Modeling Residual Covariance Structure

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

GCM, 2010
Introduction

A Composite Growth Curve Model for Cognitive Performance

Deriving the Residual Covariance Structure

Modeling the Residual Covariance Structure

Which Residual Structure?

Some Common Covariance Structures

Fixed Effects Modeling of Composite Residual Structure with R

Mixed Effects Modeling with Nonstandard Residual Covariance Structure

Modeling Residual Covariance Structure

- Introduction
- A Composite Growth Curve Model for Cognitive Performance
- Deriving the Residual Covariance Structure
  - The Residual Vector
  - Variance of a Residual
  - Covariance of Two Composite Residuals
  - Block-Diagonal Covariance Matrix
- Modeling the Residual Covariance Structure
  - Error Covariance Structure in Matrix Format
- Which Residual Structure?
  - Fixed Effects Only
  - Fixed and Random Effects — Two Modeling Choices
- Some Common Covariance Structures
  - Unstructured
  - Compound Symmetry
  - Heterogeneous Compound Symmetry
  - Autoregressive
  - Heterogeneous Autoregressive
  - Toeplitz
- Fixed Effects Modeling of Composite Residual Structure with R
- Mixed Effects Modeling with Nonstandard Residual Covariance Structure
Introduction

In this module, we examine the implications of linear combination theory for the modeling of the residual covariance structure in growth curve modeling. We discover that there are a number of possible forms for this covariance structure, and these forms require differing numbers of degrees of freedom to model. Consequently, there is a possibility that a more “compact” model may be able to account for our growth curve data.
Our discussion in this section will be built around a particular example and data set:

1. Willett (1988) examined cognitive performance on a “opposites naming” task over the course of 4 weeks.
2. In this time-structured data set, 35 people completed an inventory once every week.
3. In addition, at wave 1, each person also completed a test of general cognitive ability (COG), which (after centering) is used as a level-2 predictor to predict both slopes and intercepts.

```r
> data <- read.table("opposites_pp.txt",header=T,sep="",")
> attach(data)
> options(digits=9)
```
A Model for Cognitive Performance

Denoting $X_i = COG_i - \overline{COG}$, and $T_i = TIME_i$ to simplify the notation, the standard multilevel model is

\[
Y_{ij} = \pi_{0i} + \pi_{1i} T_j + \epsilon_{ij} \quad (1)
\]
\[
\pi_{0i} = \gamma_{00} + \gamma_{01} X_i + \zeta_{0i} \quad (2)
\]
\[
\pi_{1i} = \gamma_{10} + \gamma_{11} X_i + \zeta_{1i} \quad (3)
\]

where

\[
\epsilon_{ij} \sim iid N(0, \sigma_\epsilon^2) \quad (4)
\]

and

\[
\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim iid N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right) \quad (5)
\]
Combining Equations 1 – 5, we get the composite model

\[ Y_{ij} = \gamma_{00} + \gamma_{10} T_j + \gamma_{01} X_i + \gamma_{11} X_i \times T_j + r_{ij} \]  

(6)

where the composite residual \( r_{ij} \) is

\[ r_{ij} = \epsilon_{ij} + \zeta_{0i} + \zeta_{1i} T_j \]  

(7)

In using Equation 7 above, we will need to remember that, for given \( i \) and/or \( j \), \( T_j \) is a constant while the \( \epsilon_{ij}, \zeta_{0i}, \) and \( \zeta_{1i} \) are random variables.
The Residual Vector

Suppose we were to list the $Y_{ij}$ in order in a vector $\mathbf{y}$.

There would be a corresponding vector $\mathbf{r}$ containing the residuals.

