Gain Greater Insight into Your SAS® Software with SAS Books.

Discover all that you need on your journey to knowledge and empowerment.

support.sas.com/bookstore for additional books and resources.
# Chapter 63
## The MIXED Procedure

## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview: MIXED Procedure</td>
<td>5234</td>
</tr>
<tr>
<td>Basic Features</td>
<td>5235</td>
</tr>
<tr>
<td>Notation for the Mixed Model</td>
<td>5236</td>
</tr>
<tr>
<td>PROC MIXED Contrasted with Other SAS Procedures</td>
<td>5237</td>
</tr>
<tr>
<td>Getting Started: MIXED Procedure</td>
<td>5238</td>
</tr>
<tr>
<td>Clustered Data Example</td>
<td>5238</td>
</tr>
<tr>
<td>Syntax: MIXED Procedure</td>
<td>5244</td>
</tr>
<tr>
<td>PROC MIXED Statement</td>
<td>5246</td>
</tr>
<tr>
<td>BY Statement</td>
<td>5258</td>
</tr>
<tr>
<td>CLASS Statement</td>
<td>5258</td>
</tr>
<tr>
<td>CODE Statement</td>
<td>5259</td>
</tr>
<tr>
<td>CONTRAST Statement</td>
<td>5260</td>
</tr>
<tr>
<td>ESTIMATE Statement</td>
<td>5263</td>
</tr>
<tr>
<td>ID Statement</td>
<td>5265</td>
</tr>
<tr>
<td>LSMEANS Statement</td>
<td>5265</td>
</tr>
<tr>
<td>LSMESTIMATE Statement</td>
<td>5271</td>
</tr>
<tr>
<td>MODEL Statement</td>
<td>5272</td>
</tr>
<tr>
<td>PARMS Statement</td>
<td>5286</td>
</tr>
<tr>
<td>PERFORMANCE Statement</td>
<td>5289</td>
</tr>
<tr>
<td>PRIOR Statement</td>
<td>5290</td>
</tr>
<tr>
<td>RANDOM Statement</td>
<td>5294</td>
</tr>
<tr>
<td>REPEATED Statement</td>
<td>5298</td>
</tr>
<tr>
<td>SLICE Statement</td>
<td>5312</td>
</tr>
<tr>
<td>STORE Statement</td>
<td>5312</td>
</tr>
<tr>
<td>WEIGHT Statement</td>
<td>5312</td>
</tr>
<tr>
<td>Details: MIXED Procedure</td>
<td>5313</td>
</tr>
<tr>
<td>Mixed Models Theory</td>
<td>5313</td>
</tr>
<tr>
<td>Parameterization of Mixed Models</td>
<td>5325</td>
</tr>
<tr>
<td>Residuals and Influence Diagnostics</td>
<td>5330</td>
</tr>
<tr>
<td>Default Output</td>
<td>5337</td>
</tr>
<tr>
<td>ODS Table Names</td>
<td>5341</td>
</tr>
<tr>
<td>ODS Graphics</td>
<td>5346</td>
</tr>
<tr>
<td>Computational Issues</td>
<td>5351</td>
</tr>
<tr>
<td>Examples: MIXED Procedure</td>
<td>5355</td>
</tr>
<tr>
<td>Example 63.1: Split-Plot Design</td>
<td>5355</td>
</tr>
<tr>
<td>Example 63.2: Repeated Measures</td>
<td>5361</td>
</tr>
</tbody>
</table>
Overview: MIXED Procedure

The MIXED procedure fits a variety of mixed linear models to data and enables you to use these fitted models to make statistical inferences about the data. A mixed linear model is a generalization of the standard linear model used in the GLM procedure, the generalization being that the data are permitted to exhibit correlation and nonconstant variability. The mixed linear model, therefore, provides you with the flexibility of modeling not only the means of your data (as in the standard linear model) but their variances and covariances as well.

The primary assumptions underlying the analyses performed by PROC MIXED are as follows:

- The data are normally distributed (Gaussian).
- The means (expected values) of the data are linear in terms of a certain set of parameters.
- The variances and covariances of the data are in terms of a different set of parameters, and they exhibit a structure matching one of those available in PROC MIXED.

Since Gaussian data can be modeled entirely in terms of their means and variances/covariances, the two sets of parameters in a mixed linear model actually specify the complete probability distribution of the data. The parameters of the mean model are referred to as fixed-effects parameters, and the parameters of the variance-covariance model are referred to as covariance parameters.

The fixed-effects parameters are associated with known explanatory variables, as in the standard linear model. These variables can be either qualitative (as in the traditional analysis of variance) or quantitative (as in standard linear regression). However, the covariance parameters are what distinguishes the mixed linear model from the standard linear model.

The need for covariance parameters arises quite frequently in applications, the following being the two most typical scenarios:

- The experimental units on which the data are measured can be grouped into clusters, and the data from a common cluster are correlated.
- Repeated measurements are taken on the same experimental unit, and these repeated measurements are correlated or exhibit variability that changes.
The first scenario can be generalized to include one set of clusters nested within another. For example, if students are the experimental unit, they can be clustered into classes, which in turn can be clustered into schools. Each level of this hierarchy can introduce an additional source of variability and correlation. The second scenario occurs in longitudinal studies, where repeated measurements are taken over time. Alternatively, the repeated measures could be spatial or multivariate in nature.

PROC MIXED provides a variety of covariance structures to handle the previous two scenarios. The most common of these structures arises from the use of random-effects parameters, which are additional unknown random variables assumed to affect the variability of the data. The variances of the random-effects parameters, commonly known as variance components, become the covariance parameters for this particular structure. Traditional mixed linear models contain both fixed- and random-effects parameters, and, in fact, it is the combination of these two types of effects that led to the name mixed model. PROC MIXED fits not only these traditional variance component models but numerous other covariance structures as well.

PROC MIXED fits the structure you select to the data by using the method of restricted maximum likelihood (REML), also known as residual maximum likelihood. It is here that the Gaussian assumption for the data is exploited. Other estimation methods are also available, including maximum likelihood and MIVQUE0. The details behind these estimation methods are discussed in subsequent sections.

After a model has been fit to your data, you can use it to draw statistical inferences via both the fixed-effects and covariance parameters. PROC MIXED computes several different statistics suitable for generating hypothesis tests and confidence intervals. The validity of these statistics depends upon the mean and variance-covariance model you select, so it is important to choose the model carefully. Some of the output from PROC MIXED helps you assess your model and compare it with others.

Basic Features

PROC MIXED provides easy accessibility to numerous mixed linear models that are useful in many common statistical analyses. In the style of the GLM procedure, PROC MIXED fits the specified mixed linear model and produces appropriate statistics.

Here are some basic features of PROC MIXED:

- covariance structures, including variance components, compound symmetry, unstructured, AR(1), Toeplitz, spatial, general linear, and factor analytic
- GLM-type grammar, by using MODEL, RANDOM, and REPEATED statements for model specification and CONTRAST, ESTIMATE, and LSMEANS statements for inferences
- appropriate standard errors for all specified estimable linear combinations of fixed and random effects, and corresponding t and F tests
- subject and group effects that enable blocking and heterogeneity, respectively
- REML and ML estimation methods implemented with a Newton-Raphson algorithm
- capacity to handle unbalanced data
- ability to create a SAS data set corresponding to any table
PROC MIXED uses the Output Delivery System (ODS), a SAS subsystem that provides capabilities for displaying and controlling the output from SAS procedures. ODS enables you to convert any of the output from PROC MIXED into a SAS data set. See the section “ODS Table Names” on page 5341.

The MIXED procedure uses ODS Graphics to create graphs as part of its output. For general information about ODS Graphics, see Chapter 21, “Statistical Graphics Using ODS.” For specific information about the statistical graphics available with the MIXED procedure, see the PLOTS= option in the PROC MIXED statement and the section “ODS Graphics” on page 5346.

Notation for the Mixed Model

This section introduces the mathematical notation used throughout this chapter to describe the mixed linear model. You should be familiar with basic matrix algebra (see Searle 1982). A more detailed description of the mixed model is contained in the section “Mixed Models Theory” on page 5313.

A statistical model is a mathematical description of how data are generated. The standard linear model, as used by the GLM procedure, is one of the most common statistical models:

\[
y = X\beta + \epsilon
\]

In this expression, \( y \) represents a vector of observed data, \( \beta \) is an unknown vector of fixed-effects parameters with known design matrix \( X \), and \( \epsilon \) is an unknown random error vector modeling the statistical noise around \( X\beta \). The focus of the standard linear model is to model the mean of \( y \) by using the fixed-effects parameters \( \beta \). The residual errors \( \epsilon \) are assumed to be independent and identically distributed Gaussian random variables with mean 0 and variance \( \sigma^2 \).

The mixed model generalizes the standard linear model as follows:

\[
y = X\beta + Z\gamma + \epsilon
\]

Here, \( \gamma \) is an unknown vector of random-effects parameters with known design matrix \( Z \), and \( \epsilon \) is an unknown random error vector whose elements are no longer required to be independent and homogeneous.

To further develop this notion of variance modeling, assume that \( \gamma \) and \( \epsilon \) are Gaussian random variables that are uncorrelated and have expectations 0 and variances \( G \) and \( R \), respectively. The variance of \( y \) is thus

\[
V = ZGZ' + R
\]

Note that, when \( R = \sigma^2 I \) and \( Z = 0 \), the mixed model reduces to the standard linear model.

You can model the variance of the data, \( y \), by specifying the structure (or form) of \( Z \), \( G \), and \( R \). The model matrix \( Z \) is set up in the same fashion as \( X \), the model matrix for the fixed-effects parameters. For \( G \) and \( R \), you must select some covariance structure. Possible covariance structures include the following:

- variance components
- compound symmetry (common covariance plus diagonal)
- unstructured (general covariance)
- autoregressive
PROC MIXED Contrasted with Other SAS Procedures

PROC MIXED is a generalization of the GLM procedure in the sense that PROC GLM fits standard linear models, and PROC MIXED fits the wider class of mixed linear models. Both procedures have similar CLASS, MODEL, CONTRAST, ESTIMATE, and LSMEANS statements, but their RANDOM and REPEATED statements differ (see the following paragraphs). Both procedures use the non-full-rank model parameterization, although the sorting of classification levels can differ between the two. PROC MIXED computes only Type I–Type III tests of fixed effects, while PROC GLM computes Types I–IV.

The RANDOM statement in PROC MIXED incorporates random effects constituting the $\gamma$ vector in the mixed model. However, in PROC GLM, effects specified in the RANDOM statement are still treated as fixed as far as the model fit is concerned, and they serve only to produce corresponding expected mean squares. These expected mean squares lead to the traditional ANOVA estimates of variance components. PROC MIXED computes REML and ML estimates of variance parameters, which are generally preferred to the ANOVA estimates (Searle 1988; Harville 1988; Searle, Casella, and McCulloch 1992). Optionally, PROC MIXED also computes MIVQUE0 estimates, which are similar to ANOVA estimates.

The REPEATED statement in PROC MIXED is used to specify covariance structures for repeated measurements on subjects, while the REPEATED statement in PROC GLM is used to specify various transformations with which to conduct the traditional univariate or multivariate tests. In repeated measures situations, the mixed model approach used in PROC MIXED is more flexible and more widely applicable than either the univariate or multivariate approach. In particular, the mixed model approach provides a larger class of covariance structures and a better mechanism for handling missing values (Wolfinger and Chang 1995).

PROC MIXED subsumes the V ARCOMP procedure. PROC MIXED provides a wide variety of covariance structures, while PROC V ARCOMP estimates only simple random effects. PROC MIXED carries out several analyses that are absent in PROC V ARCOMP, including the estimation and testing of linear combinations of fixed and random effects.

The ARIMA and AUTOREG procedures provide more time series structures than PROC MIXED, although they do not fit variance component models. The CALIS procedure fits general covariance matrices, but the fixed effects structure of the model is formed differently than in PROC MIXED. The LATTICE and NESTED procedures fit special types of mixed linear models that can also be handled in PROC MIXED, although PROC MIXED might run slower because of its more general algorithm. The TSCSREG procedure analyzes time series cross-sectional data, and it fits some structures not available in PROC MIXED.

The GLIMMIX procedure fits generalized linear mixed models (GLMMs). Linear mixed models—where the data are normally distributed, given the random effects—are in the class of GLMMs. The MIXED procedure can estimate covariance parameters with ANOVA methods that are not available in the GLIMMIX procedure (see METHOD=TYPE1, METHOD=TYPE2, and METHOD=TYPE3 in the PROC MIXED statement). Also, PROC MIXED can perform a sampling-based Bayesian analysis through the PRIOR statement, and

• spatial
• general linear
• factor analytic

By appropriately defining the model matrices $X$ and $Z$, as well as the covariance structure matrices $G$ and $R$, you can perform numerous mixed model analyses.
the procedure supports certain Kronecker-type covariance structures. These features are not available in the GLIMMIX procedure. The GLIMMIX procedure, on the other hand, accommodates nonnormal data and offers a broader array of post-processing features than the MIXED procedure.

### Getting Started: MIXED Procedure

#### Clustered Data Example

Consider the following SAS data set as an introductory example:

```sas
data heights;
  input Family Gender$ Height @@;
datalines;
1 F 67 1 F 66 1 F 64 1 M 71 1 M 72 2 F 63
2 F 63 2 F 67 2 M 69 2 M 68 2 M 70 3 F 63
3 M 64 4 F 67 4 F 66 4 M 67 4 M 67 4 M 69 ;
```

The response variable `Height` measures the heights (in inches) of 18 individuals. The individuals are classified according to `Family` and `Gender`. You can perform a traditional two-way analysis of variance of these data with the following PROC MIXED statements:

```sas
proc mixed data=heights;
  class Family Gender;
  model Height = Gender Family Family*Gender;
run;
```

The `PROC MIXED` statement invokes the procedure. The `CLASS` statement instructs PROC MIXED to consider both `Family` and `Gender` as classification variables. Dummy (indicator) variables are, as a result, created corresponding to all of the distinct levels of `Family` and `Gender`. For these data, `Family` has four levels and `Gender` has two levels.

The `MODEL` statement first specifies the response (dependent) variable `Height`. The explanatory (independent) variables are then listed after the equal (=) sign. Here, the two explanatory variables are `Gender` and `Family`, and these are the main effects of the design. The third explanatory term, `Family*Gender`, models an interaction between the two main effects.

PROC MIXED uses the dummy variables associated with `Gender`, `Family`, and `Family*Gender` to construct the `X` matrix for the linear model. A column of 1s is also included as the first column of `X` to model a global intercept. There are no `Z` or `G` matrices for this model, and `R` is assumed to equal $\sigma^2 I$, where `I` is an 18 $\times$ 18 identity matrix.

The `RUN` statement completes the specification. The coding is precisely the same as with the GLM procedure. However, much of the output from PROC MIXED is different from that produced by PROC GLM.

The output from PROC MIXED is shown in Figure 63.1–Figure 63.7.

The “Model Information” table in Figure 63.1 describes the model, some of the variables that it involves, and the method used in fitting it. This table also lists the method (profile, factor, parameter, or none) for handling the residual variance.
The “Class Level Information” table in Figure 63.2 lists the levels of all variables specified in the CLASS statement. You can check this table to make sure that the data are correct.

The “Dimensions” table in Figure 63.3 lists the sizes of relevant matrices. This table can be useful in determining CPU time and memory requirements.

The “Number of Observations” table in Figure 63.4 displays information about the sample size being processed.
The “Covariance Parameter Estimates” table in Figure 63.5 displays the estimate of $\sigma^2$ for the model.

**Figure 63.5 Covariance Parameter Estimates**

<table>
<thead>
<tr>
<th>Covariance Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cov Parm</td>
</tr>
<tr>
<td>Residual</td>
</tr>
</tbody>
</table>

The “Fit Statistics” table in Figure 63.6 lists several pieces of information about the fitted mixed model, including values derived from the computed value of the restricted/residual likelihood.

**Figure 63.6 Fit Statistics**

<table>
<thead>
<tr>
<th>Fit Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 Res Log Likelihood</td>
</tr>
<tr>
<td>AIC (Smaller is Better)</td>
</tr>
<tr>
<td>AICC (Smaller is Better)</td>
</tr>
<tr>
<td>BIC (Smaller is Better)</td>
</tr>
</tbody>
</table>

The “Type 3 Tests of Fixed Effects” table in Figure 63.7 displays significance tests for the three effects listed in the `MODEL` statement. The Type 3 $F$ statistics and $p$-values are the same as those produced by the GLM procedure. However, because PROC MIXED uses a likelihood-based estimation scheme, it does not directly compute or display sums of squares for this analysis.

**Figure 63.7 Tests of Fixed Effects**

<table>
<thead>
<tr>
<th>Type 3 Tests of Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Family</td>
</tr>
<tr>
<td>Family*Gender</td>
</tr>
</tbody>
</table>
The Type 3 test for Family*Gender effect is not significant at the 5% level, but the tests for both main effects are significant.

The important assumptions behind this analysis are that the data are normally distributed and that they are independent with constant variance. For these data, the normality assumption is probably realistic since the data are observed heights. However, since the data occur in clusters (families), it is very likely that observations from the same family are statistically correlated—that is, not independent.

The methods implemented in PROC MIXED are still based on the assumption of normally distributed data, but you can drop the assumption of independence by modeling statistical correlation in a variety of ways. You can also model variances that are heterogeneous—that is, nonconstant.

For the height data, one of the simplest ways of modeling correlation is through the use of random effects. Here the family effect is assumed to be normally distributed with zero mean and some unknown variance. This is in contrast to the previous model in which the family effects are just constants, or fixed effects. Declaring Family as a random effect sets up a common correlation among all observations having the same level of Family.

Declaring Family*Gender as a random effect models an additional correlation between all observations that have the same level of both Family and Gender. One interpretation of this effect is that a female in a certain family exhibits more correlation with the other females in that family than with the other males, and likewise for a male. With the height data, this model seems reasonable.

The statements to fit this correlation model in PROC MIXED are as follows:

```plaintext
proc mixed;
   class Family Gender;
   model Height = Gender;
   random Family Family*Gender;
run;
```

Note that Family and Family*Gender are now listed in the RANDOM statement. The dummy variables associated with them are used to construct the Z matrix in the mixed model. The X matrix now consists of a column of 1s and the dummy variables for Gender.

The G matrix for this model is diagonal, and it contains the variance components for both Family and Family*Gender. The R matrix is still assumed to equal \( \sigma^2 I \), where I is an identity matrix.

The output from this analysis is as follows.

---

**Figure 63.8 Model Information**

<table>
<thead>
<tr>
<th>The Mixed Procedure</th>
<th>Model Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Set</td>
<td>WORK.HEIGHTS</td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>Height</td>
</tr>
<tr>
<td>Covariance Structure</td>
<td>Variance Components</td>
</tr>
<tr>
<td>Estimation Method</td>
<td>REML</td>
</tr>
<tr>
<td>Residual Variance Method</td>
<td>Profile</td>
</tr>
<tr>
<td>Fixed Effects SE Method</td>
<td>Model-Based</td>
</tr>
<tr>
<td>Degrees of Freedom Method</td>
<td>Containment</td>
</tr>
</tbody>
</table>
The “Model Information” table in Figure 63.8 shows that the containment method is used to compute the degrees of freedom for this analysis. This is the default method when a RANDOM statement is used; see the description of the DDFM= option for more information.

**Figure 63.9** Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>Gender</td>
<td>2</td>
<td>F M</td>
</tr>
</tbody>
</table>

The “Class Level Information” table in Figure 63.9 is the same as before. The “Dimensions” table in Figure 63.10 displays the new sizes of the X and Z matrices.

**Figure 63.10** Dimensions and Number of Observations

<table>
<thead>
<tr>
<th>Dimensions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariance Parameters</td>
<td>3</td>
</tr>
<tr>
<td>Columns in X</td>
<td>3</td>
</tr>
<tr>
<td>Columns in Z</td>
<td>12</td>
</tr>
<tr>
<td>Subjects</td>
<td>1</td>
</tr>
<tr>
<td>Max Obs per Subject</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Observations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations Read</td>
<td>18</td>
</tr>
<tr>
<td>Number of Observations Used</td>
<td>18</td>
</tr>
<tr>
<td>Number of Observations Not Used</td>
<td>0</td>
</tr>
</tbody>
</table>

The “Iteration History” table in Figure 63.11 displays the results of the numerical optimization of the restricted/residual likelihood. Six iterations are required to achieve the default convergence criterion of 1E–8.

