~ co2.uptake(A, B, C, conc)
Fixed: list(A ~ .., B ~ .., C ~ .)
Random: list(A ~ .., B ~ .., C ~ .)
Cluster: ~ plant
Data: CO2

Variance/Covariance Components Estimates:

  Structure: unstructured
  Parametrization: logcholesky
  Standard Deviation(s) of Random Effect(s)
    A     B     C
    9.510373 0.001152827 11.39466
  Correlation of Random Effects
    A   B
    B -0.06187818
    C 0.99998745 -0.06192643

  Cluster Residual Variance: 3.129989

Fixed Effects Estimates:
  A   B   C
  32.55042 0.00944257 41.61764

Number of Observations: 84
Number of Clusters: 12
```
Note that there is a very strong correlation between the $\phi_1$ and the $\phi_3$ random effects and that these are almost uncorrelated with the $\phi_2$ random effect. The scatter plot matrix of the random effects obtained using the `plot` method

```r
> plot(co2.fit1, pch = "o")
```

is shown in Figure 7. It is clear that the $\phi_1$ and $\phi_3$ random effects are virtually identical. This correlation may be due to the fact that the plant type and the chilling treatment, which were not included in the `co2.fit1` model, are affecting $\phi_1$ and $\phi_3$ in the same way. One of the main advantages of being in the $S$ environment is that all the analytical and graphical machinery present in $S$ is simultaneously available. We can use this to analyze the dependence of the individual parameters $\phi_{1i}$, $\phi_{2i}$, and $\phi_{3i}$ in model (4) on plant type and chilling factor.

Initially we create a data frame with the conditional modes of the random effects obtained in the first fit.

```r
> CO2.random <- data.frame(coef(co2.fit1)$random)
```

Then we add a column to `CO2.random` with the treatment combinations corresponding to each plant.

```r
> CO2.random$type.trt <- as.factor(rep(c("Quebec nonchilled",
+ "Quebec chilled",
+ "Mississippi nonchilled",
+ "Mississippi chilled"), rep(3,4)))
```

Finally we obtain plots of the conditional modes of the random effects versus the treatment combinations. The corresponding plots are presented in Figure 8.

```r
> plot(A ~ type.trt, data = CO2.random)
> plot(B ~ type.trt, data = CO2.random)
> plot(C ~ type.trt, data = CO2.random)
```

These plots indicate that chilled plants tend to have smaller values of $\phi_1$ and $\phi_3$, but the Mississippi plants seem to be much more affected than the Quebec plants, suggesting an interaction effect between plant type and chilling treatment. There is no clear pattern of dependence between $\phi_2$ and the treatment factors, suggesting that this parameter is not significantly affected by either plant type or chilling treatment.

We can then fit a new model in which $\phi_1$ and $\phi_3$ depend on the treatment factors, as below.