Since these residuals are random variables, they have a multivariate distribution, and we can derive the residual variance-covariance matrix using the standard rules for linear combinations. For simplicity, suppose there were just 2 people, and therefore only 8 observations. The vector $\mathbf{r}$ would look like this:
The Residual Vector

\[ \mathbf{r} = \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{14} \\ r_{21} \\ r_{22} \\ r_{23} \\ r_{24} \end{bmatrix} = \begin{bmatrix} \epsilon_{11} + \zeta_{01} + \zeta_{11} T_1 \\ \epsilon_{12} + \zeta_{01} + \zeta_{11} T_2 \\ \epsilon_{13} + \zeta_{01} + \zeta_{11} T_3 \\ \epsilon_{14} + \zeta_{01} + \zeta_{11} T_4 \\ \epsilon_{21} + \zeta_{02} + \zeta_{12} T_1 \\ \epsilon_{22} + \zeta_{02} + \zeta_{12} T_2 \\ \epsilon_{23} + \zeta_{02} + \zeta_{12} T_3 \\ \epsilon_{24} + \zeta_{02} + \zeta_{12} T_4 \end{bmatrix} \] (8)
Consider the residual $r_{ij} = \epsilon_{ij} + \zeta_{0i} + \zeta_{1i} T_j$. What is its variance? To answer this question, we simply apply the heuristic rule, remembering that for a given $j$, $T_j$ is a constant. We could begin by squaring the expression. However, recall that our beginning assumption is that every $\epsilon_{ij}$ is independent of everything, including any other $\epsilon$. So

$$\text{Var}(r_{ij}) = \text{Var}(\epsilon_{ij}) + \text{Var}(\zeta_{0i} + \zeta_{1i} T_j)$$

$$= \sigma^2_\epsilon + \text{Var}(\zeta_{0i} + T_j \zeta_{1i})$$

$$= \sigma^2_\epsilon + \text{Var}(\zeta_{0i}) + \text{Var}(T_j \zeta_{1i}) + 2 \text{Cov}(\zeta_{0i}, T_j \zeta_{1i})$$

$$= \sigma^2_\epsilon + \text{Var}(\zeta_{0i}) + T_j^2 \text{Var}(\zeta_{1i}) + 2 T_j \text{Cov}(\zeta_{0i}, \zeta_{1i})$$

$$= \sigma^2_\epsilon + \sigma^2_0 + T_j^2 \sigma^2_1 + 2 T_j \sigma_{01}$$

(9)
The Covariance of an Individual’s Residuals
Uncorrelated $\epsilon_{ij}$

For different individuals, none of the individual constituents of the $r_{ij}$ are correlated, so residuals across individuals must have zero covariance. However, for a given individual, the residuals will be correlated. Let’s derive the covariance for two residuals at different times on the same individual. Again since $\epsilon_{ij}$ and $\epsilon_{ij'}$ are independent of each other and everything else, they cannot contribute to covariance, so we can simplify the calculation by eliminating them before applying the heuristic rule

\[
\text{Cov}(r_{ij}, r_{ij'}) = \text{Cov}(\zeta_{0i} + \zeta_{1i} T_j, \zeta_{0i} + \zeta_{1i} T_{j'}) \\
= \text{Cov}(\zeta_{0i}, \zeta_{0i}) + \text{Cov}(\zeta_{1i} T_j, \zeta_{0i}) + \text{Cov}(\zeta_{0i}, \zeta_{1i} T_{j'}) + \text{Cov}(\zeta_{1i} T_j, \zeta_{1i} T_{j'}) \\
= \text{Var}(\zeta_{0i}) + T_j \text{Cov}(\zeta_{1i}, \zeta_{0i}) + T_{j'} \text{Cov}(\zeta_{0i}, \zeta_{1i}) + T_j T_{j'} \text{Cov}(\zeta_{1i}, \zeta_{1i}) \\
= \text{Var}(\zeta_{0i}) + T_j \text{Cov}(\zeta_{0i}, \zeta_{1i}) + T_{j'} \text{Cov}(\zeta_{0i}, \zeta_{1i}) + T_j T_{j'} \text{Var}(\zeta_{1i}) \\
= \sigma_0^2 + (T_j + T_{j'})\sigma_{01} + T_j T_{j'}\sigma_1^2 \\
= \sigma_0^2 + (T_j + T_{j'})\sigma_{01} + T_j T_{j'}\sigma_1^2 \\
= \sigma_0^2 + (T_j + T_{j'})\sigma_{01} + T_j T_{j'}\sigma_1^2 \\
(10)
\]
The Covariance of an Individual’s Residuals

Correlated $\epsilon_{ij}$

So far, we have assumed that the within-subject residuals are uncorrelated across time. However, there are excellent reasons to believe that frequently this will not be the case.