**Figure 63.11** REML Estimation Iteration History

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Evaluations</th>
<th>-2 Res Log Like</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>74.11074833</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>71.51614003</td>
<td>0.01441208</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>71.13845990</td>
<td>0.00412226</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>71.03613556</td>
<td>0.00058188</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>71.02281757</td>
<td>0.00001689</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>71.02245904</td>
<td>0.00000002</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>71.02245869</td>
<td>0.00000000</td>
</tr>
</tbody>
</table>

Convergence criteria met.
The “Covariance Parameter Estimates” table in Figure 63.12 displays the results of the REML fit. The Estimate column contains the estimates of the variance components for Family and Family*Gender, as well as the estimate of $\sigma^2$.

**Figure 63.12** Covariance Parameter Estimates (REML)

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>2.4010</td>
</tr>
<tr>
<td>Family*Gender</td>
<td>1.7657</td>
</tr>
<tr>
<td>Residual</td>
<td>2.1668</td>
</tr>
</tbody>
</table>

The “Fit Statistics” table in Figure 63.13 contains basic information about the REML fit.

**Figure 63.13** Fit Statistics

<table>
<thead>
<tr>
<th>Fit Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 Res Log Likelihood</td>
</tr>
<tr>
<td>AIC (Smaller is Better)</td>
</tr>
<tr>
<td>AICC (Smaller is Better)</td>
</tr>
<tr>
<td>BIC (Smaller is Better)</td>
</tr>
</tbody>
</table>

The “Type 3 Tests of Fixed Effects” table in Figure 63.14 contains a significance test for the lone fixed effect, Gender. Note that the associated $p$-value is not nearly as significant as in the previous analysis. This illustrates the importance of correctly modeling correlation in your data.

**Figure 63.14** Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Type 3 Tests of Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>Gender</td>
</tr>
</tbody>
</table>

An additional benefit of the random effects analysis is that it enables you to make inferences about gender that apply to an entire population of families, whereas the inferences about gender from the analysis where Family and Family*Gender are fixed effects apply only to the particular families in the data set.

PROC MIXED thus offers you the ability to model correlation directly and to make inferences about fixed effects that apply to entire populations of random effects.
Syntax: MIXED Procedure

The following statements are available in the MIXED procedure:

PROC MIXED < options > ;
   BY variables ;
   CLASS variable < (REF= option) > . . . < variable < (REF= option) > > < / global-options > ;
   CODE < options > ;
   ID variables ;
   MODEL dependent = < fixed-effects > < / options > ;
   RANDOM random-effects < / options > ;
   REPEATED < repeated-effect > < / options > ;
   PARMS (value-list) . . . < / options > ;
   PERFORMANCE < options > ;
   PRIOR < distribution > < / options > ;
   CONTRAST 'label' < fixed-effect values . . . >
      < | random-effect values . . . > . . . < / options > ;
   ESTIMATE 'label' < fixed-effect values . . . >
      < | random-effect values . . . > < / options > ;
   LSMEANS fixed-effects < / options > ;
   LSMESTIMATE model-effect lsmestimate-specification < / options > ;
   SLICE model-effect < / options > ;
   STORE < OUT= > item-store-name < / LABEL= 'label' > ;
   WEIGHT variable ;

Items within angle brackets ( < > ) are optional. The CONTRAST, ESTIMATE, LSMEANS, and RANDOM statements can appear multiple times; all other statements can appear only once.

The PROC MIXED and MODEL statements are required, and the MODEL statement must appear after the CLASS statement if a CLASS statement is included. The CONTRAST, ESTIMATE, LSMEANS, RANDOM, and REPEATED statements must follow the MODEL statement. The CONTRAST and ESTIMATE statements must also follow any RANDOM statements. The LSMESTIMATE, SLICE, and STORE statements are shared with many procedures. Summary descriptions of functionality and syntax for these statements are also given after the PROC MIXED statement in alphabetical order, but you can find full documentation on them in Chapter 19, “Shared Concepts and Topics.”

Table 63.1 summarizes the basic functions and important options of each PROC MIXED statement. The syntax of each statement in Table 63.1 is described in the following sections in alphabetical order after the description of the PROC MIXED statement.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Description</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROC MIXED</td>
<td>Invokes the procedure</td>
<td>DATA= specifies input data set, METHOD= specifies estimation method</td>
</tr>
<tr>
<td>BY</td>
<td>Performs multiple PROC MIXED analyses in one invocation</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 63.1 Summary of PROC MIXED Statements
## Table 63.1  continued

<table>
<thead>
<tr>
<th>Statement</th>
<th>Description</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLASS</td>
<td>Declares qualitative variables that create indicator variables in design matrices</td>
<td>None</td>
</tr>
<tr>
<td>CODE</td>
<td>Requests that the procedure write SAS DATA step code to a file or catalog entry</td>
<td>FILE= names the file where the generated code is saved, CATALOG= names the catalog entry where the generated code is saved, IMPUTE imputes predicted values for observations with missing or invalid covariates, RESIDUAL computes residuals</td>
</tr>
<tr>
<td>ID</td>
<td>Lists additional variables to be included in predicted values tables</td>
<td>None</td>
</tr>
<tr>
<td>MODEL</td>
<td>Specifies dependent variable and fixed effects, setting up X</td>
<td>S requests solution for fixed-effects parameters, DDFM= specifies denominator degrees of freedom method, OUTP= outputs predicted values to a data set, INFLUENCE computes influence diagnostics</td>
</tr>
<tr>
<td>RANDOM</td>
<td>Specifies random effects, setting up Z and G</td>
<td>SUBJECT= creates block-diagonality, TYPE= specifies covariance structure, S requests solution for random-effects parameters, G displays estimated G</td>
</tr>
<tr>
<td>REPEATED</td>
<td>Sets up R</td>
<td>SUBJECT= creates block-diagonality, TYPE= specifies covariance structure, R displays estimated blocks of R, GROUP= enables between-subject heterogeneity, LOCAL adds a diagonal matrix to R</td>
</tr>
<tr>
<td>PARMS</td>
<td>Specifies a grid of initial values for the covariance parameters</td>
<td>HOLD= and NOITER hold the covariance parameters or their ratios constant, PARMSDATA= reads the initial values from a SAS data set</td>
</tr>
<tr>
<td>PERFORMANCE</td>
<td>Specifies the number of threads to use</td>
<td>NTHREADS=</td>
</tr>
<tr>
<td>PRIOR</td>
<td>Performs a sampling-based Bayesian analysis for variance component models</td>
<td>NSAMPLE= specifies the sample size, SEED= specifies the starting seed</td>
</tr>
<tr>
<td>CONTRAST</td>
<td>Constructs custom hypothesis tests</td>
<td>E displays the L matrix coefficients</td>
</tr>
<tr>
<td>ESTIMATE</td>
<td>Constructs custom scalar estimates</td>
<td>CL produces confidence limits</td>
</tr>
<tr>
<td>LSMEANS</td>
<td>Computes least squares means for classification fixed effects</td>
<td>DIFF computes differences of the least squares means, ADJUST= performs multiple comparisons adjustments, AT changes covariates, OM changes weighting, CL produces confidence limits, SLICE= tests simple effects</td>
</tr>
</tbody>
</table>
Table 63.1  continued

<table>
<thead>
<tr>
<th>Statement</th>
<th>Description</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSMESTIMATE</td>
<td>Provides custom hypothesis tests among the least squares means</td>
<td>ADJUST= determines the method for multiple comparison adjustment of LS-mean differences, JOINT requests a joint ( F ) or chi-square test for the rows of the estimate</td>
</tr>
<tr>
<td>SLICE</td>
<td>Performs a partitioned analysis of LS–means for an interaction</td>
<td>ADJUST= determines the method for multiple comparison adjustment of LS-mean differences, DIFF requests differences of LS-means</td>
</tr>
<tr>
<td>STORE</td>
<td>Saves the context and results of the analysis</td>
<td>LABEL= adds a custom label</td>
</tr>
<tr>
<td>WEIGHT</td>
<td>Specifies a variable by which to weight ( R )</td>
<td>None</td>
</tr>
</tbody>
</table>

**PROC MIXED Statement**

PROC MIXED < options> ;

The PROC MIXED statement invokes the MIXED procedure. Table 63.2 summarizes the options available in the PROC MIXED statement. These and other options in the PROC MIXED statement are then described fully in alphabetical order.

Table 63.2  PROC MIXED Statement Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic Options</strong></td>
<td></td>
</tr>
<tr>
<td>DATA=</td>
<td>Specifies input data set</td>
</tr>
<tr>
<td>METHOD=</td>
<td>Specifies the estimation method</td>
</tr>
<tr>
<td>NOPROFILE</td>
<td>Includes scale parameter in optimization</td>
</tr>
<tr>
<td>ORDER=</td>
<td>Determines the sort order of CLASS variables</td>
</tr>
<tr>
<td><strong>Displayed Output</strong></td>
<td></td>
</tr>
<tr>
<td>ASYCORR</td>
<td>Displays asymptotic correlation matrix of covariance parameter estimates</td>
</tr>
<tr>
<td>ASYCOV</td>
<td>Displays asymptotic covariance matrix of covariance parameter estimates</td>
</tr>
<tr>
<td>CL</td>
<td>Requests confidence limits for covariance parameter estimates</td>
</tr>
<tr>
<td>COVTEST</td>
<td>Displays asymptotic standard errors and Wald tests for covariance parameters</td>
</tr>
<tr>
<td>IC</td>
<td>Displays a table of information criteria</td>
</tr>
<tr>
<td>ITDETAILS</td>
<td>Displays estimates and gradients added to “Iteration History”</td>
</tr>
<tr>
<td>LOGNOTE</td>
<td>Writes periodic status notes to the log</td>
</tr>
<tr>
<td>MMEQ</td>
<td>Displays mixed model equations</td>
</tr>
<tr>
<td>MMEQSOL</td>
<td>Displays the solution to the mixed model equations</td>
</tr>
<tr>
<td>NOCLPRINT</td>
<td>Suppresses “Class Level Information” completely or in parts</td>
</tr>
</tbody>
</table>
Table 63.2  continued

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOITPRINT</td>
<td>Suppresses “Iteration History” table</td>
</tr>
<tr>
<td>PLOTS=</td>
<td>Produces ODS statistical graphics</td>
</tr>
<tr>
<td>RANKS=</td>
<td>Displays a table of ranks of design matrices $X$ and $(XZ)$</td>
</tr>
<tr>
<td>RATIO</td>
<td>Produces ratio of covariance parameter estimates with residual variance</td>
</tr>
</tbody>
</table>

### Optimization Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAXFUNC=</td>
<td>Specifies the maximum number of likelihood evaluations</td>
</tr>
<tr>
<td>MAXITER=</td>
<td>Specifies the maximum number of iterations</td>
</tr>
</tbody>
</table>

### Computational Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONVF</td>
<td>Requests and tunes the relative function convergence criterion</td>
</tr>
<tr>
<td>CONVG</td>
<td>Requests and tunes the relative gradient convergence criterion</td>
</tr>
<tr>
<td>CONVH</td>
<td>Requests and tunes the relative Hessian convergence criterion</td>
</tr>
<tr>
<td>DFBW</td>
<td>Selects between-within degree of freedom method</td>
</tr>
<tr>
<td>EMPIRICAL</td>
<td>Computes empirical (“sandwich”) estimators</td>
</tr>
<tr>
<td>NOBOUND</td>
<td>Unbounds covariance parameter estimates</td>
</tr>
<tr>
<td>RIDGE=</td>
<td>Specifies starting value for minimum ridge value</td>
</tr>
<tr>
<td>SCORING=</td>
<td>Applies Fisher scoring where applicable</td>
</tr>
</tbody>
</table>

You can specify the following options.

**ABSOLUTE**

makes the convergence criterion absolute. By default, it is relative (divided by the current objective function value). See the CONVF, CONVG, and CONVH options in this section for a description of various convergence criteria.

**ALPHA=number**

requests that confidence limits be constructed for the covariance parameter estimates with confidence level $1 - \text{number}$. The value of \text{number} must be between 0 and 1; the default is 0.05.

**ANOVAF**

The ANOVAF option computes $F$ tests in models with REPEATED statement and without RANDOM statement by a method similar to that of Brunner, Domhof, and Langer (2002). The method consists of computing special $F$ statistics and adjusting their degrees of freedom. The technique is a generalization of the Greenhouse-Geisser adjustment in MANOVA models (Greenhouse and Geisser 1959). For more details, see the section “$F$ Tests With the ANOVAF Option” on page 5323.

**ASYCORR**

produces the asymptotic correlation matrix of the covariance parameter estimates. It is computed from the corresponding asymptotic covariance matrix (see the description of the ASYCOV option, which follows). The name of the “Asymptotic Correlation” table is AsyCorr.
ASYCOV
requests that the asymptotic covariance matrix of the covariance parameters be displayed. By default, this matrix is the observed inverse Fisher information matrix, which equals $2\mathbf{H}^{-1}$, where $\mathbf{H}$ is the Hessian (second derivative) matrix of the objective function. See the section “Covariance Parameter Estimates” on page 5339 for more information about this matrix. When you use the SCORING= option and PROC MIXED converges without stopping the scoring algorithm, PROC MIXED uses the expected Hessian matrix to compute the covariance matrix instead of the observed Hessian. The ODS name of the “Asymptotic Covariance” table is AsyCov.

CL=WALD
requests confidence limits for the covariance parameter estimates. A Satterthwaite approximation is used to construct limits for all parameters that have a lower boundary constraint of zero. These limits take the form

$$\frac{v\hat{\sigma}^2}{\chi^2_{v,1-\alpha/2}} \leq \sigma^2 \leq \frac{v\hat{\sigma}^2}{\chi^2_{v,\alpha/2}}$$

where $v = 2Z^2$, $Z$ is the Wald statistic $\frac{\hat{\sigma}^2 / \text{se}(\hat{\sigma}^2)}{\hat{\sigma}^2}$, and the denominators are quantiles of the $\chi^2$-distribution with $v$ degrees of freedom. See Milliken and Johnson (1992) and Burdick and Graybill (1992) for similar techniques.

For all other parameters, Wald Z-scores and normal quantiles are used to construct the limits. Wald limits are also provided for variance components if you specify the NOBOUND option. The optional =WALD specification requests Wald limits for all parameters.

The confidence limits are displayed as extra columns in the “Covariance Parameter Estimates” table. The confidence level is $1 - \alpha = 0.95$ by default; this can be changed with the ALPHA= option.

CONVF=number
requests the relative function convergence criterion with tolerance number. The relative function convergence criterion is

$$\frac{|f_k - f_{k-1}|}{|f_k|} \leq number$$

where $f_k$ is the value of the objective function at iteration $k$. To prevent the division by $|f_k|$, use the ABSOLUTE option. The default convergence criterion is CONVH, and the default tolerance is 1E–8.

CONVG=number
requests the relative gradient convergence criterion with tolerance number. The relative gradient convergence criterion is

$$\max_j \frac{|g_{jk}|}{|f_k|} \leq number$$

where $f_k$ is the value of the objective function, and $g_{jk}$ is the $j$th element of the gradient (first derivative) of the objective function, both at iteration $k$. To prevent division by $|f_k|$, use the ABSOLUTE option. The default convergence criterion is CONVH, and the default tolerance is 1E–8.
CONVH<\(=number\>>
requests the relative Hessian convergence criterion with tolerance \(number\). The relative Hessian convergence criterion is
\[
\frac{g_k' H^{-1} g_k}{|f_k|} \leq number
\]
where \(f_k\) is the value of the objective function, \(g_k\) is the gradient (first derivative) of the objective function, and \(H_k\) is the Hessian (second derivative) of the objective function, all at iteration \(k\).

If \(H_k\) is singular, then PROC MIXED uses the following relative criterion:
\[
\frac{g_k' g_k}{|f_k|} \leq number
\]

To prevent the division by \(|f_k|\), use the ABSOLUTE option. The default convergence criterion is CONVH, and the default tolerance is \(1E^{-8}\).

COVTEST
produces asymptotic standard errors and Wald Z-tests for the covariance parameter estimates.

DATA=\text{sas-data-set}
names the SAS data set to be used by PROC MIXED. The default is the most recently created data set.

DFBW
has the same effect as the DDFM=BW option in the MODEL statement.

EMPIRICAL
computes the estimated variance-covariance matrix of the fixed-effects parameters by using the asymptotically consistent estimator described in Huber (1967); White (1980); Liang and Zeger (1986); Diggle, Liang, and Zeger (1994). This estimator is commonly referred to as the “sandwich” estimator, and it is computed as follows:
\[
(X' \hat{\Sigma}^{-1} X)^{-1} \left( \sum_{i=1}^{S} X_i' \hat{\Sigma}_i^{-1} \hat{\epsilon}_i \hat{\epsilon}_i' \hat{\Sigma}_i^{-1} X_i \right) (X' \hat{\Sigma}^{-1} X)^{-1}
\]
Here, \(\hat{\epsilon}_i = y_i - X_i \hat{\beta}\), \(S\) is the number of subjects, and matrices with an \(i\) subscript are those for the \(i\)th subject. You must include the SUBJECT= option in either a RANDOM or REPEATED statement for this option to take effect.

When you specify the EMPIRICAL option, PROC MIXED adjusts all standard errors and test statistics involving the fixed-effects parameters. This changes output in the following tables (listed in Table 63.26): Contrast, CorrB, CovB, DiffS, Estimates, InvCovB, LSMMeans, Slices, SolutionF, Tests1–Tests3. The OUTP= and OUTPM= data sets are also affected. Finally, the Satterthwaite and Kenward-Roger degrees of freedom methods are not available if you specify the EMPIRICAL option.

IC
displays a table of various information criteria. The criteria are all in smaller-is-better form, and are described in Table 63.3.
Here $\ell$ denotes the maximum value of the (possibly restricted) log likelihood, $d$ the dimension of the model, and $n$ the number of observations. In SAS 6 of SAS/STAT software, $n$ equals the number of valid observations for maximum likelihood estimation and $n - p$ for restricted maximum likelihood estimation, where $p$ equals the rank of $X$. In later versions, $n$ equals the number of effective subjects as displayed in the “Dimensions” table, unless this value equals 1, in which case $n$ equals the number of levels of the first random effect you specify in a RANDOM statement. If the number of effective subjects equals 1 and you have no RANDOM statements, then $n$ reverts to the SAS 6 values. For AICC (a finite-sample corrected version of AIC), $n^*$ equals the SAS 6 values of $n$, unless this number is less than $d + 2$, in which case it equals $d + 2$. When $n = 0$, the value of the HQIC criterion is $-2\ell$. When $n=0$, the values of the BIC and CAIC criteria are $-2\ell$ and $-2\ell + d$, respectively.

For restricted likelihood estimation, $d$ equals $q$, the effective number of estimated covariance parameters. In SAS 6, when a parameter estimate lies on a boundary constraint, then it is still included in the calculation of $d$, but in later versions it is not. The most common example of this behavior is when a variance component is estimated to equal zero. For maximum likelihood estimation, $d$ equals $q + p$. The value of $d$ is displayed in the “Information Criteria” table as the value of Parms variable; see Table 63.27.

The ODS name of the “Information Criteria” table is InfoCrit.

INFO
is a default option. The creation of the “Model Information,” “Dimensions,” and “Number of Observations” tables can be suppressed by using the NOINFO option.

Note that in SAS 6 this option displays the “Model Information” and “Dimensions” tables.

ITDETAILS
displays the parameter values at each iteration and enables the writing of notes to the SAS log pertaining to “infinite likelihood” and “singularities” during Newton-Raphson iterations.

LOGNOTE
writes periodic notes to the log describing the current status of computations. It is designed for use with analyses requiring extensive CPU resources.

MAXFUNC=number
specifies the maximum number of likelihood evaluations in the optimization process. The default is 150.
**PROC MIXED Statement**

**MAXITER=** *number*

specifies the maximum number of iterations. The default is 50.

**METHOD=** REML | ML | MIVQUE0 | TYPE1 | TYPE2 | TYPE3

specifies the estimation method for the covariance parameters. The REML specification performs residual (restricted) maximum likelihood, and it is the default method. The ML specification performs maximum likelihood, and the MIVQUE0 specification performs minimum variance quadratic unbiased estimation of the covariance parameters.

The METHOD=TYPE*n* specifications apply only to variance component models with no SUBJECT= effects and no REPEATED statement. An analysis of variance table is included in the output, and the expected mean squares are used to estimate the variance components (see Chapter 44, “The GLM Procedure,” for further explanation). The resulting method-of-moment variance component estimates are used in subsequent calculations, including standard errors computed from ESTIMATE and LSMEANS statements. The ODS table names are Type1, Type2, and Type3, respectively.