```r
> co2.fit2 <- nlme(
+ model =
+ uptake ~ co2.upptake(A, B, C, conc),
+ fixed = list(A ~ type*trt, B ~ .,
+ C ~ type*trt),
+ random = list(A ~ ., B ~ ., C ~ .), cluster = ~ plant, data = CO2,
+ start = list(fixed = c(30, 0, 0, 0.01, 50, 0, 0, 0)))
```

The summary method provides more detailed information on the new fitted object.
> summary(co2.fit2)

Convergence at iteration: 6
Approximate Loglikelihood: -103.5041
AIC: 239.0082

Variance/Covariance Components Estimates:
Structure: unstructured
Parametrization: logcholesky

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation(s) of Random Effect(s)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. (Intercept)</td>
<td>2.276278</td>
<td>0.0003200845</td>
<td>5.981132</td>
</tr>
<tr>
<td>B</td>
<td>-0.008043761</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. (Intercept)</td>
<td>0.999984502</td>
<td>-0.008100170</td>
<td></td>
</tr>
</tbody>
</table>

Cluster Residual Variance: 3.127764

Fixed Effects Estimates:

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A. (Intercept)</td>
<td>32.45210011</td>
<td>0.7225786330</td>
<td>44.911513</td>
</tr>
<tr>
<td>A. type</td>
<td>-7.909764880</td>
<td>0.7024079993</td>
<td>-11.260927</td>
</tr>
<tr>
<td>A. trt</td>
<td>-4.231594577</td>
<td>0.7009980593</td>
<td>-6.036528</td>
</tr>
<tr>
<td>A. type: trt</td>
<td>-2.434420834</td>
<td>0.7010132656</td>
<td>-3.472717</td>
</tr>
<tr>
<td>B</td>
<td>0.009545959</td>
<td>0.0005908485</td>
<td>16.156356</td>
</tr>
<tr>
<td>C. (Intercept)</td>
<td>39.936295607</td>
<td>5.6567839253</td>
<td>7.059894</td>
</tr>
<tr>
<td>C. type</td>
<td>-10.469319722</td>
<td>4.2166574898</td>
<td>-2.482848</td>
</tr>
<tr>
<td>C. trt</td>
<td>-7.975396202</td>
<td>4.1963538181</td>
<td>-1.900554</td>
</tr>
<tr>
<td>C. type: trt</td>
<td>-12.360984497</td>
<td>4.2249903799</td>
<td>-2.925683</td>
</tr>
</tbody>
</table>

The correlation between the $\phi_1$ and the $\phi_3$ random effects remains very high, suggesting that the model is probably overparameterized and fewer random effects are needed. We will not pursue the model building for the CO$_2$ uptake data here, since our main goal is to illustrate the use of the methods for the nlme class and not to present a thorough analysis of the problem.

To compare co2.fit1 and co2.fit2 we use anova.

> anova(co2.fit1, co2.fit2)

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>AIC</th>
<th>Loglik</th>
<th>Test Lik.Ratio</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>co2.fit1</td>
<td>1</td>
<td>10</td>
<td>268.44</td>
<td>-124.22</td>
<td>41.43</td>
</tr>
<tr>
<td>co2.fit2</td>
<td>2</td>
<td>16</td>
<td>239.01</td>
<td>-103.50</td>
<td>2.3824e-07</td>
</tr>
</tbody>
</table>

We see that the inclusion of plant type and chilling treatment in the model caused a substantial increase in the loglikelihood, indicating that they have a significant effect on $\phi_1$ and $\phi_3$.

Diagnostic plots can be obtained with the x option to the plot method

> par(mfrow = c(2,2))
> plot(co2.fit2, option = "r", pch = "o")

The corresponding plot is presented in Figure 9. The first plot, observed versus fitted values, indicates that the model fits the data well — most points lie close to the y = x line. The second plot, residuals versus fitted values, does not indicate any departures from the assumptions in the model — no outliers seem to be present and the residuals are symmetrically scattered around the y = 0 line, with constant spread for different levels of the fitted values.

Predictions are returned by the predict method. For example, to obtain the population predictions of CO$_2$ uptake...
rate for Québec and Mississippi plants under chilling and no chilling, at ambient CO₂ concentrations of 50, 100, 200, and 500 μL/L, we would first define

```r
> CO2.new <- data.frame(
+   type = rep(c('Quebec', 'Mississippi'),
+              c(8, 8)),
+   trt = rep(rep(c('chilled', 'nonchilled'),
+             c(4, 4)), 2),
+   conc = rep(c(50, 100, 200, 500), 4))
```

and then use

```r
> predict(co2.fit2, CO2.new)
```

The `predict` method can also be used for plotting smooth fitted curves by calculating fitted values at closely spaced concentrations. Figure 10 presents the individual fitted curves for all twelve plants evaluated at 200 concentrations between 50 and 1000 μL/L.
3.3 Structured variance-covariance matrices and random effects blocks

Structured variance-covariance matrices and random effects blocks can also be used for nonlinear mixed-effects models. As in the \texttt{lm} function, the optional arguments \texttt{re.structure} and \texttt{re.block} provide this capability. The usage of these arguments is identical to that in \texttt{lm}, described in section 2.5.

4 Future Developments

The classes and methods described here provide tools for analyzing linear and nonlinear mixed-effects models. As they are defined within the S environment, all the powerful analytical and graphical machinery present in S is simultaneously available. The analyses of the dental data, the pixel data, and the CO$_2$ uptake data illustrate some of the available features, but many other features are available.

The code presented here was developed primarily to handle repeated measures data, i.e. data generated by observing a number of clusters repeatedly under varying experimental conditions. More general mixed effects models (e.g. with different levels of nesting) can be analyzed using the functions described here, but the code will not be computationally efficient for that purpose.

There are several directions in which the software can be extended to handle more general mixed-effects models and/or incorporate other estimation techniques. These include, but are not limited to,

- Mixed-effects models in which the variance of the within-subject errors depends on the model function, or some covariate. The current version of the code assumes constant variance for the within-subject errors.

- Mixed-effects models with autocorrelated within-subject errors (Chi and Reinsel, 1989). The current version of the code only handles the independent and identically distributed case.

- More accurate approximations to the loglikelihood in the nonlinear mixed-effects model (Pinheiro and Bates, 1995). These include Laplacian and Gaussian quadrature approximations to the integral that defines the likelihood of the data in the nonlinear mixed-effects model. The current version uses an alternating algorithm suggested by Lindstrom and Bates (1990).

- Profiling methods (Bates and Watts, 1988) for deriving confidence regions on the parameters in the model and assessing the normality of the parameter estimates. These methods are computationally intensive, especially for the nonlinear mixed-effects model, and efficient programming is needed to make their use feasible.

- Update methods for refitting the model when only small changes in the original calling sequence are necessary. These methods are particular useful for model building, when several similar models are fitted sequentially.

- Methods for deriving confidence and prediction intervals for predicted values.

We plan to incorporate all these features in future releases of the software to be contributed to the S collection at StatLib.
References


Appendix A