So, suppose, for a given individual $i$, $\text{Cov}(\epsilon_{ij}, \epsilon_{ij'}) \neq 0$.

What will be the new equation for $\text{Cov}(r_{ij}, r_{ij'})$? (C.P.)

$$\text{Cov}(r_{ij}, r_{ij'}) = \sigma_0^2 + (T_j + T_{j'})\sigma_{01} + T_j T_{j'}\sigma_1^2 + ?? \quad (11)$$
The Covariance of an Individual’s Residuals

Correlated $\epsilon_{ij}$

So far, we have assumed that the within-subject residuals are uncorrelated across time. However, there are excellent reasons to believe that frequently this will not be the case.

So, suppose, for a given individual $i$, $\text{Cov}(\epsilon_{ij}, \epsilon_{ij'}) \neq 0$.

What will be the new equation for $\text{Cov}(r_{ij}, r_{ij'})$? (C.P.)

$$\text{Cov}(r_{ij}, r_{ij'}) = \sigma_0^2 + (T_j + T_{j'})\sigma_{01} + T_j T_{j'}\sigma_1^2 + \text{Cov}(\epsilon_{ij}, \epsilon_{ij'})$$ (12)
As shown in Equations 7.9 and 7.10 on page 250 of Singer and Willett, the vector of residuals will have a covariance matrix that is typically referred to as *block diagonal*. For the first 8 observations, for example, the covariance matrix will look like this:

\[
\begin{bmatrix}
\sigma^2_{r_{11}} & \sigma_{r_{11}, r_{12}} & \sigma_{r_{11}, r_{13}} & \sigma_{r_{11}, r_{14}} & 0 & 0 & 0 & 0 \\
\sigma_{r_{12}, r_{11}} & \sigma^2_{r_{12}} & \sigma_{r_{12}, r_{13}} & \sigma_{r_{12}, r_{14}} & 0 & 0 & 0 & 0 \\
\sigma_{r_{13}, r_{11}} & \sigma_{r_{13}, r_{12}} & \sigma^2_{r_{13}} & \sigma_{r_{13}, r_{14}} & 0 & 0 & 0 & 0 \\
\sigma_{r_{14}, r_{11}} & \sigma_{r_{14}, r_{12}} & \sigma_{r_{14}, r_{13}} & \sigma^2_{r_{14}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \sigma^2_{r_{21}} & \sigma_{r_{21}, r_{22}} & \sigma_{r_{21}, r_{23}} & \sigma_{r_{21}, r_{24}} \\
0 & 0 & 0 & 0 & \sigma_{r_{22}, r_{21}} & \sigma^2_{r_{22}} & \sigma_{r_{22}, r_{23}} & \sigma_{r_{22}, r_{24}} \\
0 & 0 & 0 & 0 & \sigma_{r_{23}, r_{21}} & \sigma_{r_{23}, r_{22}} & \sigma^2_{r_{23}} & \sigma_{r_{23}, r_{24}} \\
0 & 0 & 0 & 0 & \sigma_{r_{24}, r_{21}} & \sigma_{r_{24}, r_{22}} & \sigma_{r_{24}, r_{23}} & \sigma^2_{r_{24}} \\
\end{bmatrix}
\]
Singer and Willett refer to the $4 \times 4$ block representing the within-subject covariance matrix of composite residuals as $\Sigma_r$.

Under the assumptions of the model, the compound residuals in the mixed model have a particular structural form given by Equation 12.

The simplest version, which Singer and Willett refer to as the “standard” version, is given by Equation 10.

As we have seen, this “standard” structure is a rather complicated function of model parameters and the coded values of time.