**MMEQ**

requests that the coefficient matrix and the right-hand side of the mixed model equations be displayed. If $\mathbf{G}$ is nonsingular, the coefficient matrix and the right-hand side have the following form:

$$
\begin{bmatrix}
X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z \\
Z'\hat{R}^{-1}X & Z'\hat{R}^{-1}Z + \hat{G}^{-1}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta} \\
\hat{y}
\end{bmatrix}
= 
\begin{bmatrix}
X'\hat{R}^{-1}y \\
Z'\hat{R}^{-1}y
\end{bmatrix}
$$

If $\mathbf{G}$ is singular, the coefficient matrix and right-hand side have the following modified form:

$$
\begin{bmatrix}
X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z\hat{G} \\
\hat{G}'Z'\hat{R}^{-1}X & \hat{G}'Z'\hat{R}^{-1}Z\hat{G} + \hat{G}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta} \\
\hat{\tau}
\end{bmatrix}
= 
\begin{bmatrix}
X'\hat{R}^{-1}y \\
\hat{G}'Z'\hat{R}^{-1}y
\end{bmatrix}
$$

See the section “Estimating Fixed and Random Effects in the Mixed Model” on page 5320 for further information about these equations.

**MMEQSOL**

requests that a solution to the mixed model equations be produced, in addition to the inverted coefficients matrix. If $\mathbf{G}$ is nonsingular, the formula is the same as the preceding description of the MMEQ option. If $\mathbf{G}$ is singular, $\hat{\beta}$ and $\hat{G}\hat{\tau}$ are displayed in addition to the inverse of the modified coefficient matrix.

See the section “Estimating Fixed and Random Effects in the Mixed Model” on page 5320 for further information about these equations and solution transformation.

**NAMELEN=** *number>*

specifies the length to which long effect names are shortened. The default and minimum value is 20.

**NOBOUND**

has the same effect as the NOBOUND option in the PARMS statement.

**NOCLPRINT=** *number>*

suppresses the display of the “Class Level Information” table if you do not specify *number*. If you do specify *number*, only levels with totals that are less than *number* are listed in the table.
NOINFO
suppresses the display of the “Model Information,” “Dimensions,” and “Number of Observations” tables.

NOITPRINT
suppresses the display of the “Iteration History” table.

NOPROFILE
includes the residual variance as part of the Newton-Raphson iterations. This option applies only to models that have a residual variance parameter. By default, this parameter is profiled out of the likelihood calculations, except when you have specified the HOLD= option in the PARMS statement.

ORD
displays ordinates of the relevant distribution in addition to p-values. The ordinate can be viewed as an approximate odds ratio of hypothesis probabilities.

ORDER=DATA | FORMATTED | FREQ | INTERNAL
specifies the sort order for the levels of the classification variables (which are specified in the CLASS statement). This option applies to the levels for all classification variables, except when you use the (default) ORDER=FORMATTED option with numeric classification variables that have no explicit format. In that case, the levels of such variables are ordered by their internal value.

The ORDER= option can take the following values:

<table>
<thead>
<tr>
<th>Value of ORDER=</th>
<th>Levels Sorted By</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA</td>
<td>Order of appearance in the input data set</td>
</tr>
<tr>
<td>FORMATTED</td>
<td>External formatted value, except for numeric variables with no explicit format, which are sorted by their unformatted (internal) value</td>
</tr>
<tr>
<td>FREQ</td>
<td>Descending frequency count; levels with the most observations come first in the order</td>
</tr>
<tr>
<td>INTERNAL</td>
<td>Unformatted value</td>
</tr>
</tbody>
</table>

By default, ORDER=FORMATTED. For ORDER=FORMATTED and ORDER=INTERNAL, the sort order is machine-dependent. For more information about sort order, see the chapter on the SORT procedure in the Base SAS Procedures Guide and the discussion of BY-group processing in SAS Language Reference: Concepts.

PLOTS < (global-plot-options ) > < =plot-request < (options ) >>
PLOTS < (global-plot-options ) > < = (plot-request<(options)> < . . . plot-request<(options)> >) >
requests that the MIXED procedure produce statistical graphics via the Output Delivery System, provided that ODS Graphics is enabled.

ODS Graphics must be enabled before plots can be requested. For example:
ods graphics on;
proc mixed data=heights plots=all;
    class Family Gender;
    model Height = Gender / residual;
    random Family Family*Gender;
run;
ods graphics off;

For more information about enabling and disabling ODS Graphics, see the section “Enabling and Disabling ODS Graphics” on page 606 in Chapter 21, “Statistical Graphics Using ODS.”

For examples of the basic statistical graphics produced by the MIXED procedure and aspects of their computation and interpretation, see the section “ODS Graphics” on page 5346.

The global-plot-options apply to all relevant plots generated by the MIXED procedure. The global-plot-options supported by the MIXED procedure follow.

Global Plot Options

OBSNO
uses the data set observation number to identify observations in tooltips, provided that the observation number can be determined. Otherwise, the number displayed in tooltips is the index of the observation as it is used in the analysis within the BY group.

ONLY
suppresses the default plots. Only the plots specifically requested are produced.

UNPACKPANEL
UNPACK
displays each graph separately. (By default, some graphs can appear together in a single panel.)

MAXPOINTS=NONE | number
specifies that plots with elements that require processing more than number points be suppressed. The default is MAXPOINTS=5000. No plots are suppressed if you specify MAXPOINTS=NONE.

Specific Plot Options

The following listing describes the specific plots and their options.

ALL
requests that all plots appropriate for the particular analysis be produced.

BOXPLOT <(boxplot-options)>
requests box plots for the effects in your model that consist of classification effects only. Note that these effects can involve more than one classification variable (interaction and nested effects), but they cannot contain any continuous variables. By default, the BOXPLOT request produces box plots based on (conditional) raw residuals for the qualifying effects in the MODEL, RANDOM,
and `REPEATED` statements. See the discussion of the `boxplot-options` in a later section for information about how to tune your box plot request.

**DISTANCE**< (USEINDEX) >
requests a plot of the likelihood or restricted likelihood distance. When influence diagnostics are requested with set selection according to an effect, the USEINDEX option enables you to replace the formatted tick values on the horizontal axis with integer indices of the effect levels in order to reduce the space taken up by the horizontal plot axis.

**INFLUENCEESTPLOT**< (options) >
requests panels of the deletion estimates in an influence analysis, provided that the INFLUENCE option is specified in the `MODEL` statement. No plots are produced for fixed-effects parameters associated with singular columns in the $X$ matrix or for covariance parameters associated with singularities in the `ASYCOV` matrix. By default, separate panels are produced for the fixed-effects and covariance parameters delete estimates. The FIXED and RANDOM options enable you to select these specific panels. The UNPACK option produces separate plots for each of the parameter estimates. The USEINDEX option replaces formatted tick values for the horizontal axis with integer indices.

**INFLUENCESTATPANEL**< (options) >
requests panels of influence statistics. For iterative influence analysis (see the INFLUENCE option in the `MODEL` statement), the panel shows the Cook’s $D$ and CovRatio statistics for fixed-effects and covariance parameters, enabling you to gauge impact on estimates and precision for both types of estimates. In noniterative analysis, only statistics for the fixed effects are plotted. The UNPACK option produces separate plots from the elements in the panel. The USEINDEX option replaces formatted tick values for the horizontal axis with integer indices.

**RESIDUALPANEL**< (residual-plot-options) >
requests a panel of raw residuals. By default, the conditional residuals are produced. See the discussion of `residual-plot-options` in a later section for information about how to tune this panel.

**STUDENTPANEL**< (residual-plot-options) >
requests a panel of studentized residuals. By default, the conditional residuals are produced. See the discussion of `residual-plot-options` in a later section for information about how to tune this panel.

**PEARSONPANEL**< (residual-plot-options) >
requests a panel of Pearson residuals. By default, the conditional residuals are produced. See the discussion of `residual-plot-options` in a later section for information about how to tune this panel.

**PRESS**< (USEINDEX) >
requests a plot of PRESS residuals or PRESS statistics. These are based on “leave-one-out” or “leave-set-out” prediction of the marginal mean. When influence diagnostics are requested with set selection according to an effect, the USEINDEX option enables you to replace the formatted tick values on the horizontal axis with integer indices of the effect levels in order to reduce the space taken up by the horizontal plot axis.

**VCIRYPANEL**< (residual-plot-options) >
requests a panel of residual graphics based on the scaled residuals. See the VCIRY option in the `MODEL` statement for details about these scaled residuals. Only the UNPACK and BOX options of the `residual-plot-options` are available for this type of residual panel.
NONE
suppresses all plots.

Residual Plot Options

The *residual-plot-options* determine both the composition of the panels and the type of residuals being plotted.

**BOX**
**BOXPLOT**
replaces the inset of summary statistics in the lower-right corner of the panel with a box plot of the residual (the “PROC GLIMMIX look”).

**CONDITIONAL**
**BLUP**
constructs plots from conditional residuals.

**MARGINAL**
**NOBLUP**
constructs plots from marginal residuals.

**UNPACK**
produces separate plots from the elements of the panel. The inset statistics are not part of the unpack operation.

Box Plot Options

The *boxplot-options* determine whether box plots are produced for residuals or for residuals and observed values, and for which model effects the box plots are constructed. The available *boxplot-options* are as follows.

**CONDITIONAL**
**BLUP**
constructs box plots from conditional residuals—that is, residuals using the estimated BLUPs of random effects.

**FIXED**
produces box plots for all fixed effects (MODEL statement) consisting entirely of classification variables

**GROUP**
produces box plots for all GROUP= effects (RANDOM and REPEATED statement) consisting entirely of classification variables
MARGINAL

NOBLUP
constructs box plots from marginal residuals.

NPANEL=number
provides the ability to break a box plot into multiple graphics. If number is negative, no balancing of the number of boxes takes place and number is the maximum number of boxes per graphic. If number is positive, the number of boxes per graphic is balanced. For example, suppose variable A has 125 levels, and consider the following statements:

ods graphics on;
proc mixed plots=boxplot(npanel=20);
   class A;
   model y = A;
run;

The box balancing results in six plots with 18 boxes each and one plot with 17 boxes. If number is zero, and this is the default, all levels of the effect are displayed in a single plot.

OBSERVED
adds box plots of the observed data for the selected effects.

RANDOM
produces box plots for all random effects (RANDOM statement) consisting entirely of classification variables. This does not include effects specified in the GROUP= or SUBJECT= options of the RANDOM statement.

REPEATED
produces box plots for the repeated effects (REPEATED statement). This does not include effects specified in the GROUP= or SUBJECT= options of the REPEATED statement.

STUDENT
constructs box plots from studentized residuals rather than from raw residuals.

SUBJECT
produces box plots for all SUBJECT= effects (RANDOM and REPEATED statement) consisting entirely of classification variables.

USEINDEX
uses as the horizontal axis label the index of the effect level rather than the formatted value(s). For classification variables with many levels or model effects that involve multiple classification variables, the formatted values identifying the effect levels can take up too much space as axis tick values, leading to extensive thinning. The USEINDEX option replaces tick values constructed from formatted values with the internal level number.
Multiple Plot Requests

You can list a plot request one or more times with different options. For example, the following statements request a panel of marginal raw residuals, individual plots generated from a panel of the conditional raw residuals, and a panel of marginal studentized residuals:

```sas
ods graphics on;
proc mixed plots(only)=;
   ResidualPanel(marginal)
   ResidualPanel(unpack conditional)
   StudentPanel(marginal box));
```

The inset of residual statistics is replaced in this last panel by a box plot of the studentized residuals. Similarly, if you specify the INFLUENCE option in the MODEL statement, then the following statements request statistical graphics of fixed-effects deletion estimates (in a panel), covariance parameter deletion estimates (unpacked in individual plots), and box plots for the SUBJECT= and fixed classification effects based on residuals and observed values:

```sas
ods graphics on / imagefmt=staticmap;
proc mixed plots(only)=;
   InfluenceEstPlot(fixed)
   InfluenceEstPlot(random unpack)
   BoxPlot(observed fixed subject);
```

The STATICMAP image format enables tooltips that show, for example, values of influence diagnostics associated with a particular delete estimate.

This concludes the syntax section for the PLOTS= option in the PROC MIXED statement.

**RANKS**

- displays the ranks of design matrices $X$ and $(XZ)$.

**RATIO**

- produces the ratio of the covariance parameter estimates to the estimate of the residual variance when the latter exists in the model.

**RIDGE**

- specifies the starting value for the minimum ridge value used in the Newton-Raphson algorithm. The default is 0.3125.

**SCORING<=number>**

- requests that Fisher scoring be used in association with the estimation method up to iteration `number`, which is 0 by default. When you use the SCORING= option and PROC MIXED converges without stopping the scoring algorithm, PROC MIXED uses the expected Hessian matrix to compute approximate standard errors for the covariance parameters instead of the observed Hessian. The output from the ASYCOV and ASYCORR options is similarly adjusted.
SIGITER is an alias for the NOPROFILE option.

UPDATE is an alias for the LOGNOTE option.

---

### BY Statement

**BY variables ;**

You can specify a BY statement with PROC MIXED to obtain separate analyses of observations in groups that are defined by the BY variables. When a BY statement appears, the procedure expects the input data set to be sorted in order of the BY variables. If you specify more than one BY statement, only the last one specified is used.

If your input data set is not sorted in ascending order, use one of the following alternatives:

- Sort the data by using the SORT procedure with a similar BY statement.
- Specify the NOTSORTED or DESCENDING option in the BY statement for the MIXED procedure. The NOTSORTED option does not mean that the data are unsorted but rather that the data are arranged in groups (according to values of the BY variables) and that these groups are not necessarily in alphabetical or increasing numeric order.
- Create an index on the BY variables by using the DATASETS procedure (in Base SAS software).

Because sorting the data changes the order in which PROC MIXED reads observations, the sort order for the levels of the CLASS variable might be affected if you have specified ORDER=DATA in the PROC MIXED statement. This, in turn, affects specifications in the CONTRAST or ESTIMATE statement.

For more information about BY-group processing, see the discussion in *SAS Language Reference: Concepts*. For more information about the DATASETS procedure, see the discussion in the *Base SAS Procedures Guide*.

---

### CLASS Statement

**CLASS variable <(REF= option)> ... <variable <(REF= option)>> <global-options> ;**

The CLASS statement names the classification variables to be used in the model. Typical classification variables are Treatment, Sex, Race, Group, and Replication. If you use the CLASS statement, it must appear before the MODEL statement.

Classification variables can be either character or numeric. By default, class levels are determined from the entire set of formatted values of the CLASS variables.

**NOTE:** Prior to SAS 9, class levels were determined by using no more than the first 16 characters of the formatted values. To revert to this previous behavior, you can use the TRUNCATE option in the CLASS statement.
In any case, you can use formats to group values into levels. See the discussion of the FORMAT procedure in the Base SAS Procedures Guide and the discussions of the FORMAT statement and SAS formats in SAS Formats and Informats: Reference. You can adjust the order of CLASS variable levels with the ORDER= option in the PROC MIXED statement. You can specify the following REF= option to indicate how the levels of an individual classification variable are to be ordered by enclosing it in parentheses after the variable name:

**REF='level' | FIRST | LAST**

specifies a level of the classification variable to be put at the end of the list of levels. This level thus corresponds to the reference level in the usual interpretation of the estimates with PROC MIXED’s singular parameterization. You can specify the level of the variable to use as the reference level; specify a value that corresponds to the formatted value of the variable if a format is assigned. Alternatively, you can specify REF=FIRST to designate that the first ordered level serve as the reference, or REF=LAST to designate that the last ordered level serve as the reference. To specify that REF=FIRST or REF=LAST be used for all classification variables, use the REF= global-option after the slash (/) in the CLASS statement.

You can specify the following global-options in the CLASS statement after a slash (/):

**REF=FIRST | LAST**

specifies a level of all classification variables to be put at the end of the list of levels. This level thus corresponds to the reference level in the usual interpretation of the estimates with PROC MIXED’s singular parameterization. Specify REF=FIRST to designate that the first ordered level for each classification variable serve as the reference. Specify REF=LAST to designate that the last ordered level serve as the reference. This option applies to all the variables specified in the CLASS statement. To specify different reference levels for different classification variables, use REF= options for individual variables.

**TRUNCATE**

specifies that class levels be determined by using only up to the first 16 characters of the formatted values of CLASS variables. When formatted values are longer than 16 characters, you can use this option to revert to the levels as determined in releases prior to SAS 9.

---

**CODE Statement**

```sas
CODE < options > ;
```

The CODE statement writes SAS DATA step code for computing predicted values of the fitted model either to a file or to a catalog entry. This code can then be included in a DATA step to score new data.

Table 63.4 summarizes the options available in the CODE statement.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CATALOG=</td>
<td>Names the catalog entry where the generated code is saved</td>
</tr>
<tr>
<td>DUMMIES</td>
<td>Retains the dummy variables in the data set</td>
</tr>
<tr>
<td>ERROR</td>
<td>Computes the error function</td>
</tr>
<tr>
<td>FILE=</td>
<td>Names the file where the generated code is saved</td>
</tr>
</tbody>
</table>

---
Table 63.4  continued

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FORMAT=</td>
<td>Specifies the numeric format for the regression coefficients</td>
</tr>
<tr>
<td>GROUP=</td>
<td>Specifies the group identifier for array names and statement labels</td>
</tr>
<tr>
<td>IMPUTE</td>
<td>Imputes predicted values for observations with missing or invalid covariates</td>
</tr>
<tr>
<td>LINESIZE=</td>
<td>Specifies the line size of the generated code</td>
</tr>
<tr>
<td>LOOKUP=</td>
<td>Specifies the algorithm for looking up CLASS levels</td>
</tr>
<tr>
<td>RESIDUAL</td>
<td>Computes residuals</td>
</tr>
</tbody>
</table>

For details about the syntax of the CODE statement, see the section “CODE Statement” on page 395 in Chapter 19, “Shared Concepts and Topics.”

**CONTRAST Statement**

```plaintext
CONTRAST 'label' < fixed-effect values ... >
   < | random-effect values ... >, ...< / options >;
```

The CONTRAST statement provides a mechanism for obtaining custom hypothesis tests. It is patterned after the CONTRAST statement in PROC GLM, although it has been extended to include random effects. This enables you to select an appropriate inference space (McLean, Sanders, and Stroup 1991).

You can test the hypothesis $L\phi = 0$, where $L' = (K'M')$ and $\phi' = (\beta' \gamma')$, in several inference spaces. The inference space corresponds to the choice of $M$. When $M = 0$, your inferences apply to the entire population from which the random effects are sampled; this is known as the broad inference space. When all elements of $M$ are nonzero, your inferences apply only to the observed levels of the random effects. This is known as the narrow inference space, and you can also choose it by specifying all of the random effects as fixed. The GLM procedure uses the narrow inference space. Finally, by setting to zero the portions of $M$ corresponding to selected main effects and interactions, you can choose intermediate inference spaces. The broad inference space is usually the most appropriate, and it is used when you do not specify any random effects in the CONTRAST statement.

The CONTRAST statement has the following arguments:

- **label** identifies the contrast in the table. A label is required for every contrast specified. Labels can be up to 200 characters and must be enclosed in quotes.
- **fixed-effect** identifies an effect that appears in the MODEL statement. The keyword INTERCEPT can be used as an effect when an intercept is fitted in the model. You do not need to include all effects that are in the MODEL statement.
- **random-effect** identifies an effect that appears in the RANDOM statement. The first random effect must follow a vertical bar (|); however, random effects do not have to be specified.
- **values** are constants that are elements of the $L$ matrix associated with the fixed and random effects.

The rows of $L'$ are specified in order and are separated by commas. The rows of the $K'$ component of $L'$ are specified on the left side of the vertical bars (|). These rows test the fixed effects and are, therefore, checked
for estimability. The rows of the $M'$ component of $L'$ are specified on the right side of the vertical bars. They test the random effects, and no estimability checking is necessary.

If PROC MIXED finds the fixed-effects portion of the specified contrast to be nonestimable (see the SINGULAR= option), then it displays a message in the log.

The following CONTRAST statement reproduces the $F$ test for the effect $A$ in the split-plot example (see Example 63.1):

```plaintext
contrast 'A broad'
  A  1  -1  0
  A*B .5 .5  -.5  -.5  0  0 ,
  A  1  0  -1
  A*B .5 .5  0  0  -.5  -.5 / df=6;
```

Note that no random effects are specified in the preceding contrast; thus, the inference space is broad. The resulting $F$ test has two numerator degrees of freedom because $L'$ has two rows. The denominator degrees of freedom is, by default, the residual degrees of freedom (9), but the DF= option changes the denominator degrees of freedom to 6.