```r
anova.lme <- function(object, ...) {
  anova.lme(object, ...)  # anova.lme calculates Likelihood Ratio and AIC for lme Objects
}
```

**ARGUMENTS**

- **object**: an object inheriting from class `lme`, usually returned by `lme()` or `nlme()`.
- **...**: other objects inheriting from class `lme`.

**VALUE**

if only one object is declared a data frame with the estimated values, the approximate standard errors, and the z-ratios of the fixed effects is returned. Otherwise a data frame with the degrees of freedom, the loglikelihood, and the AIC of each object is returned. Whenever two consecutive objects have different number of degrees of freedom, a likelihood ratio statistic, with the associated P-value is included in the returned data frame.

**EXAMPLE**

```r
mix.fit.ano <- anova(mix.fit, mix.fit1) # mix.fit and mix.fit1 are lme objects
mix.fit.ano # prints using print.anova.lme method
```

Response: follicles

```r
mix.fit
fixed: (Intercept), sin(2 * pi * time), cos(2 * pi * time)
random: (Intercept), sin(2 * pi * time), cos(2 * pi * time)
block: list(1, 2:3)
covariance structure: identity, diagonal
```

```r
mix.fit1
fixed: (Intercept), sin(2 * pi * time), cos(2 * pi * time)
random: (Intercept)
block: list(1:1)
covariance structure: identity
```

Model Df AIC Loglik Test Lik.Ratio P value
mix.fit 1 7 1073.1 -529.53
mix.fit1 2 5 1108.8 -549.40 1 vs. 2 39.745 2.3416e-09

**SEE ALSO**

`lme`, `lme.object`, `nlme`, `nlme.object`

```r
cluster.coef.lme <- function(object) {
  cluster.coef.lme(object)
}
```

**ARGUMENTS**

- **object**: an object inheriting from class `lme`.

**VALUE**

A matrix containing the cluster parameter estimates (ML or RML estimates of fixed effects plus conditional estimates of random effects). Rows correspond to clusters and columns to parameters.
EXAMPLE

```r
# mix.fit obtained from a call to lme()
cluster.coef.lme(mix.fit)
   (intercept) sin(2 * pi * time) cos(2 * pi * time)
1   15.536312      -1.13172662       -1.4557073
2    8.032558       -0.05075528        0.3347899
3   15.296980      -4.22924696       -1.1782081
4    6.672797      -2.86559654        0.4252792
5  11.456387       -3.66373194       -0.6038166
6  13.954620      -5.29279204       -0.2190056
7  11.219623      -3.05460072       -0.7107337
8  14.203913      -6.68393955       -0.8487961
9  12.545649       -2.91031449       -2.2115794
10 15.664186      -4.83072940       -1.7644216
11   9.475799       -1.56596434       -1.4705484
```

ARGUMENTS

**fixed:** a formula object, specifying the fixed effects part of the model, with the response on the left of a `~` operator, and the terms, separated by `*` operators, on the right. If data is given, all names used in the formula should be defined as variables in the data frame. A model with the intercept as the only random effect can be specified as `~1`.

**random:** a formula object, specifying the random effects part of the model, with the terms, separated by `*` operators, on the right of a `~` operator. This argument has the same general characteristics as fixed, but there will be no left side to the `~expression`. A model with the intercept as the only random effect can be specified as `~1`. **NOTE:** random effects are always assumed to have mean zero. A nonzero mean can be specified by including an identical term in the fixed effects part of the model.

**cluster:** an expression or formula object, specifying the experimental units over which the random effects vary. If `cluster` is given as a formula, it will have no left side to the `~expression`.

**data:** an optional data.frame in which to interpret the variables named in `fixed`, `random` and `cluster`.

**start:** an optional list containing initial values for the scaled variance-covariance matrix of the random effects (D) or the parameter vector (theta) that defines the factorization of D. Theta has to be consistent with the chosen structure and parametrization (see `re.structure` and `paramtr` below).

**est.method:** estimation method; if equal to *"ML"*, Maximum Likelihood is used, otherwise if *"RML"* is specified, Restricted Maximum Likelihood is employed. Partial matching of arguments is used, so only the first character needs to be provided. Default is *"ML"*.

**re.block:** an optional list indicating how the random effects should be blocked. Random effects pertaining to different blocks are assumed to be independent. The random effects can be referenced either by their names, or the order in which they appear in the random formula (NOTE: unless `~1` is used in `random`, an (Intercept) term will be included as the first random effect). Elements in the `re.block` list are vectors containing the names or numbers of the random effects. Within a vector all elements have to be of the same type (i.e. all names or all numbers). By default all random effects are included in the same block.
**re.structure**: an optional character vector describing the variance-covariance structure in each block of the random effects. Predefined structure names are "unstructured" for a general variance-covariance matrix with \( q(q+1)/2 \) parameters (\( q \) = the number of random effects for the block); "diagonal" for independent random effects with possibly different variances (\( q \) parameters); "identity" for independent random effects with the same variance (1 parameter); "compSym" for a compound symmetry structure, random effects with common variance and common covariance (2 parameters); "ar1" for a common variance and AR(1) autocorrelation structure (2 parameters). Users can define their own variance-covariance structure named as a class with methods for the generic functions \lmemKFactor\ and \lmemKtheta\. The generic function \lmemKFactor\ is used to obtain the factorization of \( D \) from the parameter vector \( \theta \), and \lmemKtheta\ is used to obtain \( \theta \) from \( D \) (see \lmemKFactor\.ar1 and \lmemKtheta\.ar1 for examples of the use of the generic functions). If \re.structure\ is of length 1 the name will be used for all blocks. Partial matching against the predefined names is used, so only the first character needs to be provided. Default is "unstructured".

**paramtr**: the parameterization to be used for the unstructured scaled variance-covariance matrices specified in \re.structure\; possible values are "cholesky" for Cholesky decomposition, "logcholesky" for Cholesky using logs of the diagonal elements, "spherical" for spherical coordinates of columns of Cholesky decomposition, "matrixlog" for the matrix logarithm, "householder" for a Householder approach to defining eigenvalues and eigenvectors, and "givens" for a Givens rotation approach to defining eigenvalues and eigenvectors. Partial matching of arguments is used, so only the first character needs to be provided. Default is "logcholesky".

**na.action**: an optional missing-data filter function, applied to the \model.frame\. If the data contain missing values then \na.action\ must be specified.

**max.iter**: an optional maximum number of iterations for convergence algorithm. Default is 50.

**tol**: an optional tolerance for the convergence criterion. Default is 1e-7.

**verbose**: it is passed as the \trace\ argument to \ms()\ (see documentation on that function). Default is F.

**VALUE**

an object of class \lme\ representing the fit. Generic functions such as \print(), plot()\ and \summary()\ have methods to show the results of the fit. See \lme.object\ for the components of the fit. The functions \residuals(), coefficients()\ and \fitted.values()\ can be used to extract some of its components.

**EXAMPLE**