Singer and Willett then go on to discuss other ways of directly modeling the covariance structure of the composite residuals.
Recall how the model for the $i$th individual can be expressed in matrix format as

$$
y_i = X_i \beta + Z_i b_i + \epsilon_i = X_i \beta + r_i
$$

$$
|b_i| \sim N(0, \Psi), \quad |\epsilon_i| \sim N(0, \sigma^2 \Lambda_i)
$$

where $X_i$ is the fixed effects regressor matrix for the $i$th person, and $Z_i$ is the random effects regressor matrix, which usually contains a subset (perhaps all) of the columns of $X_i$. The vector $\beta$ contains fixed effects, while $b_i$ contains the random effects. Note that the general form has the covariance matrix of the $\epsilon_i$ as $\sigma^2 \Lambda_i$, where $\Lambda_i$ is a correlation matrix, while the “standard assumption” has $\Lambda = I$. 

James H. Steiger
From standard matrix algebra, since \( r_i = Z_i b_i + \epsilon_i \), we have

\[
\text{Cov}(r_i) = Z_i \Psi Z_i' + \sigma^2 \Lambda_i \tag{14}
\]

The standard assumption of uncorrelated within-subject errors (the \( \epsilon_{ij} \)) gives

\[
\text{Cov}(r_i) = Z_i \Psi Z_i' + \sigma^2 I \tag{15}
\]
Example Computation

For example, consider the result of fitting the model on page 247, using, as Singer and Willett did, REML.

```R
> library(lme4)
> fit.1 <- lmer( OPP ~ TIME + CCOG + TIME:CCOG + (1+TIME|ID))
> fit.1
Linear mixed model fit by REML
Formula: OPP ~ TIME + CCOG + TIME:CCOG + (1 + TIME | ID)
AIC  BIC  logLik  deviance
1276.28 1299.82 -630.142  1266.96 1260.28
Random effects:
Groups Name Variance Std.Dev. Corr
ID (Intercept) 1236.4132 35.16267
TIME 107.2492 10.35612 -0.522
Residual 159.4771 12.62843
Number of obs: 140, groups: ID, 35

Fixed effects:
(Intercept) 164.37429  6.206096 26.48594
TIME  26.95998  1.993950 13.52089
CCOG -0.11355  0.504012 -0.22530
TIME:CCOG  0.432858  0.161933  2.67306

Correlation of Fixed Effects:
(Intr) TIME  CCOG
TIME -0.522
CCOG 0.000  0.000
TIME:CCOG 0.000  0.000 -0.522
```
Example Computation

We can get the estimated $\Psi$ with a bit more precision with the `VarCorr` function:

```r
> VarCorr(fit.1)

$ID

 (Intercept)   TIME
(Intercept) 1236.413173 -178.233362
TIME -178.233362  107.249200
attr("stddev")

 (Intercept)   TIME
35.1626673  10.3561190
attr("correlation")

 (Intercept)   TIME
(Intercept) 1.000000000 -0.489452056
TIME -0.489452056  1.000000000

attr("sc")
sigmaREML
12.6284257

> Psi <- matrix(c(1236.413173, -178.233362,-178.233362,107.2492),2,2)
```
Since the data are time-structured, all the $Z_i$ are the same, i.e.,

$$Z = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad (16)$$

```r
> Z <- matrix(c(1,1,1,1,0,1,2,3),4,2)
> sigma <- 12.6284257
> I <- diag(1,4,4)
```
Example Computation

We can compute the covariance matrix of the composite residuals in R easily as

```r
> cov.r <- Z %*% Psi %*% t(Z) + sigma^2 * I
> cov.r
```

```
[1,] 1395.890309 1058.179811  879.946449  701.713087  771.732627  692.664773
[2,] 1058.179811 1146.672785  916.211487  845.227325  916.211487 1058.179811
[3,]  879.946449  916.211487 1111.953661  988.741563 1058.179811  916.211487
[4,]  701.713087  845.227325  988.741563 1291.732937  771.732627  916.211487
```
For ease of comparison with Equation 7.14 on page 255 of Willett and Singer, we can round to one decimal place. There are some discrepancies.

```r
> Singer.Willett.7.14 <- matrix(c(1395.9,1058.2,880,701.7,1058.2,1146.8,916.2,845.2,880,916.2,1112.3,988.8,701.7,845.2,988.8,1294.4),4,4)
> round(Singer.Willett.7.14 - cov.r,1)

[1,] 0.0 0.0 0.1 0.0
[2,] 0.0 0.1 0.0 0.0
[3,] 0.1 0.0 0.3 0.1
[4,] 0.0 0.0 0.1 2.7
```