The following CONTRAST statement reproduces the $F$ test for $A$ when Block and $A*Block$ are considered fixed effects (the narrow inference space):

```plaintext
contrast 'A narrow'
  A  1  -1  0
  A*B .5 .5  -.5  -.5  0  0 |
  A*Block .25 .25 .25 .25  -.25  -.25  -.25  -.25  0  0  0  0 ,
  A  1  0  -1
  A*B .5 .5  0  0  -.5  -.5 |
  A*Block .25 .25 .25 .25  0  0  0  0  -.25  -.25  -.25  -.25 ;
```

The preceding contrast does not contain coefficients for $B$ and Block, because they cancel out in estimated differences between levels of $A$. Coefficients for $B$ and Block are necessary to estimate the mean of one of the levels of $A$ in the narrow inference space (see Example 63.1).

If the elements of $L$ are not specified for an effect that contains a specified effect, then the elements of the specified effect are automatically “filled in” over the levels of the higher-order effect. This feature is designed to preserve estimability for cases where there are complex higher-order effects. The coefficients for the higher-order effect are determined by equitably distributing the coefficients of the lower-level effect, as in the construction of least squares means. In addition, if the intercept is specified, it is distributed over all classification effects that are not contained by any other specified effect. If an effect is not specified and does not contain any specified effects, then all of its coefficients in $L$ are set to 0. You can override this behavior by specifying coefficients for the higher-order effect.

If too many values are specified for an effect, the extra ones are ignored; if too few are specified, the remaining ones are set to 0. If no random effects are specified, the vertical bar can be omitted; otherwise, it must be present. If a SUBJECT= effect is used in the RANDOM statement, then the coefficients specified for the effects in the RANDOM statement are equitably distributed across the levels of the SUBJECT effect. You can use the E option to see exactly which $L$ matrix is used.

The SUBJECT and GROUP options in the CONTRAST statement are useful for the case when a SUBJECT= or GROUP= variable appears in the RANDOM statement, and you want to contrast different subjects or
groups. By default, CONTRAST statement coefficients on random effects are distributed equally across subjects and groups.

PROC MIXED handles missing level combinations of classification variables similarly to the way PROC GLM does. Both procedures delete fixed-effects parameters corresponding to missing levels in order to preserve estimability. However, PROC MIXED does not delete missing level combinations for random-effects parameters because linear combinations of the random-effects parameters are always estimable. These conventions can affect the way you specify your CONTRAST coefficients.

The CONTRAST statement computes the statistic

\[
F = \frac{\left[ \hat{\beta} \right]' \left( L' \hat{C} L \right)^{-1} L' \left[ \hat{\beta} \right]}{r}
\]

where \( r = \text{rank}(L' \hat{C} L) \), and approximates its distribution with an \( F \) distribution. In this expression, \( \hat{C} \) is an estimate of the generalized inverse of the coefficient matrix in the mixed model equations. See the section “Inference and Test Statistics” on page 5322 for more information about this \( F \) statistic.

The numerator degrees of freedom in the \( F \) approximation are \( r = \text{rank}(L' \hat{C} L) \), and the denominator degrees of freedom are taken from the “Tests of Fixed Effects” table and corresponds to the final effect you list in the CONTRAST statement. You can change the denominator degrees of freedom by using the \( \text{DF=} \) option.

You can specify the following \textit{options} in the CONTRAST statement after a slash (/).

**CHISQ**

requests that chi-square tests be performed in addition to any \( F \) tests. A chi-square statistic equals its corresponding \( F \) statistic times the associate numerator degrees of freedom, and the same degrees of freedom are used to compute the \( p \)-value for the chi-square test. This \( p \)-value is always less than that for the \( F \) -test, as it effectively corresponds to an \( F \) test with infinite denominator degrees of freedom.

**DF=\text{number}**

specifies the denominator degrees of freedom for the \( F \) test. The default is the denominator degrees of freedom taken from the “Tests of Fixed Effects” table and corresponds to the final effect you list in the CONTRAST statement.

**E**

requests that the \( L \) matrix coefficients for the contrast be displayed. The ODS name of the “\( L \) Matrix Coefficients” table is Coef.

**GROUP \texttt{coeffs}**

**GRP \texttt{coeffs}**

sets up random-effect contrasts between different groups when a \texttt{GROUP=} variable appears in the \texttt{RANDOM} statement. By default, CONTRAST statement coefficients on random effects are distributed equally across groups.

**SINGULAR=\text{number}**

 tunes the estimability checking. If \( v \) is a vector, define \( \text{ABS}(v) \) to be the absolute value of the element of \( v \) with the largest absolute value. If \( \text{ABS}(K' - K'T) \) is greater than \( c^*\text{\texttt{\text{number}}} \) for any row of \( K' \) in the contrast, then \( K \) is declared nonestimable. Here \( T \) is the Hermite form matrix \((X'X)^{-1}X'X\), and \( c \) is \( \text{ABS}(K') \) except when it equals 0, and then \( c \) is 1. The value for \texttt{\text{number}} must be between 0 and 1; the default is 1E–4.
sets up random-effect contrasts between different subjects when a SUBJECT= variable appears in the RANDOM statement. By default, CONTRAST statement coefficients on random effects are distributed equally across subjects.

**SUBJECT coeffs**

**SUB coeffs**

The ESTIMATE statement is exactly like a CONTRAST statement, except only one-row \( L \) matrices are permitted. The actual estimate, \( L \hat{\beta} \), is displayed along with its approximate standard error. An approximate \( t \) test that \( L \hat{\beta} = 0 \) is also produced.

PROC MIXED selects the degrees of freedom to match those displayed in the “Tests of Fixed Effects” table for the final effect you list in the ESTIMATE statement. You can modify the degrees of freedom by using the DF= option.

If PROC MIXED finds the fixed-effects portion of the specified estimate to be nonestimable, then it displays “Non-est” for the estimate entries.

The following examples of ESTIMATE statements compute the mean of the first level of \( A \) in the split-plot example (see Example 63.1) for various inference spaces:

```plaintext
estimate 'A1 mean narrow' intercept 1
   A 1 B .5 .5 A*B .5 .5 |
   block .25 .25 .25 .25
   A*Block .25 .25 .25 .25
   0 0 0 0
   0 0 0 0;

estimate 'A1 mean intermed' intercept 1
   A 1 B .5 .5 A*B .5 .5 |
   Block .25 .25 .25 .25;

estimate 'A1 mean broad' intercept 1
   A 1 B .5 .5 A*B .5 .5;
```

The construction of the \( L \) vector for an ESTIMATE statement follows the same rules as listed under the CONTRAST statement.

Table 63.5 summarizes the options available in the ESTIMATE statement.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALPHA=</td>
<td>Specifies the confidence level</td>
</tr>
<tr>
<td>CL</td>
<td>Constructs ( t )-type confidence limits</td>
</tr>
<tr>
<td>DF=</td>
<td>Specifies the degrees of freedom</td>
</tr>
<tr>
<td>DIVISOR=</td>
<td>Specifies a value by which to divide all coefficients</td>
</tr>
<tr>
<td>E</td>
<td>Displays the ( L ) matrix coefficients</td>
</tr>
<tr>
<td>GROUP</td>
<td>Sets up random-effect contrasts between different groups</td>
</tr>
</tbody>
</table>
Table 63.5  continued

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOWER</td>
<td>Performs lower-tailed tests</td>
</tr>
<tr>
<td>SINGULAR=</td>
<td>Tunes the estimability checking</td>
</tr>
<tr>
<td>SUBJECT</td>
<td>Sets up random-effect contrasts between different subjects</td>
</tr>
<tr>
<td>UPPER</td>
<td>Performs upper-tailed tests</td>
</tr>
</tbody>
</table>

You can specify the following options in the ESTIMATE statement after a slash (/).

**ALPHA=number**
requests that a $t$-type confidence interval be constructed with confidence level $1 - \text{number}$. The value of \text{number} must be between 0 and 1; the default is 0.05.

**CL**
requests that $t$-type confidence limits be constructed. The confidence level is 0.95 by default; this can be changed with the **ALPHA=** option.

**DF=number**
specifies the degrees of freedom for the $t$ test and confidence limits. The default is the denominator degrees of freedom taken from the “Tests of Fixed Effects” table and corresponds to the final effect you list in the ESTIMATE statement.

**DIVISOR=number**
specifies a value by which to divide all coefficients so that fractional coefficients can be entered as integer numerators.

**E**
requests that the L matrix coefficients be displayed. The ODS name of this “L Matrix Coefficients” table is “Coef.”

**GROUP coeffs**

**GRP coeffs**
sets up random-effect contrasts between different groups when a **GROUP=** variable appears in the **RANDOM** statement. By default, ESTIMATE statement coefficients on random effects are distributed equally across groups.

**LOWER**

**LOWERTAILED**
requests that the $p$-value for the $t$ test be based only on values less than the $t$ statistic. A two-tailed test is the default. A lower-tailed confidence limit is also produced if you specify the **CL** option.

**SINGULAR=number**
tunes the estimability checking as documented for the **SINGULAR=** option in the **CONTRAST** statement.
SUBJECT coeffs
SUB coeffs
sets up random-effect contrasts between different subjects when a SUBJECT= variable appears in the RANDOM statement. By default, ESTIMATE statement coefficients on random effects are distributed equally across subjects. For example, the ESTIMATE statement in the following code from Example 63.5 constructs the difference between the random slopes of the first two batches.

```plaintext
proc mixed data=rc;
  class batch;
  model y = month / s;
  random int month / type=un sub=batch s;
  estimate 'slope b1 - slope b2' | month 1 / subject 1 -1;
run;
```

UPPER
UPPERTAILED
requests that the p-value for the t test be based only on values greater than the t statistic. A two-tailed test is the default. An upper-tailed confidence limit is also produced if you specify the CL option.

---

**ID Statement**

```
ID variables ;
```

The ID statement specifies which variables from the input data set are to be included in the OUTP= and OUTPM= data sets from the MODEL statement. If you do not specify an ID statement, then all variables are included in these data sets. Otherwise, only the variables you list in the ID statement are included. Specifying an ID statement with no variables prevents any variables from being included in these data sets.

**LSMEANS Statement**

```
LSMEANS fixed-effects </options> ;
```

The LSMEANS statement computes least squares means (LS-means) of fixed effects. As in the GLM procedure, LS-means are predicted population margins—that is, they estimate the marginal means over a balanced population. In a sense, LS-means are to unbalanced designs as class and subclass arithmetic means are to balanced designs. The L matrix constructed to compute them is the same as the L matrix formed in PROC GLM; however, the standard errors are adjusted for the covariance parameters in the model.

Each LS-mean is computed as $L\hat{\beta}$, where $L$ is the coefficient matrix associated with the least squares mean and $\hat{\beta}$ is the estimate of the fixed-effects parameter vector (see the section “Estimating Fixed and Random Effects in the Mixed Model” on page 5320). The approximate standard errors for the LS-mean is computed as the square root of $L(XX^T)^{-1}L'$.

LS-means can be computed for any effect in the MODEL statement that involves CLASS variables. You can specify multiple effects in one LSMEANS statement or in multiple LSMEANS statements, and all LSMEANS statements must appear after the MODEL statement. As in the ESTIMATE statement, the L
matrix is tested for estimability, and if this test fails, PROC MIXED displays “Non-est” for the LS-means entries.

Assuming the LS-mean is estimable, PROC MIXED constructs an approximate $t$ test to test the null hypothesis that the associated population quantity equals zero. By default, the denominator degrees of freedom for this test are the same as those displayed for the effect in the “Tests of Fixed Effects” table (see the section “Default Output” on page 5337).

Table 63.6 summarizes the options available in the LSMEANS statement. All LSMEANS options are subsequently discussed in alphabetical order.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construction and Computation of LS-Means</strong></td>
<td></td>
</tr>
<tr>
<td>AT</td>
<td>Modifies covariate value in computing LS-means</td>
</tr>
<tr>
<td>BYLEVEL</td>
<td>Computes separate margins</td>
</tr>
<tr>
<td>DIFF</td>
<td>Requests differences of LS-means</td>
</tr>
<tr>
<td>OM</td>
<td>Specifies weighting scheme for LS-mean computation</td>
</tr>
<tr>
<td>SINGULAR=</td>
<td>Tunes estimability checking</td>
</tr>
<tr>
<td>SLICE=</td>
<td>Partitions $F$ tests (simple effects)</td>
</tr>
<tr>
<td><strong>Degrees of Freedom and $p$-values</strong></td>
<td></td>
</tr>
<tr>
<td>ADJDFE=</td>
<td>Determines whether to compute row-wise denominator degrees of freedom with DDFM=SATTERTHWAITE or DDFM=KENWARDROGER</td>
</tr>
<tr>
<td>ADJUST=</td>
<td>Determines the method for multiple comparison adjustment of LS-mean differences</td>
</tr>
<tr>
<td>ALPHA=$\alpha$</td>
<td>Determines the confidence level ($1 - \alpha$)</td>
</tr>
<tr>
<td>DF=</td>
<td>Assigns specific value to degrees of freedom for tests and confidence limits</td>
</tr>
<tr>
<td><strong>Statistical Output</strong></td>
<td></td>
</tr>
<tr>
<td>CL</td>
<td>Constructs confidence limits for means and or mean differences</td>
</tr>
<tr>
<td>CORR</td>
<td>Displays correlation matrix of LS-means</td>
</tr>
<tr>
<td>COV</td>
<td>Displays covariance matrix of LS-means</td>
</tr>
<tr>
<td>E</td>
<td>Prints the $L$ matrix</td>
</tr>
</tbody>
</table>

You can specify the following options in the LSMEANS statement after a slash (/).

**ADJDFE=SOURCE | ROW**

specifies how denominator degrees of freedom are determined when $p$-values and confidence limits are adjusted for multiple comparisons with the ADJUST= option. When you do not specify the ADJDFE= option, or when you specify ADJDFE=S, the denominator degrees of freedom for multiplicity-adjusted results are the denominator degrees of freedom for the LS-mean effect in the “Type 3 Tests of Fixed Effects” table. When you specify ADJDFE=ROW, the denominator degrees of freedom for multiplicity-adjusted results correspond to the degrees of freedom displayed in the DF column of the “Differences of Least Squares Means” table.
The ADJDFE=ROW setting is particularly useful if you want multiplicity adjustments to take into account that denominator degrees of freedom are not constant across LS-mean differences. This can be the case, for example, when the DDFM=SATTERTHWAITE or DDFM=KENWARDROGER degrees-of-freedom method is in effect.

In one-way models with heterogeneous variance, combining certain ADJUST= options with the ADJDFE=ROW option corresponds to particular methods of performing multiplicity adjustments in the presence of heteroscedasticity. For example, the following statements fit a heteroscedastic one-way model and perform Dunnett’s T3 method (Dunnett 1980), which is based on the studentized maximum modulus (ADJUST=SMM):

```
proc mixed;
  class A;
  model y = A / ddfm=satterth;
  repeated / group=A;
  lsmeans A / adjust=smm adjdfe=row;
run;
```

If you combine the ADJDFE=ROW option with ADJUST=SIDAK, the multiplicity adjustment corresponds to the T2 method of Tamhane (1979), while ADJUST=TUKEY corresponds to the method of Games-Howell (Games and Howell 1976). Note that ADJUST=TUKEY gives the exact results for the case of fractional degrees of freedom in the one-way model, but it does not take into account that the degrees of freedom are subject to variability. A more conservative method, such as ADJUST=SMM, might protect the overall error rate better.

Unless the ADJUST= option of the LSMEANS statement is specified, the ADJDFE= option has no effect.

**ADJUST=BON**
**ADJUST=DUNNETT**
**ADJUST=SCHEFFE**
**ADJUST=SIDAK**
**ADJUST=SIMULATE< (sim-options)>**
**ADJUST=SMM | GT2**
**ADJUST=TUKEY**

requests a multiple comparison adjustment for the p-values and confidence limits for the differences of LS-means. By default, PROC MIXED adjusts all pairwise differences unless you specify ADJUST=DUNNETT, in which case PROC MIXED analyzes all differences with a control level. The ADJUST= option implies the DIFF option.

The BON (Bonferroni) and SIDAK adjustments involve correction factors described in Chapter 44, “The GLM Procedure,” and Chapter 65, “The MULTTEST Procedure;” also see Westfall and Young (1993) and Westfall et al. (1999). When you specify ADJUST=TUKEY and your data are unbalanced, PROC MIXED uses the approximation described in Kramer (1956). Similarly, when you specify ADJUST=DUNNETT and the LS-means are correlated, PROC MIXED uses the factor-analytic covariance approximation described in Hsu (1992). The preceding references also describe the SCHEFFE and SMM adjustments.

The SIMULATE adjustment computes adjusted p-values and confidence limits from the simulated distribution of the maximum or maximum absolute value of a multivariate t random vector. All
covariance parameters except the residual variance are fixed at their estimated values throughout the simulation, potentially resulting in some underdispersion. The simulation estimates \( q \), the true \((1 - \alpha)\) quantile, where \( 1 - \alpha \) is the confidence coefficient. The default \( \alpha \) is 0.05, and you can change this value with the \texttt{ALPHA=} option in the \texttt{LSMEANS} statement.

The number of samples is set so that the tail area for the simulated \( q \) is within \( \gamma \) of \( 1 - \alpha \) with \( 100(1 - \epsilon)\% \) confidence. In equation form,

\[
P(\left| F(\hat{q}) - (1 - \alpha) \right| \leq \gamma) = 1 - \epsilon
\]

where \( \hat{q} \) is the simulated \( q \) and \( F \) is the true distribution function of the maximum; see Edwards and Berry (1987) for details. By default, \( \gamma = 0.005 \) and \( \epsilon = 0.01 \), placing the tail area of \( \hat{q} \) within 0.005 of 0.95 with 99% confidence. The \texttt{ACC=} and \texttt{EPS=} \texttt{sim-options} reset \( \gamma \) and \( \epsilon \), respectively; the \texttt{NSAMP=} \texttt{sim-option} sets the sample size directly; and the \texttt{SEED=} \texttt{sim-option} specifies an integer used to start the pseudo-random number generator for the simulation. If you do not specify a seed, or if you specify a value less than or equal to zero, the seed is generated from reading the time of day from the computer clock. For additional descriptions of these and other simulation options, see the section “\texttt{LSMEANS Statement}” on page 3461 in Chapter 44, “The GLM Procedure.”

\textbf{ALPHA=}\texttt{number}

requests that a \( t \)-type confidence interval be constructed for each of the LS-means with confidence level \( 1 - \texttt{number} \). The value of \texttt{number} must be between 0 and 1; the default is 0.05.

\textbf{AT variable = value}

\textbf{AT (variable-list)= (value-list)}

\textbf{AT MEANS}

enables you to modify the values of the covariates used in computing LS-means. By default, all covariate effects are set equal to their mean values for computation of standard LS-means. The AT option enables you to assign arbitrary values to the covariates. Additional columns in the output table indicate the values of the covariates.

If there is an effect containing two or more covariates, the AT option sets the effect equal to the product of the individual means rather than the mean of the product (as with standard LS-means calculations). The AT MEANS option sets covariates equal to their mean values (as with standard LS-means) and incorporates this adjustment to crossproducts of covariates.

As an example, consider the following invocation of PROC MIXED:

```plaintext
proc mixed;
  class A;
  model Y = A X1 X2 X1*X2;
  lsmeans A;
  lsmeans A / at means;
  lsmeans A / at X1=1.2;
  lsmeans A / at (X1 X2)=(1.2 0.3);
run;
```

For the first two LSMEANS statements, the LS-means coefficient for \( X1 \) is \( \bar{x}_1 \) (the mean of \( X1 \)) and for \( X2 \) is \( \bar{x}_2 \) (the mean of \( X2 \)). However, for the first LSMEANS statement, the coefficient for \( X1 \cdot X2 \) is \( \bar{x}_1 \cdot \bar{x}_2 \), but for the second LSMEANS statement, the coefficient is \( \bar{x}_1 \cdot \bar{x}_2 \). The third LSMEANS
statement sets the coefficient for X1 equal to 1.2 and leaves it at $\sqrt{2}$ for X2, and the final LSMEANS statement sets these values to 1.2 and 0.3, respectively.

If a WEIGHT variable is present, it is used in processing AT variables. Also, observations with missing dependent variables are included in computing the covariate means, unless these observations form a missing cell and the FULLX option in the MODEL statement is not in effect. You can use the E option in conjunction with the AT option to check that the modified LS-means coefficients are the ones you want.

The AT option is disabled if you specify the BYLEVEL option.

BYLEVEL
requests PROC MIXED to process the OM data set by each level of the LS-mean effect (LSMEANS effect) in question. For more details, see the OM option later in this section.

CL
requests that $t$-type confidence limits be constructed for each of the LS-means. The confidence level is 0.95 by default; this can be changed with the ALPHA= option.