```
# Example from Lindstrom and Bates (1988) J.A.S.A.
# (see references section)
mix.fit <- lme(follicles ~ sin(2*pi*time) + cos(2*pi*time),
           ~sin(2*pi*time)+ cos(2*pi*time), ~ mare,
           ovary.data, est.method="RML",
           re.block= list(1, 2:3),
           re.structure = c("i","d"))
mix.fit
```

**Call:**

Fixed: follicles ~ sin(2 * pi * time) + cos(2 * pi * time)
Random: ~ sin(2 * pi * time) + cos(2 * pi * time)
Cluster: ~ mare
Data: ovary.data

**Variance/Covariance Components Estimates:**

Block: 1
Structure: identity
Standard Deviation(s) of Random Effect(s)
(Intercept)
3.163948
Block: 2
Structure: diagonal
Standard Deviation(s) of Random Effect(s)
\[ \sin(2 \times \pi \times \text{time}) \cos(2 \times \pi \times \text{time}) \]
\[
\begin{array}{ccc}
2.089664 & 1.054145 & \\
2.089664 & 1.054145 & \\
\end{array}
\]
Cluster Residual Variance: 9.12222

Fixed Effects Estimates:
(Intercept) \[ \sin(2 \times \pi \times \text{time}) \cos(2 \times \pi \times \text{time}) \]
\[
\begin{array}{ccc}
12.18717 & -3.298127 & 0.882068 \\
12.18717 & -3.298127 & 0.882068 \\
\end{array}
\]
Number of Observations: 308
Number of Clusters: 11

REFERENCES

<table>
<thead>
<tr>
<th>lme.object</th>
<th>Linear Mixed-effects Model Object</th>
<th>lme.object</th>
</tr>
</thead>
</table>

This class of objects is returned from the \texttt{lme()} function to represent a fitted linear mixed-effects model. Objects of this class have methods for the generic functions \texttt{anova()}, \texttt{coef()}, \texttt{cluster.coef()}, \texttt{fitted()}, \texttt{plot()}, \texttt{predict()}, \texttt{print()}, \texttt{residuals()}, and \texttt{summary()}.  

COMPONENTS
The following components must be included in a legitimate \texttt{lme} object. The residuals, fitted values and coefficients can be extracted by the generic functions of the same name, or by the \texttt{*-S*} operator.

\texttt{coefficients:} a list with two components, \texttt{fixed} and \texttt{random}, where the first is a vector containing the estimated coefficients for the fixed effects - the names of the coefficients are the same as those in the \texttt{fixed} formula of the call to \texttt{lme()}, and the second is a matrix containing the estimated coefficients for the random effects. The columns refer to the parameters in the \texttt{random} formula, and the rows to the \texttt{cluster} levels; the names of the coefficients are the same as in the \texttt{random} formula (columns) and the \texttt{cluster} levels (rows).

\texttt{fitted.values:} a list containing the population and cluster fitted values. The population fitted values are evaluated at the converged estimates of the fixed effects and the mean of the random effects (i.e. the random effects are set to zero). The cluster fitted values are evaluated at the converged estimates of the fixed effects and the conditional estimates of the random effects.

\texttt{residuals:} a list containing the population and cluster residuals from the fit. The population residuals are the observed values minus the population fitted values and the cluster residuals are the observed values minus the cluster fitted values.

\texttt{var.ran:} Random effects variance-covariance-correlation matrix estimate. The variances estimates for the random effects are displayed on the main diagonal, covariances above the diagonal and correlations below the diagonal.

\texttt{var.fix:} the conditional variance-covariance matrix of the fixed effects (i.e. the variance-covariance matrix of the fixed effects given the random effects).

\texttt{sigma:} estimate of the cluster residual standard deviation.

\texttt{call:} a list containing an image of the \texttt{lme()} call.

\texttt{cluster:} a vector containing the clusters levels.
nobs, nclus: the number of observations (nobs) and clusters (nclus) used in the fit.

loglik: the loglikelihood value at convergence.

est.method: estimation method used in the fit, either "ML" or "REML".

fixed, random: formulas representing the fixed and random components of the model.

niter: number of iterations used in iterative algorithm.

re.block, re.structure: the blocking and covariance structures used for the random effects.

paramtr: the parametrization used for the unstructured variance-covariance matrices of the blocks of random effects.

sizetheta: length of the parameter vector theta used to obtain the scaled variance-covariance matrix of the random effects D.

SEE ALSO

nlme