The discrepancies are probably attributable to rounding differences in most cases, but careful tracing of the calculation of element (4,4) of the matrix on page 252 shows what appears to be a “digit transfer” in their rounded computation. Specifically, using the rounded values on page 252, one obtains 1292.4, rather than the printed value of 1294.4.
In modeling the covariance structure of residuals, one has several choices which can and should be motivated by both theoretical and practical concerns.

Let’s review some of the major choices.
If the model has only fixed effects, then one may write

\[ y_i = X_i \beta + \epsilon_i \]  

(17)

In this case, the classic assumption is that

\[ \text{Cov}(y_i) = \text{Cov}(\epsilon_i) = \sigma^2 I. \]

A more relaxed assumption is that

\[ \text{Cov}(y_i) = \text{Cov}(\epsilon_i) = \sigma^2 \Lambda_i \]

In either case, the definition of the “residual” is clear, it is the \( \epsilon_i \) term.
If the model includes fixed and random effects, then the covariance matrix of the $y_i$ is determined by both the random effects term $Z_i b_i$ and the error term $\epsilon_i$, as shown above in Equation 14.
Singer and Willett present the scalar algebra equivalent of Equation 14, and call it the “standard structure” for the covariance matrix of the “composite residual” $r_i = Z_i b_i + \epsilon_i$. They then propose to go on and model other structures for the covariance matrix of $r_i$. It is important to realize that when modeling “other structures,” you have, more or less, dispensed with the random effects term, and are no longer fitting a random effects model. You are now, in fact, fitting the model of Equation 13.
Modeling the Composite Residual

This choice may be reasonable in some contexts. However, if you are comparing the “standard structure” of Equation 15 with some other structure, the models are not nested, because you have, by directly altering the covariance structure of the $r_i$ instead of the $\epsilon_i$, implicitly, wiped out the random effects term and simultaneously changed the model for the covariance structure of the $\epsilon_i$. 

James H. Steiger
In a model with both fixed and random effects, and alternative to modeling the covariance structure of the composite residual is to model the covariance structure of the $\epsilon_i$, that is, model the structure of the matrix $\Lambda_i$. This option is not discussed by Singer and Willett in their Chapter 7 treatment.
Pinheiro and Bates (2000), in their book “Mixed-Effects Models in S and S-Plus,” delineate carefully between the options of

1. Dropping the random effects contribution and modeling the covariance structure of the \( r_i \), using the `gls` function, and

2. Keeping the random effects contribution and modeling the covariance structure of the \( \epsilon_i \), using the `lme` function
Choosing between the Two Residual Modeling Options

The choice between an \texttt{lme} model and a \texttt{gls} model should take into account more than just information criteria and likelihood tests. A mixed-effects model has a hierarchical structure which, in many applications, provides a more intuitive way of accounting for within-group dependency than the direct modeling of the marginal variance-covariance structure of the response in the \texttt{gls} approach. Furthermore, the mixed-effects estimation gives, as a byproduct, estimates for the random effects, which may be of interest in themselves. The \texttt{gls} model focuses on marginal inference and is more appealing when a hierarchical structure for the data is not believed to be present, or is not relevant in the analysis, and one is more interested in parameters associated with the error variance-covariance structure, as in time-series analysis and spatial statistics. (Pinheiro \& Bates, 2000, pp 254–255.)
This is just a general, positive-definite covariance matrix. It adds a significant number of free parameters to the fitting process, since a $p \times p$ covariance matrix has $p(p + 1)/2$ non-redundant elements.

$$
\begin{bmatrix}
\sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\
\sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\
\sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\
\sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2
\end{bmatrix}
$$