CORR
 displays the estimated correlation matrix of the least squares means as part of the “Least Squares Means” table.

COV
 displays the estimated covariance matrix of the least squares means as part of the “Least Squares Means” table.

DF=number
 specifies the degrees of freedom for the $t$ test and confidence limits. The default is the denominator degrees of freedom taken from the “Tests of Fixed Effects” table corresponding to the LS-means effect unless the DDFM=SATTERTHWAITE or DDFM=KENWARDROGER option is in effect in the MODEL statement. For these DDFM= methods, degrees of freedom are determined separately for each test; see the DDFM= option for more information.

DIFF<=difftype>
PDIFF<=difftype>
requests that differences of the LS-means be displayed. The optional difftype specifies which differences to produce, with possible values being ALL, CONTROL, CONTROLL, and CONTROLU. The difftype ALL requests all pairwise differences, and it is the default. The difftype CONTROL requests the differences with a control, which, by default, is the first level of each of the specified LSMEANS effects.

To specify which levels of the effects are the controls, list the quoted formatted values in parentheses after the keyword CONTROL. For example, if the effects A, B, and C are classification variables, each having two levels, 1 and 2, the following LSMEANS statement specifies the (1,2) level of A*B and the (2,1) level of B*C as controls:

   lsmeans A*B B*C / diff=control('1' '2' '2' '1');
For multiple effects, the results depend upon the order of the list, and so you should check the output to make sure that the controls are correct.

Two-tailed tests and confidence limits are associated with the CONTROL `difftype`. For one-tailed results, use either the CONTROLL or CONTROLU `difftype`. The CONTROLL `difftype` tests whether the noncontrol levels are significantly smaller than the control; the upper confidence limits for the control minus the noncontrol levels are considered to be infinity and are displayed as missing. Conversely, the CONTROLU `difftype` tests whether the noncontrol levels are significantly larger than the control; the upper confidence limits for the noncontrol levels minus the control are considered to be infinity and are displayed as missing.

If you want to perform multiple comparison adjustments on the differences of LS-means, you must specify the ADJUST= option.

The differences of the LS-means are displayed in a table titled “Differences of Least Squares Means.” The ODS table name is Diffs.

`E` requests that the `L` matrix coefficients for all LSMEANS effects be displayed. The ODS name of this “`L` Matrix Coefficients” table is Coef.

`OM=<OM-data-set>`

`OBSMARGINS=<OM-data-set>` specifies a potentially different weighting scheme for the computation of LS-means coefficients. The standard LS-means have equal coefficients across classification effects; however, the OM option changes these coefficients to be proportional to those found in `OM-data-set`. This adjustment is reasonable when you want your inferences to apply to a population that is not necessarily balanced but has the margins observed in `OM-data-set`.

By default, `OM-data-set` is the same as the analysis data set. You can optionally specify another data set that describes the population for which you want to make inferences. This data set must contain all model variables except for the dependent variable (which is ignored if it is present). In addition, the levels of all CLASS variables must be the same as those occurring in the analysis data set. Specifying an `OM-data-set` enables you to construct arbitrarily weighted LS-means.

In computing the observed margins, PROC MIXED uses all observations for which there are no missing or invalid independent variables, including those for which there are missing dependent variables. Also, if `OM-data-set` has a WEIGHT variable, PROC MIXED uses weighted margins to construct the LS-means coefficients. If `OM-data-set` is balanced, the LS-means are unchanged by the OM option.

The BYLEVEL option modifies the observed-margins LS-means. Instead of computing the margins across all of the `OM-data-set`, PROC MIXED computes separate margins for each level of the LSMEANS effect in question. In this case the resulting LS-means are actually equal to raw means for fixed-effects models and certain balanced random-effects models, but their estimated standard errors account for the covariance structure that you have specified. If the AT option is specified, the BYLEVEL option disables it.

You can use the E option in conjunction with either the OM or BYLEVEL option to check that the modified LS-means coefficients are the ones you want. It is possible that the modified LS-means are not estimable when the standard ones are, or vice versa. Nonestimable LS-means are noted as “Non-est” in the output.
**PDIF**

is the same as the DIFF option.

**SINGULAR=** *number*

tunes the estimability checking as documented for the SINGULAR= option in the CONTRAST statement.

**SLICE=** *fixed-effect | (fixed-effects)*

specifies effects by which to partition interaction LSMEANS effects. This can produce what are known as tests of simple effects (Winer 1971). For example, suppose that A*B is significant, and you want to test the effect of A for each level of B. The appropriate LSMEANS statement is as follows:

```
lsmeans A*B / slice=B;
```

This code tests for the simple main effects of A for B, which are calculated by extracting the appropriate rows from the coefficient matrix for the A*B LS-means and by using them to form an F test. See the section “Inference and Test Statistics” for more information about this F test.

The SLICE option produces a table titled “Tests of Effect Slices.” The ODS table name is Slices.

---

**LSMEANS Statement**

```
LSMEANS model-effect < 'label' > values < divisor=n >
< ... 'label' values < divisor=n >>
< / options > ;
```

The LSMEANS statement provides a mechanism for obtaining custom hypothesis tests among least squares means.

Table 63.7 summarizes the options available in the LSMEANS statement.

**Table 63.7** LSMEANS Statement Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construction and Computation of LS-Means</strong></td>
<td></td>
</tr>
<tr>
<td>AT</td>
<td>Modifies covariate values in computing LS-means</td>
</tr>
<tr>
<td>BYLEVEL</td>
<td>Computes separate margins</td>
</tr>
<tr>
<td>DIVISOR=</td>
<td>Specifies a list of values to divide the coefficients</td>
</tr>
<tr>
<td>OM=</td>
<td>Specifies the weighting scheme for LS-means computation as determined by a data set</td>
</tr>
<tr>
<td>SINGULAR=</td>
<td>Tunes estimability checking</td>
</tr>
</tbody>
</table>
### Table 63.7  continued

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Degrees of Freedom and p-values</strong></td>
<td></td>
</tr>
<tr>
<td>ADJUST=</td>
<td>Determines the method for multiple-comparison adjustment of LS-means differences</td>
</tr>
<tr>
<td>ALPHA=α</td>
<td>Determines the confidence level (1 - \alpha)</td>
</tr>
<tr>
<td>LOWER</td>
<td>Performs one-sided, lower-tailed inference</td>
</tr>
<tr>
<td>STEPDOWN</td>
<td>Adjusts multiple-comparison p-values further in a step-down fashion</td>
</tr>
<tr>
<td>TESTVALUE=</td>
<td>Specifies values under the null hypothesis for tests</td>
</tr>
<tr>
<td>UPPER</td>
<td>Performs one-sided, upper-tailed inference</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Statistical Output</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CL</td>
<td>Constructs confidence limits for means and mean differences</td>
</tr>
<tr>
<td>CORR</td>
<td>Displays the correlation matrix of LS-means</td>
</tr>
<tr>
<td>COV</td>
<td>Displays the covariance matrix of LS-means</td>
</tr>
<tr>
<td>E</td>
<td>Prints the L matrix</td>
</tr>
<tr>
<td>ELSM</td>
<td>Prints the K matrix</td>
</tr>
<tr>
<td>JOINT</td>
<td>Produces a joint F or chi-square test for the LS-means and LS-means differences</td>
</tr>
<tr>
<td>SEED=</td>
<td>Specifies the seed for computations that depend on random numbers</td>
</tr>
</tbody>
</table>

For details about the syntax of the `LSMEANS` statement, see the section “LSMEANS Statement” on page 476 in Chapter 19, “Shared Concepts and Topics.”

---

**MODEL Statement**

```plaintext
MODEL dependent = <fixed-effects> <options> ;
```

The MODEL statement names a single dependent variable and the fixed effects, which determine the X matrix of the mixed model (see the section “Parameterization of Mixed Models” on page 5325 for details). The specification of effects is the same as in the GLM procedure; however, unlike PROC GLM, you do not specify random effects in the MODEL statement. The MODEL statement is required.

An intercept is included in the fixed-effects model by default. If no fixed effects are specified, only this intercept term is fit. The intercept can be removed by using the NOINT option.

Table 63.8 summarizes the options available in the MODEL statement. These are subsequently discussed in detail in alphabetical order.

### Table 63.8  Summary of MODEL Statement Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model Building</strong></td>
<td></td>
</tr>
<tr>
<td>NOINT</td>
<td>Excludes fixed-effect intercept from model</td>
</tr>
</tbody>
</table>
Table 63.8  continued

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistical Computations</strong></td>
<td></td>
</tr>
<tr>
<td>ALPHA=(\alpha)</td>
<td>Determines the confidence level ((1 - \alpha)) for fixed effects</td>
</tr>
<tr>
<td>ALPHAP=(\alpha)</td>
<td>Determines the confidence level ((1 - \alpha)) for predicted values</td>
</tr>
<tr>
<td>CHISQ</td>
<td>Requests chi-square tests</td>
</tr>
<tr>
<td>DDF=</td>
<td>Specifies denominator degrees of freedom (list)</td>
</tr>
<tr>
<td>DDFM=</td>
<td>Specifies the method for computing denominator degrees of freedom</td>
</tr>
<tr>
<td>HTYPE=</td>
<td>Selects the type of hypothesis test</td>
</tr>
<tr>
<td>INFLUENCE</td>
<td>Requests influence and case-deletion diagnostics</td>
</tr>
<tr>
<td>NOTEST</td>
<td>Suppresses hypothesis tests for the fixed effects</td>
</tr>
<tr>
<td>OUTP=</td>
<td>Specifies output data set for predicted values and related quantities</td>
</tr>
<tr>
<td>OUTPM=</td>
<td>Specifies output data set for predicted means and related quantities</td>
</tr>
<tr>
<td>RESIDUAL</td>
<td>Adds Pearson-type and studentized residuals to output data sets</td>
</tr>
<tr>
<td>VCIRY</td>
<td>Adds scaled marginal residual to output data sets</td>
</tr>
</tbody>
</table>

| **Statistical Output** |                                                                                     |
| CL                  | Displays confidence limits for fixed-effects parameter estimates                  |
| CORRB               | Displays correlation matrix of fixed-effects parameter estimates                 |
| COVB                | Displays covariance matrix of fixed-effects parameter estimates                  |
| COVBI               | Displays inverse covariance matrix of fixed-effects parameter estimates          |
| E, E1, E2, E3       | Displays L matrix coefficients                                                   |
| INTERCEPT           | Adds a row for the intercept to test tables                                     |
| SOLUTION            | Displays fixed-effects parameter estimates (and scale parameter in GLM models)  |

| **Singularity Tolerances** |                                                                                     |
| SINGCHOL=             | Tunes sensitivity in computing Cholesky roots                                    |
| SINGRES=              | Tunes singularity criterion for residual variance                                |
| SINGULAR=             | Tunes the sensitivity in sweeping                                                |
| ZETA=                 | Tunes the sensitivity in forming Type 3 functions                                 |

You can specify the following *options* in the MODEL statement after a slash (/).

**ALPHA=**\(\text{number}\)

requests that a \(t\)-type confidence interval be constructed for each of the fixed-effects parameters with confidence level \(1 - \text{number}\). The value of \(\text{number}\) must be between 0 and 1; the default is 0.05.

**ALPHAP=**\(\text{number}\)

requests that a \(t\)-type confidence interval be constructed for the predicted values with confidence level \(1 - \text{number}\). The value of \(\text{number}\) must be between 0 and 1; the default is 0.05.

**CHISQ**

requests that chi-square tests be performed for all specified effects in addition to the \(F\) tests. Type 3 tests are the default; you can produce the Type 1 and Type 2 tests by using the HTYPE= option.
requests that t-type confidence limits be constructed for each of the fixed-effects parameter estimates. The confidence level is 0.95 by default; this can be changed with the \texttt{ALPHA=} option.

\texttt{CONTAIN} has the same effect as the \texttt{DDFM=}CONTAIN\ option.

\texttt{CORRB} produces the approximate correlation matrix of the fixed-effects parameter estimates. The ODS name of this table is CorrB.

\texttt{COVB} produces the approximate variance-covariance matrix of the fixed-effects parameter estimates $\hat{\beta}$. By default, this matrix equals $(X'V^{-1}X)^{-}$ and results from sweeping $(X y)'V^{-1}(X y)$ on all but its last pivot and removing the y border. The \texttt{EMPIRICAL} option in the \texttt{PROC MIXED} statement changes this matrix into “empirical sandwich” form. The ODS name of this table is CovB. If the degrees-of-freedom method of Kenward and Roger (1997) is in effect \texttt{(DDFM=}KENWARDROGER\), the COVB matrix changes because the method entails an adjustment of the variance-covariance matrix of the fixed effects by the method proposed by Prasad and Rao (1990); Harville and Jeske (1992). See also Kackar and Harville (1984).

\texttt{COVBI} produces the inverse of the approximate variance-covariance matrix of the fixed-effects parameter estimates. The ODS name of this table is InvCovB.

\texttt{DDF=}value-list enables you to specify your own denominator degrees of freedom for the fixed effects. The \texttt{value-list} specification is a list of numbers or missing values (.) separated by commas. The degrees of freedom should be listed in the order in which the effects appear in the “Tests of Fixed Effects” table. If you want to retain the default degrees of freedom for a particular effect, use a missing value for its location in the list. For example, the following statement assigns 3 denominator degrees of freedom to $A$ and 4.7 to $A*B$, while those for $B$ remain the same:

\begin{verbatim}
model Y = A B A*B / ddf=3,.,4.7;
\end{verbatim}

If you specify \texttt{DDFM=}SATERTHWAITE or \texttt{DDFM=}KENWARDROGER, the DDF= option has no effect.

\texttt{DDFM=} \texttt{CONTAIN} \texttt{BETWITHIN} \texttt{RESIDUAL} \texttt{SATERTHWAITE} \texttt{KENWARDROGER\< (FIRSTORDER) \> \texttt{KENWARDROGER\< (LINEAR) \>}

specifies the method for computing the denominator degrees of freedom for the tests of fixed effects resulting from the \texttt{MODEL}, \texttt{CONTRAST}, \texttt{ESTIMATE}, and \texttt{LSMEANS} statements.

Table 63.9 lists syntax aliases for the degrees-of-freedom methods.
The DDFM=CONTAIN option invokes the containment method to compute denominator degrees of freedom, and it is the default when you specify a RANDOM statement. The containment method is carried out as follows: Denote the fixed effect in question A, and search the RANDOM effect list for the effects that syntactically contain A. For example, the random effect B(A) contains A, but the random effect C does not, even if it has the same levels as B(A).

Among the random effects that contain A, compute their rank contribution to the (X Z) matrix. The DDF assigned to A is the smallest of these rank contributions. If no effects are found, the DDF for A is set equal to the residual degrees of freedom, \( N - \text{rank}(X Z) \). This choice of DDF matches the tests performed for balanced split-plot designs and should be adequate for moderately unbalanced designs.

**CAUTION:** If you have a Z matrix with a large number of columns, the overall memory requirements and the computing time after convergence can be substantial for the containment method. If it is too large, you might want to use the DDFM=BETWITHIN option.

The DDFM=BETWITHIN option is the default for REPEATED statement specifications (with no RANDOM statements). It is computed by dividing the residual degrees of freedom into between-subject and within-subject portions. PROC MIXED then checks whether a fixed effect changes within any subject. If so, it assigns within-subject degrees of freedom to the effect; otherwise, it assigns the between-subject degrees of freedom to the effect (see Schluchter and Elashoff 1990). If there are multiple within-subject effects containing classification variables, the within-subject degrees of freedom are partitioned into components corresponding to the subject-by-effect interactions.

One exception to the preceding method is the case where you have specified no RANDOM statements and a REPEATED statement with the TYPE=UN option. In this case, all effects are assigned the between-subject degrees of freedom to provide for better small-sample approximations to the relevant sampling distributions. DDFM=KENWARDROGER might be a better option to try for this case.

The DDFM=RESIDUAL option performs all tests by using the residual degrees of freedom, \( n - \text{rank}(X) \), where \( n \) is the number of observations.

The DDFM=SATTERTHWAITAE option performs a general Satterthwaite approximation for the denominator degrees of freedom, computed as follows. Suppose \( \theta \) is the vector of unknown parameters in V, and suppose \( C = (X'V^{-1}X)^{-} \), where \(^{-}\) denotes a generalized inverse. Let \( \hat{C} \) and \( \hat{\theta} \) be the corresponding estimates.

Consider the one-dimensional case, and consider \( \ell \) to be a vector defining an estimable linear combination of \( \beta \). The Satterthwaite degrees of freedom for the t statistic

\[
t = \frac{\ell \hat{\beta}}{\sqrt{\ell \hat{C} \ell'}}
\]
is computed as

$$\nu = \frac{2(\ell_0'\hat{C}\ell')^2}{g'Ag}$$

where $g$ is the gradient of $\ell C \ell'$ with respect to $\theta$, evaluated at $\hat{\theta}$, and $A$ is the asymptotic variance-covariance matrix of $\hat{\theta}$ obtained from the second derivative matrix of the likelihood equations.

For the multidimensional case, let $L$ be an estimable contrast matrix and denote the rank of $L\hat{C}L'$ as $q > 1$. The Satterthwaite denominator degrees of freedom for the $F$ statistic

$$F = \frac{\hat{\beta}'L'(L\hat{C}L')^{-1}L\hat{\beta}}{q}$$

are computed by first performing the spectral decomposition $L\hat{C}L' = P'DP$, where $P$ is an orthogonal matrix of eigenvectors and $D$ is a diagonal matrix of eigenvalues, both of dimension $q \times q$. Define $\ell_m$ to be the $m$th row of $PL$, and let

$$v_m = \frac{2(D_m)^2}{g_m'A g_m}$$

where $D_m$ is the $m$th diagonal element of $D$ and $g_m$ is the gradient of $\ell_m C \ell_m'$ with respect to $\theta$, evaluated at $\hat{\theta}$. Then let

$$E = \sum_{m=1}^{q} \frac{v_m}{v_m - 2} I(v_m > 2)$$

where the indicator function eliminates terms for which $v_m \leq 2$. The degrees of freedom for $F$ are then computed as

$$\nu = \frac{2E}{E - q}$$

provided $E > q$; otherwise $\nu$ is set to zero.

This method is a generalization of the techniques described in Giesbrecht and Burns (1985); McLean and Sanders (1988); Fai and Cornelius (1996). The method can also include estimated random effects. In this case, append $\hat{\gamma}$ to $\hat{\beta}$ and change $C$ to be the inverse of the coefficient matrix in the mixed model equations. The calculations require extra memory to hold $c$ matrices that are the size of the mixed model equations, where $c$ is the number of covariance parameters. In the notation of Table 63.29, this is approximately $8q(p + g)(p + g)/2$ bytes. Extra computing time is also required to process these matrices. The Satterthwaite method implemented here is intended to produce an accurate $F$ approximation; however, the results can differ from those produced by PROC GLM. Also, the small sample properties of this approximation have not been extensively investigated for the various models available with PROC MIXED.

The DDFM=KENWARDROGER option performs the degrees of freedom calculations detailed by Kenward and Roger (1997). This approximation involves inflating the estimated variance-covariance matrix of the fixed and random effects by the method proposed by Prasad and Rao (1990) and Harville and Jeske (1992), see also Kackar and Harville (1984). Satterthwaite-type degrees of freedom are then computed based on this adjustment. By default, the observed information matrix of the covariance parameter estimates is used in the calculations. For covariance structures that have nonzero
second derivatives with respect to the covariance parameters, the Kenward-Roger covariance matrix adjustment includes a second-order term. This term can result in standard error shrinkage. Also, the resulting adjusted covariance matrix can then be indefinite and is not invariant under reparameterization. The FIRSTORDER or LINEAR suboption of the DDFM=KENWARDROGER option eliminates the second derivatives from the calculation of the covariance matrix adjustment. The LINEAR suboption is an alias for FIRSTORDER. For the case of scalar estimable functions, the resulting estimator is referred to as the Prasad-Rao estimator in Harville and Jeske (1992). The following are examples of covariance structures that generally lead to nonzero second derivatives: TYPE=ANTE(1), TYPE=AR(1), TYPE=ARH(1), TYPE=ARMA(1,1), TYPE=CSH, TYPE=FA, TYPE=FA0(q), TYPE=TOEPH, TYPE=UNR, and all TYPE=SP() structures.

When the asymptotic variance matrix of the covariance parameters is found to be singular, a generalized inverse is used. Covariance parameters with zero variance then do not contribute to the degrees-of-freedom adjustment for DDFM=SATTERTHWAITE and DDFM=KENWARDROGER, and a message is written to the log.