```
nlme <- function(model, fixed, random, cluster,
    data = sys.parent(), start, est.method = c("ML", "REML"),
    re.block, re.structure = c("unstructured", "diagonal",
        "identity", "compasyym", "ar1"), paramtr = c("logcholesky",
        "cholesky", "spherical", "matrixlog", "householder",
        "givens"), control, na.action, na.pattern, verbose = F)
```

ARGUMENTS

model: a nonlinear model formula with the response on the left of a ~ operator, and an expression involving parameters and covariates on the right. If data is given, all names used in the formula should be defined as parameters or variables in the data frame.

fixed: a list of formulas of the parameter fixed effects. Usually there will be a fixed effects formula for each parameter. A . on the right hand side of a formula indicates a single fixed effect for that parameter, a formula is evaluated as a linear formula in the frame (data).

random: a list of formulas of the parameter random effects. There need not be a random effect formula for each parameter. A . on the right hand side of a formula indicates a single random effect for that parameter, a formula is evaluated as a linear formula in the frame (data). NOTE: random effects are always assumed to have mean zero. A nonzero mean can be specified by including an identical term in the fixed effects part of the model.

cluster: an expression or formula object, specifying the experimental units over which the random effects vary. If cluster is given as a formula, it will have no left side to the ~ expression.

data: an optional data.frame in which to interpret the variables named in fixed, random and cluster and covariates referenced in the first two.

start: a list of initial estimates for the parameters. The fixed component is required. The order of initial estimates in that component corresponds to the order in the fixed argument. Optional components are random, D (scaled variance-covariance matrix of the random effects), and theta (the factorized form of the scaled variance-covariance matrix of the random effects). If present, the random component should be a matrix with as many rows as there are random effects and as many columns as there are clusters. Theta has to be consistent with the chosen structure and parametrization (see re.structure and paramtr below).
est.method: estimation method; if equal to "ML", Maximum Likelihood is used, otherwise if "REML" is specified, Residual Maximum Likelihood is employed. Partial matching of arguments is used, so only the first character needs to be provided. Default is "ML".

re.block: an optional list indicating how the random effects should be blocked. Random effects pertaining to different blocks are assumed to be independent. The random effects can be referenced either by their names, or the order in which they appear in the random formula (NOTE: unless a -1 is used in random, an (intercept) term will be included as the first random effect). Elements in the re.block list are vectors containing the names or numbers of the random effects. Within a vector all elements have to be of the same type (i.e. all names or all numbers). By default all random effects are included in the same block.

re.structure: an optional character vector describing the variance-covariance structure in each block of the random effects. Predefined structure names are "unstructured" for a general variance-covariance matrix with q(q+1)/2 parameters (q = the number of random effects for the block); "diagonal" for independent random effects with possibly different variances (q parameters); "identity" for independent random effects with the same variance (1 parameter); "compsym" for a compound symmetry structure, random effects with common variance and common covariance (2 parameters); "ar1" for a common variance and AR(1) autocorrelation structure (2 parameters). Users can define their own variance-covariance structure named as a class with methods for the generic functions lmemkFactor and lmemktheta. The generic function lmemkFactor is used to obtain the factorization of D from the parameter vector theta, and lmemktheta is used to obtain theta from D (see lmemkFactor.ar1 and lmemktheta.ar1 for examples of the use of the generic functions). If re.structure is of length 1 the name will be used for all blocks. Partial matching against the predefined names is used, so only the first character needs to be provided. Default is "unstructured".

parametr: the parameterization to be used for the unstructured scaled variance-covariance matrices specified in re.structure; possible values are "cholesky" for Cholesky decomposition, "logcholesky" for Cholesky using logs of the diagonal elements, "spherical" for spherical coordinates of columns of Cholesky decomposition, "matrixlog" for the matrix logarithm, "householder" for a Householder approach to defining eigenvalues and eigenvectors, and "givens" for a Givens rotation approach to defining eigenvalues and eigenvectors. Partial matching of arguments is used, so only the first character needs to be provided. Default is "logcholesky".

control: an optional list of control values for the nonlinear estimation algorithm.

na.action: an optional missing-data filter function, applied to the model.frame. If the data contain missing values then na.action must be specified.

na.pattern: an optional expression or formula object, specifying which returned values are to be regarded as missing.

verbose: if T (TRUE) information on the evolution of the iterative algorithm (number of iterations, convergence criterion, etc.) will be printed. Default is F.

VALUE
an object of class nlme representing the fit. Generic functions such as print(), plot(), predict, and summary() have methods to show the results of the fit. See nlme.object for the components of the fit. The functions residuals(), coefficients() and fitted.values() can be used to extract some of its components.

SEE ALSO
nlme.control
EXAMPLE