(18)
A Composite Growth Curve Model for Cognitive Performance
Deriving the Residual Covariance Structure
Modeling the Residual Covariance Structure
Which Residual Structure?
Some Common Covariance Structures

Fixed Effects Modeling of Composite Residual Structure
Mixed Effects Modeling with Nonstandard Residual Covariance Structures

Unstructured
Compound Symmetry
Heterogeneous Compound Symmetry
Autoregressive
Heterogeneous Autoregressive
Toeplitz

Compound Symmetry

\[
\begin{bmatrix}
\sigma^2 + \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\
\sigma_1^2 & \sigma^2 + \sigma_1^2 & \sigma_1^2 \\
\sigma_1^2 & \sigma_1^2 & \sigma^2 + \sigma_1^2 \\
\end{bmatrix}
\]

(19)

James H. Steiger
Modeling Residual Covariance Structure
Heterogeneous Compound Symmetry

\[
\begin{bmatrix}
\sigma_1^2 & \sigma_2 \sigma_1 \rho & \sigma_3 \sigma_1 \rho & \sigma_4 \sigma_1 \rho \\
\sigma_2 \sigma_1 \rho & \sigma_2^2 & \sigma_3 \sigma_2 \rho & \sigma_4 \sigma_2 \rho \\
\sigma_3 \sigma_1 \rho & \sigma_3 \sigma_2 \rho & \sigma_3^2 & \sigma_4 \sigma_3 \rho \\
\sigma_4 \sigma_1 \rho & \sigma_4 \sigma_2 \rho & \sigma_4 \sigma_3 \rho & \sigma_4^2
\end{bmatrix}
\]

(20)
Some Common Covariance Structures

Autoregressive

\[
\begin{pmatrix}
\sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2 & \sigma^2 \rho^3 \\
\sigma^2 \rho & \sigma^2 & \sigma^2 \rho & \sigma^2 \rho^2 \\
\sigma^2 \rho^2 & \sigma^2 \rho & \sigma^2 & \sigma^2 \rho \\
\sigma^2 \rho^3 & \sigma^2 \rho^2 & \sigma^2 \rho & \sigma^2
\end{pmatrix}
\]
Heterogeneous Autoregressive

$$
\begin{bmatrix}
\sigma_1^2 & \sigma_2 \sigma_1 & \sigma_3 \sigma_1^2 & \sigma_4 \sigma_1^3 \\
\sigma_2 \sigma_1 & \sigma_2^2 & \sigma_3 \sigma_2^2 & \sigma_4 \sigma_2^2 \\
\sigma_3 \sigma_1^2 & \sigma_3 \sigma_2^2 & \sigma_3^2 & \sigma_4 \sigma_3 \\
\sigma_4 \sigma_1^3 & \sigma_4 \sigma_2^2 & \sigma_4 \sigma_3 & \sigma_4^2
\end{bmatrix}
$$

(22)

James H. Steiger

Modeling Residual Covariance Structure
Introduction
A Composite Growth Curve Model for Cognitive Performance
Deriving the Residual Covariance Structure
Modeling the Residual Covariance Structure
Which Residual Structure?

Some Common Covariance Structures
Fixed Effects Modeling of Composite Residual Structure
Mixed Effects Modeling with Nonstandard Residual Covariance Structure

Toeplitz

\[
\begin{bmatrix}
\sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\
\sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\
\sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\
\sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \\
\end{bmatrix}
\]

(23)
Using the R function `gls` in the `nlme` library, we can model the covariance structure of a fixed-effects linear model. Pinheiro and Bates (2000) refer to this as the *extended linear model*, because it replaces the normal assumption that the $\epsilon_i$ have a covariance matrix of $\sigma^2 I$ with a more complex model. This model may assume that variances are equal, or it may allow them to be unequal. Various models for the correlation structure of the errors are supported. This option is discussed in detail by Pinheiro and Bates (2000), Section 5.4.
Mixed Effects Modeling with Nonstandard Residual Covariance Structure

The R function `lme` in the `nlme` library has a facility for modeling the covariance structure of residuals within the mixed model framework. This capability is discussed in Chapter 5 of Pinheiro and Bates (2000).