This method changes output in the following tables (listed in Table 63.26): Contrast, CorrB, CovB, Diffs, Estimates, InvCovB, LSMeans, Slices, SolutionF, SolutionR, Tests1–Tests3. The OUTP= and OUTPM= data sets are also affected.

E requests that Type 1, Type 2, and Type 3 L matrix coefficients be displayed for all specified effects. The ODS name of the table is Coef.

E1 requests that Type 1 L matrix coefficients be displayed for all specified effects. The ODS name of the table is Coef.

E2 requests that Type 2 L matrix coefficients be displayed for all specified effects. The ODS name of the table is Coef.

E3 requests that Type 3 L matrix coefficients be displayed for all specified effects. The ODS name of the table is Coef.

FULLX requests that columns of the X matrix that consist entirely of zeros not be eliminated from X; otherwise, they are eliminated by default. For a column corresponding to a missing cell to be added to X, its particular levels must be present in at least one observation in the analysis data set along with a missing dependent variable. The use of the FULLX option can affect coefficient specifications in the CONTRAST and ESTIMATE statements, as well as covariate coefficients from LSMEANS statements specified with the AT MEANS option.

HTYPE=value-list indicates the type of hypothesis test to perform on the fixed effects. Valid entries for values in the list are 1, 2, and 3; the default value is 3. You can specify several types by separating the values with a comma or a space. The ODS table names are Tests1 for the Type 1 tests, Tests2 for the Type 2 tests, and Tests3 for the Type 3 tests.
INFLUENCE<(influence-options)> 

specifies that influence and case deletion diagnostics are to be computed.

The INFLUENCE option computes influence diagnostics by noniterative or iterative methods. The noniterative diagnostics rely on recomputation formulas under the assumption that covariance parameters or their ratios remain fixed. With the possible exception of a profiled residual variance, no covariance parameters are updated. This is the default behavior because of its computational efficiency. However, the impact of an observation on the overall analysis can be underestimated if its effect on covariance parameters is not assessed. Toward this end, iterative methods can be applied to gauge the overall impact of observations and to obtain influence diagnostics for the covariance parameter estimates.

If you specify the INFLUENCE option without further suboptions, PROC MIXED computes single-case deletion diagnostics and influence statistics for each observation in the data set by updating estimates for the fixed-effects parameter estimates, and also the residual variance, if it is profiled. The EFFECT=, SELECT=, ITER=, SIZE=, and KEEP= suboptions provide additional flexibility in the computation and reporting of influence statistics. Table 63.10 briefly describes important suboptions and their effect on the influence analysis.

<table>
<thead>
<tr>
<th>Description</th>
<th>Suboption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute influence diagnostics for individual observations</td>
<td>Default</td>
</tr>
<tr>
<td>Measure influence of sets of observations chosen according to a classification variable or effect</td>
<td>EFFECT=</td>
</tr>
<tr>
<td>Remove pairs of observations and report the results sorted by degree of influence</td>
<td>SIZE=2</td>
</tr>
<tr>
<td>Remove triples, quadruples of observations, etc.</td>
<td>SIZE=</td>
</tr>
<tr>
<td>Allow selection of individual observations, observations sharing specific levels of effects, and construction of tuples from specified subsets of observations</td>
<td>SELECT=</td>
</tr>
<tr>
<td>Update fixed effects and covariance parameters by refitting the mixed model, adding up to ( n ) iterations</td>
<td>ITER=( n &gt; 0 )</td>
</tr>
<tr>
<td>Compute influence diagnostics for the covariance parameters</td>
<td>ITER=( n &gt; 0 )</td>
</tr>
<tr>
<td>Update only fixed effects and the residual variance, if it is profiled</td>
<td>ITER=0</td>
</tr>
<tr>
<td>Add the reduced-data estimates to the data set created with ODS OUTPUT</td>
<td>ESTIMATES</td>
</tr>
</tbody>
</table>

The modifiers and their default values are discussed in the following paragraphs. The set of computed influence diagnostics varies with the suboptions. The most extensive set of influence diagnostics is obtained when \( \text{ITER}=n \) with \( n > 0 \).

You can produce statistical graphics of influence diagnostics when ODS Graphics is enabled. For general information about ODS Graphics, see Chapter 21, “Statistical Graphics Using ODS.” For specific information about the graphics available in the MIXED procedure, see the section “ODS Graphics” on page 5346.

You can specify the following influence-options in parentheses:
**EFFECT=effect**

specifies an effect according to which observations are grouped. Observations sharing the same level of the effect are removed from the analysis as a group. The effect must contain only classification variables, but they need not be contained in the model.

Removing observations can change the rank of the $(X'V^{-1}X)^{-1}$ matrix. This is particularly likely to happen when multiple observations are eliminated from the analysis. If the rank of the estimated variance-covariance matrix of $\hat{\beta}$ changes or its singularity pattern is altered, no influence diagnostics are computed.

**ESTIMATES**

**EST**

specifies that the updated parameter estimates should be written to the ODS output data set. The values are not displayed in the “Influence” table, but if you use ODS OUTPUT to create a data set from the listing, the estimates are added to the data set. If ITER=0, only the fixed-effects estimates are saved. In iterative influence analyses, fixed-effects and covariance parameters are stored. The $p$ fixed-effects parameter estimates are named Parm1–Parmp, and the $q$ covariance parameter estimates are named CovP1–CovPq. The order corresponds to that in the “Solution for Fixed Effects” and “Covariance Parameter Estimates” tables. If parameter updates fail—for example, because of a loss of rank or a nonpositive definite Hessian—missing values are reported.

**ITER=n**

controls the maximum number of additional iterations PROC MIXED performs to update the fixed-effects and covariance parameter estimates following data point removal. If you specify $n > 0$, then statistics such as DFFITS, MDFFITS, and the likelihood distances measure the impact of observation(s) on all aspects of the analysis. Typically, the influence will grow compared to values at ITER=0. In models without RANDOM or REPEATED effects, the ITER= option has no effect.

This documentation refers to analyses when $n > 0$ simply as iterative influence analysis, even if final covariance parameter estimates can be updated in a single step (for example, when METHOD=MIVQUE0 or METHOD=TYPE3). This nomenclature reflects the fact that only if $n > 0$ are all model parameters updated, which can require additional iterations. If $n > 0$ and METHOD=REML (default) or METHOD=ML, the procedure updates fixed effects and variance-covariance parameters after removing the selected observations with additional Newton-Raphson iterations, starting from the converged estimates for the entire data. The process stops for each observation or set of observations if the convergence criterion is satisfied or the number of further iterations exceeds $n$. If $n > 0$ and METHOD=TYPE1, TYPE2, or TYPE3, ANOVA estimates of the covariance parameters are recomputed in a single step.

Compared to noniterative updates, the computations are more involved. In particular for large data sets or a large number of random effects (or both), iterative updates require considerably more resources. A one-step (ITER=1) or two-step update might be a good compromise. The output includes the number of iterations performed, which is less than $n$ if the iteration converges. If the process does not converge in $n$ iterations, you should be careful in interpreting the results, especially if $n$ is fairly large.

Bounds and other restrictions on the covariance parameters carry over from the full-data model. Covariance parameters that are not iterated in the model fit to the full data (the NOITER or HOLD= option in the PARMS statement) are likewise not updated in the refit. In certain models,
such as random-effects models, the ratios between the covariance parameters and the residual variance are maintained rather than the actual value of the covariance parameter estimate (see the section “Influence Diagnostics” on page 5331).

**KEEP=n**

determines how many observations are retained for display and in the output data set or how many tuples if you specify SIZE=. The output is sorted by an influence statistic as discussed for the SIZE= suboption.

**SELECT=value-list**
specifies which observations or effect levels are chosen for influence calculations. If the SELECT= suboption is not specified, diagnostics are computed as follows:

- for all observations, if EFFECT= or SIZE= are not given
- for all levels of the specified effect, if EFFECT= is specified
- for all tuples of size $k$ formed from the observations in value-list, if SIZE=$k$ is specified

When you specify an effect with the EFFECT= option, the values in value-list represent indices of the levels in the order in which PROC MIXED builds classification effects. Which observations in the data set correspond to this index depends on the order of the variables in the CLASS statement, not the order in which the variables appear in the interaction effect. See the section “Parameterization of Mixed Models” on page 5325 to understand precisely how the procedure indexes nested and crossed effects and how levels of classification variables are ordered. The actual values of the classification variables involved in the effect are shown in the output so you can determine which observations were removed.

If the EFFECT= suboption is not specified, the SELECT= value list refers to the sequence in which observations are read from the input data set or from the current BY group if there is a BY statement. This indexing is not necessarily the same as the observation numbers in the input data set, for example, if a WHERE clause is specified or during BY processing.

**SIZE=n**
instructs PROC MIXED to remove groups of observations formed as tuples of size $n$. For example, SIZE=2 specifies all $n \times (n - 1)/2$ unique pairs of observations. The number of tuples for SIZE=$k$ is $n!/(k!(n-k)!)$ and grows quickly with $n$ and $k$. Using the SIZE= option can result in considerable computing time. The MIXED procedure displays by default only the 50 tuples with the greatest influence. Use the KEEP= option to override this default and to retain a different number of tuples in the listing or ODS output data set. Regardless of the KEEP= specification, all tuples are evaluated and the results are ordered according to an influence statistic. This statistic is the (restricted) likelihood distance as a measure of overall influence if ITER=$n > 0$ or when a residual variance is profiled. When likelihood distances are unavailable, the results are ordered by the PRESS statistic.

To reduce computational burden, the SIZE= option can be combined with the SELECT=value-list modifier. For example, the following statements evaluate all $15 = 6 \times 5/2$ pairs formed from observations 13, 14, 18, 30, 31, and 33 and display the five pairs with the greatest influence:

```plaintext
proc mixed;
  class a m f;
  model penetration = a m /
```
influence(size=2 keep=5 select=13,14,18,30,31,33);
random f(m);
run;

If any observation in a tuple contains missing values or has otherwise not contributed to the analysis, the tuple is not evaluated. This guarantees that the displayed results refer to the same number of observations, so that meaningful statistics are available by which to order the results. If computations fail for a particular tuple—for example, because the \((X'V^{-1}X)^{-}\) matrix changes rank or the \(G\) matrix is not positive definite—no results are produced. Results are retained when the maximum number of iterative updates is exceeded in iterative influence analyses.

The SIZE= suboption cannot be combined with the EFFECT= suboption. As in the case of the EFFECT= suboption, the statistics being computed are those appropriate for removal of multiple data points, even if SIZE=1.

The ODS name of the “Influence Diagnostics” table is Influence. The variables in this table depend on whether you specify the EFFECT=, SIZE=, or KEEP= suboption and whether covariance parameters are iteratively updated. When ITER=0 (the default), certain influence diagnostics are meaningful only if the residual variance is profiled. Table 63.11 and Table 63.12 summarize the statistics obtained depending on the model and modifiers. The last column in these tables gives the variable name in the ODS OUTPUT INFLUENCE= data set. Restricted likelihood distances are reported instead of the likelihood distance unless METHOD=ML. See the section “Influence Diagnostics” on page 5331 for details about the individual statistics.

Table 63.11 Statistics Computed with INFLUENCE Option, Noniterative Analysis (ITER=0)

<table>
<thead>
<tr>
<th>Suboption</th>
<th>(\sigma^2) Profiled</th>
<th>Statistic</th>
<th>Variable Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>Yes</td>
<td>Observed value</td>
<td>Observed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Predicted value</td>
<td>Predicted</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marginal residual</td>
<td>Residual</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Leverage</td>
<td>Leverage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PRESS residual</td>
<td>PRESSRes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Internally studentized marginal residual</td>
<td>Student</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Externally studentized marginal residual</td>
<td>RStudent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE without deleted observations</td>
<td>RMSE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cook’s (D)</td>
<td>CookD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DFFITS</td>
<td>DFFITS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CovRatio</td>
<td>COVRATIO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Restricted) likelihood distance</td>
<td>RLD, LD</td>
</tr>
<tr>
<td>Default</td>
<td>No</td>
<td>Observed value</td>
<td>Observed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Predicted value</td>
<td>Predicted</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marginal residual</td>
<td>Residual</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Leverage</td>
<td>Leverage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PRESS residual</td>
<td>PRESSRes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Internally studentized marginal residual</td>
<td>Student</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cook’s (D)</td>
<td>CookD</td>
</tr>
</tbody>
</table>
### Table 63.11 continued

<table>
<thead>
<tr>
<th>Suboption</th>
<th>Profiled</th>
<th>Statistic</th>
<th>Variable Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFFECT, or KEEP</td>
<td>Yes</td>
<td>Observations in level (tuple)</td>
<td>Nobs</td>
</tr>
<tr>
<td>SIZE</td>
<td>PRESS statistic</td>
<td>PRESS</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>Cook’s D</td>
<td>CookD</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>MDFFITS</td>
<td>MDFFITS</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>CovRatio</td>
<td>COVRATIO</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>COVTRACE</td>
<td>COVTRACE</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>RMSE without deleted level (tuple)</td>
<td>RMSE</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>(Restricted) likelihood distance</td>
<td>RLD, LD</td>
<td></td>
</tr>
</tbody>
</table>

| EFFECT, SIZE, or KEEP | No | Observations in level (tuple) | Nobs |
| EFFECT, SIZE, or KEEP | PRESS statistic | PRESS |
| EFFECT, SIZE, or KEEP | Cook’s D | CookD |

### Table 63.12 Statistics Computed with INFLUENCE Option, Iterative Analysis (ITER=n > 0)

<table>
<thead>
<tr>
<th>Suboption</th>
<th>Statistic</th>
<th>Variable Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>Number of iterations</td>
<td>Iter</td>
</tr>
<tr>
<td>Default</td>
<td>Observed value</td>
<td>Observed</td>
</tr>
<tr>
<td>Default</td>
<td>Predicted value</td>
<td>Predicted</td>
</tr>
<tr>
<td>Default</td>
<td>Marginal residual</td>
<td>Residual</td>
</tr>
<tr>
<td>Default</td>
<td>Leverage</td>
<td>Leverage</td>
</tr>
<tr>
<td>Default</td>
<td>PRESS residual</td>
<td>PRESSres</td>
</tr>
<tr>
<td>Default</td>
<td>Internally studentized marginal residual</td>
<td>Student</td>
</tr>
<tr>
<td>Default</td>
<td>Externally studentized marginal residual</td>
<td>RStudent</td>
</tr>
<tr>
<td>Default</td>
<td>RMSE without deleted obs (if possible)</td>
<td>RMSE</td>
</tr>
<tr>
<td>Default</td>
<td>Cook’s D</td>
<td>CookD</td>
</tr>
<tr>
<td>Default</td>
<td>DFFITS</td>
<td>DFFITS</td>
</tr>
<tr>
<td>Default</td>
<td>CovRatio</td>
<td>COVRATIO</td>
</tr>
<tr>
<td>Default</td>
<td>Cook’s D CovParms</td>
<td>CookDCP</td>
</tr>
<tr>
<td>Default</td>
<td>CovRatio CovParms</td>
<td>COVRATIOCP</td>
</tr>
<tr>
<td>Default</td>
<td>MDFFITS CovParms</td>
<td>MDFFITSCP</td>
</tr>
<tr>
<td>Default</td>
<td>(Restricted) likelihood distance</td>
<td>RLD, LD</td>
</tr>
</tbody>
</table>

| EFFECT, SIZE, or KEEP | Observations in level (tuple) | Nobs |
| EFFECT, SIZE, or KEEP | Number of iterations | Iter |
| EFFECT, SIZE, or KEEP | PRESS statistic | PRESS |
| EFFECT, SIZE, or KEEP | RMSE without deleted level (tuple) | RMSE |
| EFFECT, SIZE, or KEEP | Cook’s D | CookD |
| EFFECT, SIZE, or KEEP | MDFFITS | MDFFITS |
| EFFECT, SIZE, or KEEP | CovRatio | COVRATIO |
| EFFECT, SIZE, or KEEP | COVTRACE | COVTRACE |
| EFFECT, SIZE, or KEEP | Cook’s D CovParms | CookDCP |
### Table 63.12  continued

<table>
<thead>
<tr>
<th>Suboption</th>
<th>Statistic</th>
<th>Variable Name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CovRatio CovParms</td>
<td>COVRATIOCP</td>
</tr>
<tr>
<td></td>
<td>MDFFITS CovParms</td>
<td>MDFFITSCP</td>
</tr>
<tr>
<td></td>
<td>(Restricted) likelihood distance</td>
<td>RLD, LD</td>
</tr>
</tbody>
</table>

**INTERCEPT**

adds a row to the tables for Type 1, 2, and 3 tests corresponding to the overall intercept.

**LCOMPONENTS**

requests an estimate for each row of the \( L \) matrix used to form tests of fixed effects. Components corresponding to Type 3 tests are the default; you can produce the Type 1 and Type 2 component estimates with the HTYPE= option.

Tests of fixed effects involve testing of linear hypotheses of the form \( L\beta = 0 \). The matrix \( L \) is constructed from Type 1, 2, or 3 estimable functions. By default the MIXED procedure constructs Type 3 tests. In many situations, the individual rows of the matrix \( L \) represent contrasts of interest. For example, in a one-way classification model, the Type 3 estimable functions define differences of factor-level means. In a balanced two-way layout, the rows of \( L \) correspond to differences of cell means.

For example, suppose factors \( A \) and \( B \) have \( a \) and \( b \) levels, respectively. The following statements produce \((a - 1)\) one degree of freedom tests for the rows of \( L \) associated with the Type 1 and Type 3 estimable functions for factor \( A \), \((b - 1)\) tests for the rows of \( L \) associated with factor \( B \), and a single test for the Type 1 and Type 3 coefficients associated with regressor \( X \):

```r
class A B;
model y = A B x / htype=1,3 lcomponents;
```

The denominator degrees of freedom associated with a row of \( L \) are the same as those in the corresponding “Tests of Fixed Effects” table, except for DDFM=KENWARDROGER and DDFM=SATTERTHWAITIE. For these degree of freedom methods, the denominator degrees of freedom are computed separately for each row of \( L \).

The ODS name of the table containing all requested component tests is LComponents. See Example 63.9 for applications of the LCOMPONENTS option.

**NOCONTAIN**

has the same effect as the DDFM=RESIDUAL option.

**NOINT**

requests that no intercept be included in the model. An intercept is included by default.

**NOTEST**

specifies that no hypothesis tests be performed for the fixed effects.
OUTP=SAS-data-set
OUTPRED=SAS-data-set

specifies an output data set containing predicted values and related quantities. This option replaces the P option from SAS 6.

Predicted values are formed by using the rows from $(X \ Z)$ as $L$ matrices. Thus, predicted values from the original data are $X\hat{\beta} + Z\hat{y}$. Their approximate standard errors of prediction are formed from the quadratic form of $L$ with $\hat{C}$ defined in the section “Statistical Properties” on page 5321. The $L95$ and $U95$ variables provide a $t$-type confidence interval for the predicted values, and they correspond to the $L95M$ and $U95M$ variables from the GLM and REG procedures for fixed-effects models. The residuals are the observed minus the predicted values. Predicted values for data points other than those observed can be obtained by using missing dependent variables in your input data set.

Specifications that have a REPEATED statement with the SUBJECT= option and missing dependent variables compute predicted values by using empirical best linear unbiased prediction (EBLUP). Using hats ($\hat{}$) to denote estimates, the EBLUP formula is

$$\hat{m} = X_m\hat{\beta} + \hat{C}_m\hat{V}^{-1}(y - X\hat{\beta})$$

where $m$ represents a hypothetical realization of a missing data vector with associated design matrix $X_m$. The matrix $C_m$ is the model-based covariance matrix between $m$ and the observed data $y$, and other notation is as presented in the section “Mixed Models Theory” on page 5313.

The estimated prediction variance is as follows:

$$\text{Var}(\hat{m} - m) = \hat{V}_m - \hat{C}_m\hat{V}^{-1}\hat{C}_m' + [X_m - \hat{C}_m\hat{V}^{-1}X](X'\hat{V}^{-1}X)^{-1}[X_m - \hat{C}_m\hat{V}^{-1}X]'$$

where $V_m$ is the model-based variance matrix of $m$. For further details, see Henderson (1984) and Harville (1990). This feature can be useful for forecasting time series or for computing spatial predictions.

By default, all variables from the input data set are included in the OUTP= data set. You can select a subset of these variables by using the ID statement.

OUTPM=SAS-data-set
OUTPREDM=SAS-data-set

specifies an output data set containing predicted means and related quantities. This option replaces the PM option from SAS 6.