```r
# Example from Lindstrom and Bates (1990) Biometrics
# (see references section)
Orange.fit <- nlme(Circ ~ A/(1 + exp((B - Time)/C)),
                   fixed = list(A ~ ., B ~ ., C ~ .),
                   random = list(A ~ ., B ~ .),
                   cluster = ~ Tree, data = Orange,
                   start = list(fixed = c(160, 700, 350)),
                   re.structure = "c")

Orange.fit

Call:
  Model: Circ ~ A/(1 + exp((B - Time)/C))
  Fixed: list(A ~ ., B ~ ., C ~ .)
  Random: list(A ~ ., B ~ .)
  Cluster: ~ Tree
  Data: Orange

Variance/Covariance Components Estimates:
  Structure: compound symmetry
  Standard Deviation(s) of Random Effect(s)
  A    B
  34.32682 34.32682

Correlation of Random Effects
  A    B
  0.795578

Cluster Residual Variance: 56.68832
```

REFERENCES


```
nlme.control

Arguments:

maxiter: maximum number of iterations during fit. Default is 50.

pnls.maxiter: maximum number of iterations during penalized nonlinear least squares (PNLS) step. Default is 7.

lme.maxiter: maximum number of iterations during linear mixed-effects (LME) step. Default is 50.

minscale: minimum factor by which to shrink the default step size in an attempt to decrease the sum of squares in the PNLS step. Default 0.001.

tolerance: tolerance for the convergence criterion in the algorithm. Default is 0.001.