The output data set is of the same form as that resulting from the OUTP= option, except that the predicted values do not incorporate the EBLUP values $Z\hat{y}$. They also do not use the EBLUPs for specifications that have a REPEATED statement with the SUBJECT= option and missing dependent variables. The predicted values are formed as $X\hat{\beta}$ in the OUTPM= data set, and standard errors are quadratic forms in the approximate variance-covariance matrix of $\hat{\beta}$ as displayed by the COVB option.

By default, all variables from the input data set are included in the OUTPM= data set. You can select a subset of these variables by using the ID statement.
requests that Pearson-type and (internally) studentized residuals be added to the OUTP= and OUTPM= data sets. Studentized residuals are raw residuals standardized by their estimated standard error. When residuals are internally studentized, the data point in question has contributed to the estimation of the covariance parameter estimates on which the standard error of the residual is based. Externally studentized marginal residuals can be computed with the INFLUENCE option. Pearson-type residuals scale the residual by the standard deviation of the response.

The option has no effect unless the OUTP= or OUTPM= option is specified or unless ODS Graphics is enabled. For general information about ODS Graphics, see Chapter 21, “Statistical Graphics Using ODS.” For specific information about the graphics available in the MIXED procedure, see the section “ODS Graphics” on page 5346. For computational details about studentized and Pearson residuals in MIXED, see the section “Residual Diagnostics” on page 5330.

SINGCHOL=number

tunes the sensitivity in computing Cholesky roots. If a diagonal pivot element is less than $D \times \text{number}$ as PROC MIXED performs the Cholesky decomposition on a matrix, the associated column is declared to be linearly dependent upon previous columns and is set to 0. The value $D$ is the original diagonal element of the matrix. The default for number is 1E4 times the machine epsilon; this product is approximately 1E–12 on most computers.

SINGRES=number

sets the tolerance for which the residual variance is considered to be zero. The default is 1E4 times the machine epsilon; this product is approximately 1E–12 on most computers.

SINGULAR=number

tunes the sensitivity in sweeping. If a diagonal pivot element is less than $D \times \text{number}$ as PROC MIXED sweeps a matrix, the associated column is declared to be linearly dependent upon previous columns, and the associated parameter is set to 0. The value $D$ is the original diagonal element of the matrix. The default is 1E4 times the machine epsilon; this product is approximately 1E–12 on most computers.

SOLUTION

requests that a solution for the fixed-effects parameters be produced. Using notation from the section “Mixed Models Theory” on page 5313, the fixed-effects parameter estimates are $\hat{\beta}$ and their approximate standard errors are the square roots of the diagonal elements of $(X'\tilde{V}^{-1}X)^{-1}$. You can output this approximate variance matrix with the COVB option or modify it with the EMPIRICAL option in the PROC MIXED statement or the DDFM=KENWARDROGER option in the MODEL statement.

Along with the estimates and their approximate standard errors, a $t$ statistic is computed as the estimate divided by its standard error. The degrees of freedom for this $t$ statistic matches the one appearing in the “Tests of Fixed Effects” table under the effect containing the parameter. The “Pr > |t|” column contains the two-tailed $p$-value corresponding to the $t$ statistic and associated degrees of freedom. You can use the CL option to request confidence intervals for all of the parameters; they are constructed around the estimate by using a radius of the standard error times a percentage point from the $t$ distribution.
VCIRY requests that responses and marginal residuals be scaled by the inverse Cholesky root of the marginal variance-covariance matrix. The variables ScaledDep and ScaledResid are added to the OUTPM= data set. These quantities can be important in bootstrapping of data or residuals. Examination of the scaled residuals is also helpful in diagnosing departures from normality. Notice that the results of this scaling operation can depend on the order in which the MIXED procedure processes the data.

The VCIRY option has no effect unless you also use the OUTPM= option or unless ODS Graphics is enabled. For general information about ODS Graphics, see Chapter 21, “Statistical Graphics Using ODS.” For specific information about the graphics available in the MIXED procedure, see the section “ODS Graphics” on page 5346.

XPVIX is an alias for the COVBI option.

XPVIXI is an alias for the COVB option.

ZETA=number tunes the sensitivity in forming Type 3 functions. Any element in the estimable function basis with an absolute value less than number is set to 0. The default is 1E–8.

---

PARMS Statement

PARMS (value-list)…</options> ;

The PARMS statement specifies initial values for the covariance parameters, or it requests a grid search over several values of these parameters. You must specify the values in the order in which they appear in the “Covariance Parameter Estimates” table.

The value-list specification can take any of several forms:

- \( m \) a single value
- \( m_1, m_2, \ldots, m_n \) several values
- \( m \) to \( n \) a sequence where \( m \) equals the starting value, \( n \) equals the ending value, and the increment equals 1
- \( m \) to \( n \) by \( i \) a sequence where \( m \) equals the starting value, \( n \) equals the ending value, and the increment equals \( i \)
- \( m_1, m_2 \) to \( m_3 \) mixed values and sequences

You can use the PARMS statement to input known parameters. Referring to the split-plot example (Example 63.1), suppose the three variance components are known to be 60, 20, and 6. The SAS statements to fix the variance components at these values are as follows:

```sas
proc mixed data=sp noprofile;
  class Block A B;
  model Y = A B A*B;
  random Block A*Block;
  parms (60) (20) (6) / noiter;
run;
```
The **NOPROFILE** option requests PROC MIXED to refrain from profiling the residual variance parameter during its calculations, thereby enabling its value to be held at 6 as specified in the PARMS statement. The **NOITER** option prevents any Newton-Raphson iterations so that the subsequent results are based on the given variance components. You can also specify known parameters of G by using the **GDATA=** option in the **RANDOM** statement.

If you specify more than one set of initial values, PROC MIXED performs a grid search of the likelihood surface and uses the best point on the grid for subsequent analysis. Specifying a large number of grid points can result in long computing times. The grid search feature is also useful for exploring the likelihood surface. (See Example 63.3.)

The results from the PARMS statement are the values of the parameters on the specified grid (denoted by CovP1–CovPn), the residual variance (possibly estimated) for models with a residual variance parameter, and various functions of the likelihood.

The ODS name of the “Parameter Search” table is ParmSearch.

Table 63.13 summarizes the **options** available in the PARMS statement.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HOLD=</strong></td>
<td>Holds parameter values equal to the specified values</td>
</tr>
<tr>
<td><strong>LOGDETH</strong></td>
<td>Evaluates the log determinant of the Hessian matrix</td>
</tr>
<tr>
<td><strong>LOWERB=</strong></td>
<td>Specifies lower boundary constraints</td>
</tr>
<tr>
<td><strong>NOBOUND</strong></td>
<td>Removes boundary constraints on covariance parameters</td>
</tr>
<tr>
<td><strong>NOITER</strong></td>
<td>Performs inferences using the best value from the grid search</td>
</tr>
<tr>
<td><strong>NOPRINT</strong></td>
<td>Suppresses the “Parameter Search” table</td>
</tr>
<tr>
<td><strong>NOPROFILE</strong></td>
<td>Specifies a different computational method for the residual variance</td>
</tr>
<tr>
<td><strong>PARMSDATA=</strong></td>
<td>Reads in covariance parameter values from a SAS data set</td>
</tr>
<tr>
<td><strong>RATIOS</strong></td>
<td>Indicates that ratios with the residual variance are specified</td>
</tr>
<tr>
<td><strong>UPPERB=</strong></td>
<td>Specifies upper boundary constraints</td>
</tr>
</tbody>
</table>

You can specify the following **options** in the PARMS statement after a slash (/).

**HOLD=** value-list

**EQCONS=** value-list

specifies which parameter values PROC MIXED should hold to equal the specified values. For example, the following statement constrains the first and third covariance parameters to equal 5 and 2, respectively:

```
parms (5) (3) (2) (3) / hold=1,3;
```

**LOGDETH**

evaluates the log determinant of the Hessian matrix for each point specified in the PARMS statement. A Log Det H column is added to the “Parameter Search” table.
LOWEBR=value-list

enables you to specify lower boundary constraints on the covariance parameters. The value-list specification is a list of numbers or missing values (_) separated by commas. You must list the numbers in the order that PROC MIXED uses for the covariance parameters, and each number corresponds to the lower boundary constraint. A missing value instructs PROC MIXED to use its default constraint, and if you do not specify numbers for all of the covariance parameters, PROC MIXED assumes the remaining ones are missing.

An example for which this option is useful is when you want to constrain the $G$ matrix to be positive definite in order to avoid the more computationally intensive algorithms required when $G$ becomes singular. The corresponding statements for a random coefficients model are as follows:

```
proc mixed;
  class person;
  model y = time;
  random int time / type=fa0(2) sub=person;
  parms / lowerb=1e-4,.,1e-4;
run;
```

Here the TYPE=FA0(2) structure is used in order to specify a Cholesky root parameterization for the $2 \times 2$ unstructured blocks in $G$. This parameterization ensures that the $G$ matrix is nonnegative definite, and the PARMS statement then ensures that it is positive definite by constraining the two diagonal terms to be greater than or equal to 1E–4.

**NOBOUND**

requests the removal of boundary constraints on covariance parameters. For example, variance components have a default lower boundary constraint of 0, and the NOBOUND option allows their estimates to be negative.

**NOITER**

requests that no Newton-Raphson iterations be performed and that PROC MIXED use the best value from the grid search to perform inferences. By default, iterations begin at the best value from the PARMS grid search.

**NOPRINT**

suppresses the display of the “Parameter Search” table.

**NOPROFILE**

specifies a different computational method for the residual variance during the grid search. By default, PROC MIXED estimates this parameter by using the profile likelihood when appropriate. This estimate is displayed in the Variance column of the “Parameter Search” table. The NOPROFILE option suppresses the profiling and uses the actual value of the specified variance in the likelihood calculations.

**OLS**

requests starting values corresponding to the usual general linear model. Specifically, all variances and covariances are set to zero except for the residual variance, which is set equal to its ordinary least squares (OLS) estimate. This option is useful when the default MIVQUE0 procedure produces poor starting values for the optimization process.
PARMSDATA=SAS-data-set
PDATA=SAS-data-set
reads in covariance parameter values from a SAS data set. The data set should contain the Est or Covp1–Covpn variables.

RATIOS
indicates that ratios with the residual variance are specified instead of the covariance parameters themselves. The default is to use the individual covariance parameters.

UPPERB=value-list
enables you to specify upper boundary constraints on the covariance parameters. The value-list specification is a list of numbers or missing values (.) separated by commas. You must list the numbers in the order that PROC MIXED uses for the covariance parameters, and each number corresponds to the upper boundary constraint. A missing value instructs PROC MIXED to use its default constraint, and if you do not specify numbers for all of the covariance parameters, PROC MIXED assumes that the remaining ones are missing.

PERFORMANCE Statement

PERFORMANCE < performance-options> ;

The PERFORMANCE statement defines performance parameters for multithreaded computing in solving large mixed model equations and other large matrix sweep operations.

You can specify the following performance-options in the PERFORMANCE statement:

CPUCOUNT=ACTUAL | num
specifies how many processors the procedure assumes to be available on the machine. You can specify any integer from 1 to 256 for num. CPUCOUNT=ACTUAL is equivalent to setting num to be the number of physical processors available. This number can be less than the physical number of CPUs if the SAS process has been restricted by system administration tools. Setting CPUCOUNT= to a number greater than the actual number of available CPUs might result in reduced performance. This option overrides the CPUCOUNT= SAS system option.

NTHREADS=n
specifies the number of threads for analytic computations and overrides the SAS system option THREADS | NOTHREADS. If you do not specify the NTHREADS= option, the number of threads are determined based on the number of CPUs on the machine by creating one thread per CPU for the analytic computations. The largest number that can be specified for n is 256.

By default, PROC MIXED executes in multiple concurrent threads unless you specify the NOTHREADS system option or you force single-threaded execution with NTHREADS=1 or the NOTHREADS options in the PERFORMANCE statement.

THREADS=YES | NO
specifies whether PROC MIXED executes in multiple concurrent threads. The default is YES.

NOTREAD
turns off the multithreaded computation.
PRIOR Statement

PRIOR <distribution> </options>;

The PRIOR statement enables you to carry out a sampling-based Bayesian analysis in PROC MIXED. It currently operates only with variance component models. Other TYPE= structures are not supported. The analysis produces a SAS data set containing a pseudo-random sample from the joint posterior density of the variance components and other parameters in the mixed model.

The posterior analysis is performed after all other PROC MIXED computations. It begins with the “Posterior Sampling Information” table, which provides basic information about the posterior sampling analysis, including the prior densities, sampling algorithm, sample size, and random number seed. The ODS name of this table is Posterior.

By default, PROC MIXED uses an independence chain algorithm in order to generate the posterior sample (Tierney 1994). This algorithm works by generating a pseudo-random proposal from a convenient base distribution, chosen to be as close as possible to the posterior. The proposal is then retained in the sample with probability proportional to the ratio of weights constructed by taking the ratio of the true posterior to the base density. If a proposal is not accepted, then a duplicate of the previous observation is added to the chain.

In selecting the base distribution, PROC MIXED makes use of the fact that the fixed-effects parameters can be analytically integrated out of the joint posterior, leaving the marginal posterior density of the variance components. In order to better approximate the marginal posterior density of the variance components, PROC MIXED transforms them by using the MIVQUE(0) equations. You can display the selected transformation with the PTRANS option or specify your own with the TDATA= option. The density of the transformed parameters is then approximated by a product of inverted gamma densities (see Gelfand et al. 1990).

To determine the parameters for the inverted gamma densities, PROC MIXED evaluates the logarithm of the posterior density over a grid of points in each of the transformed parameters, and you can display the results of this search with the PSEARCH option. PROC MIXED then performs a linear regression of these values on the logarithm of the inverted gamma density. The resulting base densities are displayed in the “Base Densities” table; the ODS name of this table is Base. You can input different base densities with the BDATA= option.

At the end of the sampling, the “Acceptance Rates” table displays the acceptance rate computed as the number of accepted samples divided by the total number of samples generated. The ODS name of the “Acceptance Rates” table is AccRates.

The OUT= option specifies the output data set containing the posterior sample. PROC MIXED automatically includes all variance component parameters in this data set (labeled COVP1–COVPn), the Type 3 F statistics constructed as in Ghosh (1992) discussing Schervish (1992) (labeled T3F_n), the log values of the posterior (labeled LOGF), the log of the base sampling density (labeled LOGG), and the log of their ratio (labeled LOGRATIO). If you specify the SOLUTION option in the MODEL statement, the data set also contains a random sample from the posterior density of the fixed-effects parameters (labeled BETAn); and if you specify the SOLUTION option in the RANDOM statement, the table contains a random sample from the posterior density of the random-effects parameters (labeled GAMn). PROC MIXED also generates additional variables corresponding to any CONTRAST, ESTIMATE, or LSMEANS statement that you specify.

Subsequently, you can use SAS/INSIGHT or the UNIVARIATE, CAPABILITY, or KDE procedure to analyze the posterior sample.
The prior density of the variance components is, by default, a noninformative version of Jeffreys’ prior (Box and Tiao 1973). You can also specify informative priors with the DATA= option or a flat (equal to 1) prior for the variance components. The prior density of the fixed-effects parameters is assumed to be flat (equal to 1), and the resulting posterior is conditionally multivariate normal (conditioning on the variance component parameters) with mean \((X'V^{-1}X)^{-1}X'y\) and variance \((X'V^{-1}X)^{-1}\).

Table 63.14 summarizes the options available in the PRIOR statement.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA=</td>
<td>Inputs the prior densities of the variance components</td>
</tr>
<tr>
<td>JEFFREYS</td>
<td>Specifies a noninformative reference version of Jeffreys’ prior</td>
</tr>
<tr>
<td>FLAT</td>
<td>Specifies a prior density equal to 1 everywhere</td>
</tr>
<tr>
<td>ALG=</td>
<td>Specifies the algorithm used for generating the posterior sample</td>
</tr>
<tr>
<td>BDATA=</td>
<td>Inputs the base densities used by the sampling algorithm</td>
</tr>
<tr>
<td>GRID=</td>
<td>Specifies a grid of values over which to evaluate the posterior density</td>
</tr>
<tr>
<td>GRIDT=</td>
<td>Specifies a transformed grid of values over which to evaluate the posterior density</td>
</tr>
<tr>
<td>IFACTOR=</td>
<td>An alias for the SFACTOR= option</td>
</tr>
<tr>
<td>LOGNOTE=</td>
<td>Writes a note to the log after generating the sample</td>
</tr>
<tr>
<td>LOORBOUND=</td>
<td>Specifies the bounding constant for rejection sampling</td>
</tr>
<tr>
<td>NSAMPLE=</td>
<td>Specifies the number of posterior samples to generate</td>
</tr>
<tr>
<td>NSEARCH=</td>
<td>Specifies the number of posterior evaluations</td>
</tr>
<tr>
<td>OUT=</td>
<td>Creates an output data set containing the sample from the posterior density</td>
</tr>
<tr>
<td>OUTG=</td>
<td>Creates an output data set from the grid evaluations</td>
</tr>
<tr>
<td>OUTGT=</td>
<td>Creates an output data set from the transformed grid evaluations</td>
</tr>
<tr>
<td>PSEARCH=</td>
<td>Displays the search used to determine the parameters for the inverted gamma densities</td>
</tr>
<tr>
<td>PTRANS=</td>
<td>Displays the transformation of the variance components</td>
</tr>
<tr>
<td>SEED=</td>
<td>Specifies an integer used to start the pseudo-random number generator</td>
</tr>
<tr>
<td>SFACTOR=</td>
<td>Adjusts the search range of the transformed parameters</td>
</tr>
<tr>
<td>TDATA=</td>
<td>Inputs the transformation used by the sampling algorithm</td>
</tr>
<tr>
<td>TRANS=</td>
<td>Specifies the algorithm that determines the transformation of the covariance parameters</td>
</tr>
<tr>
<td>UPDATE=</td>
<td>An alias for the LOGNOTE= option</td>
</tr>
</tbody>
</table>

The distribution argument in the PRIOR statement determines the prior density for the variance component parameters of your mixed model. Valid values are as follows.

**DATA=**

 enables you to input the prior densities of the variance components used by the sampling algorithm. This data set must contain the **Type** and **Parm1–Parmn** variables, where \(n\) is the largest number of parameters among each of the base densities. The format of the DATA= data set matches that created by PROC MIXED in the “Base Densities” table, so you can output the densities from one run and use them as input for a subsequent run.
JEFFREYS
specifies a noninformative reference version of Jeffreys’ prior constructed by using the square root of
the determinant of the expected information matrix as in (1.3.92) of Box and Tiao (1973). This is the
default prior.

FLAT
specifies a prior density equal to 1 everywhere, making the likelihood function the posterior.

You can specify the following options in the PRIOR statement after a slash (/).

ALG=IC | INDCHAIN
ALG=IS | IMPSAMP
ALG=RS | REJSAMP
ALG=RWC | RWCHAIN

specifies the algorithm used for generating the posterior sample. The ALG=IC option requests an
independence chain algorithm, and it is the default. The option ALG=IS requests importance sampling,
ALG=RS requests rejection sampling, and ALG=RWC requests a random walk chain. For more
information about these techniques, see Ripley (1987); Smith and Gelfand (1992); Tierney (1994).

BDATA=
enables you to input the base densities used by the sampling algorithm. This data set must contain the
Type and Parm1–Parmn variables, where n is the largest number of parameters among each of the base
densities. The format of the BDATA= data set matches that created by PROC MIXED in the “Base
Densities” table, so you can output the densities from one run and use them as input for a subsequent
run.

GRID=(value-list)
specifies a grid of values over which to evaluate the posterior density. The value-list syntax is the same
as in the PARMS statement, and you must specify an output data set name with the OUTG= option.

GRIDT=(value-list)
specifies a transformed grid of values over which to evaluate the posterior density. The value-list
syntax is the same as in the PARMS statement, and you must specify an output data set name with the
OUTGT= option.

IFACTOR=number
is an alias for the SFACTOR= option.

LOGNOTE=number
instructs PROC MIXED to write a note to the SAS log after it generates the sample corresponding to
each multiple of number. This is useful for monitoring the progress of CPU-intensive runs.

LOGRBOUND=number
specifies the bounding constant for rejection sampling. The value of number equals the maximum of
\log\{f/g\} over the variance component parameter space, where f is the posterior density and g is the
product inverted gamma densities used to perform rejection sampling.

When performing the rejection sampling, you might encounter the following message:

WARNING: The log ratio bound of LL was violated at sample XX.
When this occurs, PROC MIXED reruns an optimization algorithm to determine a new log upper bound and then restarts the rejection sampling. The resulting OUT= data set contains all observations that have been generated; therefore, assuming that you have requested $N$ samples, you should retain only the final $N$ observations in this data set for analysis purposes.