pnls.tolerance: tolerance for the convergence criterion in the PNLS step. Default is 0.001.
```
lme.tolerance: tolerance for the convergence criterion in the LME step, passed as the rel.tolerance argument to ms() (see documentation on that function). Default is 1e-6.

return.object: if T (TRUE) the fitted object is returned by nlme() when the maximum number of iterations is attained without convergence. Default is F.

VALUE

a list containing components for each of the possible arguments, either the value supplied by the user or the default.

EXAMPLE

# increase the maximum number of iterations in the PNLS step
# and return the object when maximum number of iterations is
# attained without convergence
nlme.control(pnls.maxiter = 20, return.object = T)

<table>
<thead>
<tr>
<th>nlme.object</th>
<th>Nonlinear Mixed-effects Model Object</th>
<th>nlme.object</th>
</tr>
</thead>
</table>

This class of objects is returned from the nlme() function to represent a fitted nonlinear mixed-effects model. Objects of this class have methods for the generic functions anova(), coef(), cluster.coef(), fitted(), plot(), predict(), print(), residuals(), and summary(). Objects from this class inherit from the lme class.

COMPONENTS

The following components must be included in a legitimate nlme object. The residuals, fitted values and coefficients can be extracted by the generic functions of the same name, or by the ‘$’ operator.

coefficients: a list with two components, fixed and random, where the first is a vector containing the estimated coefficients for the fixed effects and the second is a matrix containing the estimated coefficients for the random effects. The columns refer to the parameters in the random formula, and the rows to the cluster levels; the names of the coefficients are the same as those in the fixed or random formulas when just a . is used on the right hand side, or the original name with the covariate name appended to it (i.e. name.covariate) otherwise.

fitted.values: a list containing the population and cluster fitted values. The population fitted values are evaluated at the converged estimates of the fixed effects and the mean of the random effects (i.e. the random effects are set to zero). The cluster fitted values are evaluated at the converged estimates of the fixed effects and the conditional estimates of the random effects.

residuals: a list containing the population and cluster residuals from the fit. The population residuals are the observed values minus the population fitted values and the cluster residuals are the observed values minus the cluster fitted values.

var.ran: Random effects variance-covariance-correlation matrix estimate. The variances estimates for the random effects are displayed on the main diagonal, covariances above the diagonal and correlations below the diagonal.

var.fix: the conditional variance-covariance matrix of the fixed effects (i.e. the variance-covariance matrix of the fixed effects given the random effects).

sigma: estimate of the cluster residual standard deviation.

loglik: the approximate log likelihood at convergence.

cluster: a vector containing the clusters labels.
call: a list containing an image of the \texttt{nlme()} call.

\texttt{nobs, nclus:} the number of observations (nobs) and clusters (ncclus) used in the fit.

\texttt{est.method:} estimation method used in the fit, either "ML" or "RML".

\texttt{niter:} number of iterations used in iterative algorithm.

\texttt{re.block, re.structure:} the blocking and covariance structures used for the random effects.

\texttt{parametr:} the parametrization used for the unstructured variance-covariance matrices of the blocks of random effects.

\texttt{sizetheta:} length of the parameter vector theta used to obtain the scaled variance-covariance matrix of the random effects D.

\texttt{criterion:} a vector containing the values of the convergence criteria for the fixed and random effects and the factor of the scaled covariance matrix of the random effects.

\texttt{SEE ALSO}

\texttt{nlme}

\begin{verbatim}
plot.lme

\texttt{plot.lme(object, option = c("c","r","s"), resid.type = c("c","p"),
    which = 1:3, ...)}

ARGUMENTS

\texttt{object:} an object inheriting from class \texttt{lme}.

\texttt{option:} if "r" or "s", diagnostic plots for the fitted values and the raw ("r") or standardized ("s") residuals are returned; otherwise, if option is set to "c", the function returns plots for the random effects coefficients. Default is "c".

\texttt{resid.type:} if "c", the cluster residuals are used, otherwise, if "p", the population residuals are used. Default is "r". See the help file for \texttt{lme.object} for definitions of cluster and population residuals and fitted values.

\texttt{which:} an optional integer vector specifying which plots should be generated when option is set to "r" or "s". Available options include observed versus fitted values (1), residuals versus fitted values (2), and residuals by cluster (3). Default is 1:3 (all plots generated).

\texttt{...:} optional plot settings arguments.

VALUE

For either the raw residuals or the standardized residuals, three diagnostic plots are returned: \texttt{response vs. fitted.values}, \texttt{residuals vs. fitted.values} and \texttt{residuals vs. cluster}. For the random effects coefficients, if just one coefficient is present in the model, the normal plot of the coefficients is returned; if two coefficients are used, a scatter plot is returned; otherwise a scatter plot matrix with all pairwise combinations of the random coefficients is returned.

\texttt{EXAMPLE}

\begin{verbatim}
> plot.lme(mix.fit)
> plot.lme(mix.fit,"s","p")
\end{verbatim}
\end{verbatim}
predict.lme

Make Predictions from a Fitted lme Object

predict.lme(object, cluster, data, na.action)

ARGUMENTS

object: an object inheriting from class lme.

cluster: an optional expression or formula object, specifying the cluster associated with each observation for which the prediction is required; if cluster is given as a formula, there should be no left hand side to the ~expression. If cluster is not specified only the population predictions are calculated. Missing values are allowed for cluster, in which case only the population prediction is calculated and the cluster prediction is set to NA. See the help file for lme.object or nlme.object for definitions of cluster and population residuals and fitted values.

data: a data.frame containing the values at which predictions are required. Only those predictors, referred to in the right side of the fixed and random formulas in object, need be present by name in data.

na.action: an optional missing-data filter function, applied to the model.frame. If the data contain missing values the na.action must be specified.

VALUE

da data.frame containing either two components: cluster - the clusters associated with each prediction and fit - a data.frame with columns population - the predictions based solely on the fixed effects estimates and cluster - the predictions using both the fixed effects estimates and the posterior modes of the random effects, when cluster is specified or only the population column, otherwise.

EXAMPLE

# defining the new data frame for predictions
ovary.new <-
data.frame(time = c(0, 0.25, 0.5, 0.75, 1),
mare = c(1, 1, 2, 2, 3))
predict.lme(mix.fit, ~mare, ovary.new)
cluster fit.cluster fit.population
1 1 14.080605 11.305098
2 1 14.404586 8.889039
3 2 7.697768 13.069234
4 2 8.083313 15.485293
5 3 14.118772 11.305098

SEE ALSO

lme, lme.object

print.lme

Use print() on a lme Object

print.lme(x, ...)

This function is a method for the generic function print() for class "lme". It can be invoked by calling print(x) for an object x of the appropriate class, or directly by calling print.lme(), regardless of the class of the object. See print or print.default for the general behavior of this function and for the interpretation of x.
summary.lme

**Summary of a Linear Mixed-Effects Model**

`summary.lme(object, re = T)`

**Arguments**
- `object`: an object inheriting from class `lme()`.
- `re`: if `T` (TRUE), the conditional modes of the random effects are included in the returned object. Default is `T`.

**Value**
A structure with the appropriate summary information. The components include `correlations`, `fixed.table`, `random.table`, `call`, `nobs`, `nclust`, `loglik`, `AIC`, `niter`, `sigma`, `est.method`, `paramtr`, `residuals`, and `residuals.type`. The `AIC` component gives the Akaike Information Criteria, defined as `-2*loglikelihood + 2*(number of parameters in the model)`. The object returned is of class `summary.lme`. A print method is available for this class. Note: the `z` ratio in the fixed.table component is included simply as a guideline for screening variables. The recommended approach for deleting variables from the model is to refit the reduced model and test the significance of the reduction in the loglikelihood, when compared to the original fit.

**Example**
```r
mix.fit.summ <- summary(mix.fit)  # mix.fit is a lme object
mix.fit.summ  # prints using print.summary.lme method
Call:
  Fixed: follicles ~ sin(2 * pi * time) + cos(2 * pi * time)
  Random:  ~ sin(2 * pi * time) + cos(2 * pi * time)
Cluster:  ~ mare
  Data: ovari.data
Estimation Method: RML
Convergence at iteration: 7
Restricted Loglikelihood: -629.5315
Restricted AIC: 1073.063
Variance/Covariance Components Estimates:
  Block: 1
    Structure: identity
    Standard Deviation(s) of Random Effect(s)
    (Intercept)
      3.163948
  Block: 2
    Structure: diagonal
    Standard Deviation(s) of Random Effect(s)
    sin(2 * pi * time)  cos(2 * pi * time)
      2.089664  1.054145
Cluster Residual Variance: 9.12222
Fixed Effects Estimates:

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Approx. Std.Error</th>
<th>z ratio (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>12.187166</td>
<td>0.9707129</td>
<td>12.554862</td>
</tr>
<tr>
<td>sin(2 * pi * time)</td>
<td>-3.298127</td>
<td>0.6805604</td>
<td>-4.846193</td>
</tr>
<tr>
<td>cos(2 * pi * time)</td>
<td>-0.882068</td>
<td>0.3992836</td>
<td>-2.209127</td>
</tr>
</tbody>
</table>
Conditional Correlations of Fixed Effects Estimates

<table>
<thead>
<tr>
<th>(Intercept)</th>
<th>sin(2 * pi * time)</th>
<th>cos(2 * pi * time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sin(2 * pi * time)</td>
<td>-4.050516e-17</td>
<td>3.125134e-02</td>
</tr>
<tr>
<td>cos(2 * pi * time)</td>
<td>4.115554e-17</td>
<td>3.125134e-02</td>
</tr>
</tbody>
</table>
```
Random Effects (Conditional Estimates):

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>sin(2 * pi * time)</th>
<th>cos(2 * pi * time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.3491465</td>
<td>2.1664005</td>
<td>-0.57363934</td>
</tr>
<tr>
<td>2</td>
<td>-4.1546083</td>
<td>3.2473718</td>
<td>1.21685789</td>
</tr>
<tr>
<td>3</td>
<td>3.1098142</td>
<td>-0.9311199</td>
<td>-0.29614018</td>
</tr>
<tr>
<td>4</td>
<td>-5.5143693</td>
<td>0.4325305</td>
<td>1.30734712</td>
</tr>
<tr>
<td>5</td>
<td>-0.7307790</td>
<td>-0.3656049</td>
<td>0.27825137</td>
</tr>
<tr>
<td>6</td>
<td>1.7674541</td>
<td>-1.9946650</td>
<td>0.66306239</td>
</tr>
<tr>
<td>7</td>
<td>-0.9675429</td>
<td>0.2435264</td>
<td>0.17133429</td>
</tr>
<tr>
<td>8</td>
<td>2.0167471</td>
<td>-3.3858125</td>
<td>0.03327185</td>
</tr>
<tr>
<td>9</td>
<td>0.3584834</td>
<td>0.3878126</td>
<td>-1.32951141</td>
</tr>
<tr>
<td>10</td>
<td>3.4770206</td>
<td>-1.5326023</td>
<td>-0.88235360</td>
</tr>
<tr>
<td>11</td>
<td>-2.7113664</td>
<td>1.7321627</td>
<td>-0.58848038</td>
</tr>
</tbody>
</table>

Standardized Cluster Residuals:

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Q1</th>
<th>Med</th>
<th>Q3</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.688067</td>
<td>-0.5996717</td>
<td>-0.03194583</td>
<td>0.5487485</td>
<td>4.267131</td>
</tr>
</tbody>
</table>

Number of Observations: 308
Number of Clusters: 11

SEE ALSO

lme, lme.object, print.lme