**NSAMPLE=number**

specifies the number of posterior samples to generate. The default is 1000, but more accurate results are obtained with larger samples such as 10000.

**NSEARCH=number**

specifies the number of posterior evaluations PROC MIXED makes for each transformed parameter in determining the parameters for the inverted gamma densities. The default is 20.

**OUT=SAS-data-set**

creates an output data set containing the sample from the posterior density.

**OUTG=SAS-data-set**

creates an output data set from the grid evaluations specified in the GRID= option.

**OUTGT=SAS-data-set**

creates an output data set from the transformed grid evaluations specified in the GRIDT= option.

**PSEARCH**

displays the search used to determine the parameters for the inverted gamma densities. The ODS name of the table is Search.

**PTRANS**

displays the transformation of the variance components. The ODS name of the table is Trans.

**SEED=number**

specifies an integer used to start the pseudo-random number generator for the simulation. If you do not specify a seed, or if you specify a value less than or equal to zero, the seed is by default generated from reading the time of day from the computer clock. You should use a positive seed (less than $2^{31} - 1$) whenever you want to duplicate the sample in another run of PROC MIXED.

**SFACTOR=number**

enables you to adjust the range over which PROC MIXED searches the transformed parameters in order to determine the parameters for the inverted gamma densities. PROC MIXED determines the range by first transforming the estimates from the standard PROC MIXED analysis (REML, ML, or MIVQUE0, depending upon which estimation method you select). It then multiplies and divides the transformed estimates by $2 \times number$ to obtain upper and lower bounds, respectively. Transformed values that produce negative variance components in the original scale are not included in the search. The default value is 1; number must be greater than 0.5.

**TDATA=SAS-data-set**

enables you to input the transformation of the covariance parameters used by the sampling algorithm. This data set should contain the CovP1–CovPn variables. The format of the TDATA= data set matches that created by PROC MIXED in the Trans table, so you can output the transformation from one run and use it as input for a subsequent run.
TRANS=EXPECTED | MIVQUE0 | OBSERVED
specifies the particular algorithm used to determine the transformation of the covariance parameters. The default is MIVQUE0, indicating a transformation based on the MIVQUE(0) equations. The other two options indicate the type of Hessian matrix used in constructing the transformation via a Cholesky root.

UPDATE=number
is an alias for the LOGNOTE= option.

### RANDOM Statement

```
RANDOM random-effects / options;
```

The RANDOM statement defines the random effects constituting the \( \mathbf{y} \) vector in the mixed model. It can be used to specify traditional variance component models (as in the VARCOMP procedure) and to specify random coefficients. The random effects can be classification or continuous, and multiple RANDOM statements are possible.

Using notation from the section “Mixed Models Theory” on page 5313, the purpose of the RANDOM statement is to define the \( \mathbf{Z} \) matrix of the mixed model, the random effects in the \( \mathbf{y} \) vector, and the structure of \( \mathbf{G} \). The \( \mathbf{Z} \) matrix is constructed exactly as the \( \mathbf{X} \) matrix for the fixed effects, and the \( \mathbf{G} \) matrix is constructed to correspond with the effects constituting \( \mathbf{Z} \). The structure of \( \mathbf{G} \) is defined by using the TYPE= option.

You can specify INTERCEPT (or INT) as a random effect to indicate the intercept. PROC MIXED does not include the intercept in the RANDOM statement by default as it does in the MODEL statement.

Table 63.15 summarizes the options available in the RANDOM statement. All options are subsequently discussed in alphabetical order.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construction of Covariance Structure</strong></td>
<td></td>
</tr>
<tr>
<td>GDATA=</td>
<td>Requests that the ( \mathbf{G} ) matrix be read from a SAS data set</td>
</tr>
<tr>
<td>GROUP=</td>
<td>Varies covariance parameters by groups</td>
</tr>
<tr>
<td>LDATA=</td>
<td>Specifies data set with coefficient matrices for TYPE=LIN</td>
</tr>
<tr>
<td>NOFULLZ</td>
<td>Eliminates columns in ( \mathbf{Z} ) corresponding to missing values</td>
</tr>
<tr>
<td>RATIOS</td>
<td>Indicates that ratios are specified in the GDATA= data set</td>
</tr>
<tr>
<td>SUBJECT=</td>
<td>Identifies the subjects in the model</td>
</tr>
<tr>
<td>TYPE=</td>
<td>Specifies the covariance structure</td>
</tr>
<tr>
<td><strong>Statistical Output</strong></td>
<td></td>
</tr>
<tr>
<td>ALPHA=( \alpha )</td>
<td>Determines the confidence level ( (1 - \alpha) )</td>
</tr>
<tr>
<td>CL</td>
<td>Requests confidence limits for predictors of random effects</td>
</tr>
<tr>
<td>G</td>
<td>Displays the estimated ( \mathbf{G} ) matrix</td>
</tr>
<tr>
<td>GC</td>
<td>Displays the Cholesky root (lower) of estimated ( \mathbf{G} ) matrix</td>
</tr>
<tr>
<td>GCI</td>
<td>Displays the inverse Cholesky root (lower) of estimated ( \mathbf{G} ) matrix</td>
</tr>
<tr>
<td>GCORR</td>
<td>Displays the correlation matrix corresponding to estimated ( \mathbf{G} ) matrix</td>
</tr>
</tbody>
</table>
Table 63.15 continued

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GI</td>
<td>Displays the inverse of the estimated G matrix</td>
</tr>
<tr>
<td>SOLUTION</td>
<td>Displays solutions $\hat{\gamma}$ of the G-side random effects</td>
</tr>
<tr>
<td>V</td>
<td>Displays blocks of the estimated V matrix</td>
</tr>
<tr>
<td>VC</td>
<td>Displays the lower-triangular Cholesky root of blocks of the estimated V matrix</td>
</tr>
<tr>
<td>VCI</td>
<td>Displays the inverse Cholesky root of blocks of the estimated V matrix</td>
</tr>
<tr>
<td>VCORR</td>
<td>Displays the correlation matrix corresponding to blocks of the estimated V matrix</td>
</tr>
<tr>
<td>VI</td>
<td>Displays the inverse of the blocks of the estimated V matrix</td>
</tr>
</tbody>
</table>

You can specify the following *options* in the RANDOM statement after a slash (/).

**ALPHA=number**
requests that a $t$-type confidence interval be constructed for each of the random-effect estimates with confidence level $1 - \text{number}$. The value of number must be between 0 and 1; the default is 0.05.

**CL**
requests that $t$-type confidence limits be constructed for each of the random-effect estimates. The confidence level is 0.95 by default; this can be changed with the ALPHA= option.

**G**
requests that the estimated G matrix be displayed. PROC MIXED displays blanks for values that are 0. If you specify the SUBJECT= option, then the block of the G matrix corresponding to the first subject is displayed. The ODS name of the table is G.

**GC**
displays the lower-triangular Cholesky root of the estimated G matrix according to the rules listed under the G option. The ODS name of the table is CholG.

**GCI**
displays the inverse Cholesky root of the estimated G matrix according to the rules listed under the G option. The ODS name of the table is InvCholG.

**GCORR**
displays the correlation matrix corresponding to the estimated G matrix according to the rules listed under the G option. The ODS name of the table is GCorr.

**GDATA=SAS-data-set**
requests that the G matrix be read in from a SAS data set. This G matrix is assumed to be known; therefore, only R-side parameters from effects in the REPEATED statement are included in the Newton-Raphson iterations. If no REPEATED statement is specified, then only a residual variance is estimated.

The information in the GDATA= data set can appear in one of two ways. The first is a sparse representation for which you include Row, Col, and Value variables to indicate the row, column, and value of G, respectively. All unspecified locations are assumed to be 0. The second representation
is for dense matrices. In it you include Row and Col1–Coln variables to indicate, respectively, the row and columns of G, which is a symmetric matrix of order n. For both representations, you must specify effects in the RANDOM statement that generate a Z matrix that contains n columns. (See Example 63.4.)

If you have more than one RANDOM statement, only one GDATA= option is required in any one of them, and the data set you specify must contain the entire G matrix defined by all of the RANDOM statements.

If the GDATA= data set contains variance ratios instead of the variances themselves, then use the RATIOS option.

Known parameters of G can also be input by using the PARMS statement with the HOLD= option.

GI displays the inverse of the estimated G matrix according to the rules listed under the G option. The ODS name of the table is InvG.

GROUP=effect
GRP=effect defines an effect specifying heterogeneity in the covariance structure of G. All observations having the same level of the group effect have the same covariance parameters. Each new level of the group effect produces a new set of covariance parameters with the same structure as the original group. You should exercise caution in defining the group effect, because strange covariance patterns can result from its misuse. Also, the group effect can greatly increase the number of estimated covariance parameters, which can adversely affect the optimization process.

Continuous variables are permitted as arguments to the GROUP= option. PROC MIXED does not sort by the values of the continuous variable; rather, it considers the data to be from a new subject or group whenever the value of the continuous variable changes from the previous observation. Using a continuous variable decreases execution time for models with a large number of subjects or groups and also prevents the production of a large “Class Level Information” table.

LDATA=SAS-data-set
reads the coefficient matrices associated with the TYPE=LIN(number) option. The data set must contain the variables Parm, Row, Col1–Coln or Parm, Row, Col, Value. The Parm variable denotes which of the number coefficient matrices is currently being constructed, and the Row, Col1–Coln, or Row, Col, Value variables specify the matrix values, as they do with the GDATA= option. Unspecified values of these matrices are set equal to 0.

NOFULLZ eliminates the columns in Z corresponding to missing levels of random effects involving CLASS variables. By default, these columns are included in Z.

RATIOS indicates that ratios with the residual variance are specified in the GDATA= data set instead of the covariance parameters themselves. The default GDATA= data set contains the individual covariance parameters.
SOLUTION

requests that the solution for the random-effects parameters be produced. Using notation from the section “Mixed Models Theory” on page 5313, these estimates are the empirical best linear unbiased predictors (EBLUPs) $\hat{y} = GZV^{-1}(y - X\hat{\beta})$. They can be useful for comparing the random effects from different experimental units and can also be treated as residuals in performing diagnostics for your mixed model.

The numbers displayed in the SE Pred column of the “Solution for Random Effects” table are not the standard errors of the $\hat{y}$ displayed in the Estimate column; rather, they are the standard errors of predictions $\hat{y}_i - y_i$, where $\hat{y}_i$ is the $i$th EBLUP and $y_i$ is the $i$th random-effect parameter.

SUBJECT=effect

identifies the subjects in your mixed model. Complete independence is assumed across subjects; thus, for the RANDOM statement, the SUBJECT= option produces a block-diagonal structure in $G$ with identical blocks. The $Z$ matrix is modified to accommodate this block diagonality. In fact, specifying a subject effect is equivalent to nesting all other effects in the RANDOM statement within the subject effect.

Continuous variables are permitted as arguments to the SUBJECT= option. PROC MIXED does not sort by the values of the continuous variable; rather, it considers the data to be from a new subject or group whenever the value of the continuous variable changes from the previous observation. Using a continuous variable decreases execution time for models with a large number of subjects or groups and also prevents the production of a large “Class Level Information” table.

When you specify the SUBJECT= option and a classification random effect, computations are usually much quicker if the levels of the random effect are duplicated within each level of the SUBJECT= effect.

TYPE= covariance-structure

specifies the covariance structure of $G$. Valid values for covariance-structure and their descriptions are listed in Table 63.17 and Table 63.18. Although a variety of structures are available, most applications call for either TYPE=VC or TYPE=UN. The TYPE=VC (variance components) option is the default structure, and it models a different variance component for each random effect.

The TYPE=UN (unstructured) option is useful for correlated random coefficient models. For example, the following statement specifies a random intercept-slope model that has different variances for the intercept and slope and a covariance between them:

```
random intercept age / type=un subject=person;
```

You can also use TYPE=FA0(2) here to request a $G$ estimate that is constrained to be nonnegative definite.

If you are constructing your own columns of $Z$ with continuous variables, you can use the TYPE=TOEP(1) structure to group them together to have a common variance component. If you want to have different covariance structures in different parts of $G$, you must use multiple RANDOM statements with different TYPE= options.
V<\texttt{value-list}>
requests that blocks of the estimated V matrix be displayed. The first block determined by the
\texttt{SUBJECT=} effect is the default displayed block. PROC MIXED displays entries that are 0 as blanks
in the table.

You can optionally use the \texttt{value-list} specification, which indicates the subjects for which blocks of V
are to be displayed. For example, the following statement displays block matrices for the first, third,
and seventh persons:

\begin{verbatim}
random int time / type=un subject=person v=1,3,7;
\end{verbatim}

The ODS table name is V.

VC<\texttt{value-list}>
displays the Cholesky root of the blocks of the estimated V matrix. The \texttt{value-list} specification is the
same as in the V option. The ODS table name is CholV.

VCI<\texttt{value-list}>
displays the inverse of the Cholesky root of the blocks of the estimated V matrix. The \texttt{value-list}
specification is the same as in the V option. The ODS table name is InvCholV.

VCORR<\texttt{value-list}>
displays the correlation matrix corresponding to the blocks of the estimated V matrix. The \texttt{value-list}
specification is the same as in the V option. The ODS table name is VCorr.

VI<\texttt{value-list}>
displays the inverse of the blocks of the estimated V matrix. The \texttt{value-list} specification is the same as
in the V option. The ODS table name is InvV.

---

**REPEATED Statement**

\begin{verbatim}
REPEATED <repeated-effect> / options;
\end{verbatim}

The REPEATED statement is used to specify the R matrix in the mixed model. Its syntax is different from
that of the REPEATED statement in PROC GLM. If no REPEATED statement is specified, R is assumed to
be equal to $\sigma^2 I$.

For many repeated measures models, no repeated effect is required in the REPEATED statement. Simply use
the \texttt{SUBJECT=} option to define the blocks of R and the \texttt{TYPE=} option to define their covariance structure.
In this case, the repeated measures data must be similarly ordered for each subject, and you must indicate
all missing response variables with periods in the input data set unless they all fall at the end of a subject’s
repeated response profile. These requirements are necessary in order to inform PROC MIXED of the proper
location of the observed repeated responses.

Specifying a repeated effect is useful when you do not want to indicate missing values with periods in the
input data set. The repeated effect must contain only classification variables. Make sure that the levels of
the repeated effect are different for each observation within a subject; otherwise, PROC MIXED constructs
identical rows in R corresponding to the observations with the same level. This results in a singular R and an
infinite likelihood.
Whether you specify a REPEATED effect or not, the rows of $R$ for each subject are constructed in the order in which they appear in the input data set.

Table 63.16 summarizes the options available in the REPEATED statement. All options are subsequently discussed in alphabetical order.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construction of Covariance Structure</strong></td>
<td></td>
</tr>
<tr>
<td>GROUP=</td>
<td>Defines an effect specifying heterogeneity in the R-side covariance structure</td>
</tr>
<tr>
<td>LDATA=</td>
<td>Specifies data set with coefficient matrices for TYPE=LIN</td>
</tr>
<tr>
<td>LOCAL</td>
<td>Requests that a diagonal matrix be added to $R$</td>
</tr>
<tr>
<td>LOCALW</td>
<td>Specifies that only the local effects are weighted</td>
</tr>
<tr>
<td>NONLOCALW</td>
<td>Specifies that only the nonlocal effects are weighted</td>
</tr>
<tr>
<td>SUBJECT=</td>
<td>Identifies the subjects in the R-side model</td>
</tr>
<tr>
<td>TYPE=</td>
<td>Specifies the R-side covariance structure</td>
</tr>
<tr>
<td><strong>Statistical Output</strong></td>
<td></td>
</tr>
<tr>
<td>HLM</td>
<td>Produces a table of Hotelling-Lawley-McKeon statistics (McKeon 1974)</td>
</tr>
<tr>
<td>HLPS</td>
<td>Produces a table of Hotelling-Lawley-Pillai-Samson statistics (Pillai and Samson 1959)</td>
</tr>
<tr>
<td>R</td>
<td>Displays blocks of the estimated $R$ matrix</td>
</tr>
<tr>
<td>RC</td>
<td>Display the Cholesky root (lower) of blocks of the estimated $R$ matrix</td>
</tr>
<tr>
<td>RCI</td>
<td>Displays the inverse Cholesky root (lower) of blocks of the estimated $R$ matrix</td>
</tr>
<tr>
<td>RCORR</td>
<td>Displays the correlation matrix corresponding to blocks of the estimated $R$ matrix</td>
</tr>
<tr>
<td>RI</td>
<td>Displays the inverse of blocks of the estimated $R$ matrix</td>
</tr>
</tbody>
</table>

You can specify the following options in the REPEATED statement after a slash (/).

**GROUP=**<em>effect</em>

**GRP=**<em>effect</em>

defines an effect that specifies heterogeneity in the covariance structure of $R$. All observations that have the same level of the GROUP effect have the same covariance parameters. Each new level of the GROUP effect produces a new set of covariance parameters with the same structure as the original group. You should exercise caution in properly defining the GROUP effect, because strange covariance patterns can result with its misuse. Also, the GROUP effect can greatly increase the number of estimated covariance parameters, which can adversely affect the optimization process.

Continuous variables are permitted as arguments to the GROUP= option. PROC MIXED does not sort by the values of the continuous variable; rather, it considers the data to be from a new subject or group whenever the value of the continuous variable changes from the previous observation. Using a continuous variable decreases execution time for models with a large number of subjects or groups and also prevents the production of a large “Class Level Information” table.
produces a table of Hotelling-Lawley-McKeon statistics (McKeon 1974) for all fixed effects whose levels change across data having the same level of the SUBJECT= effect (the within-subject fixed effects). This option applies only when you specify a REPEATED statement with the TYPE=UN option and no RANDOM statements. For balanced data, this model is equivalent to the multivariate model for repeated measures in PROC GLM.

The Hotelling-Lawley-McKeon statistic has a slightly better \( F \) approximation than the Hotelling-Lawley-Pillai-Samson statistic (see the description of the HLPS option, which follows). Both of the Hotelling-Lawley statistics can perform much better in small samples than the default \( F \) statistic (Wright 1994).

Separate tables are produced for Type 1, 2, and 3 tests, according to the ones you select. The ODS table names are HLM1, HLM2, and HLM3, respectively.

**HLPS**
produces a table of Hotelling-Lawley-Pillai-Samson statistics (Pillai and Samson 1959) for all fixed effects whose levels change across data having the same level of the SUBJECT= effect (the within-subject fixed effects). This option applies only when you specify a REPEATED statement with the TYPE=UN option and no RANDOM statements. For balanced data, this model is equivalent to the multivariate model for repeated measures in PROC GLM, and this statistic is the same as the Hotelling-Lawley Trace statistic produced by PROC GLM.

Separate tables are produced for Type 1, 2, and 3 tests, according to the ones you select. The ODS table names are HLPS1, HLPS2, and HLPS3, respectively.

**LDATA=SAS-data-set**
reads the coefficient matrices associated with the TYPE=LIN(number) option. The data set must contain the variables Parm, Row, Col1-Coln or Parm, Row, Col, Value. The Parm variable denotes which of the number coefficient matrices is currently being constructed, and the Row, Col1-Coln, or Row, Col, Value variables specify the matrix values, as they do with the RANDOM statement option GDATA=. Unspecified values of these matrices are set equal to 0.

**LOCAL**

**LOCAL=EXP(<effects>)**

**LOCAL=POM(POM-data-set)**
requests that a diagonal matrix be added to \( R \). With just the LOCAL option, this diagonal matrix equals \( \sigma^2 I \), and \( \sigma^2 \) becomes an additional variance parameter that PROC MIXED profiles out of the likelihood provided that you do not specify the NOPROFILE option in the PROC MIXED statement. The LOCAL option is useful if you want to add an observational error to a time series structure (Jones and Boadi-Boateng 1991) or a nugget effect to a spatial structure Cressie (1993).

The LOCAL=EXP(<effects>) option produces exponential local effects, also known as dispersion effects, in a log-linear variance model. These local effects have the form

\[
\sigma^2 \text{diag}[\exp(U\delta)]
\]

where \( U \) is the full-rank design matrix corresponding to the effects that you specify and \( \delta \) are the parameters that PROC MIXED estimates. An intercept is not included in \( U \) because it is accounted for by \( \sigma^2 \). PROC MIXED constructs the full-rank \( U \) in terms of 1s and –1s for classification effects. Be sure to scale continuous effects in \( U \) sensibly.
The LOCAL=POM(*POM-data-set*) option specifies the power-of-the-mean structure. This structure possesses a variance of the form $\sigma^2 |x_i' \hat{\beta}^*|^\theta$ for the $i$th observation, where $x_i$ is the $i$th row of $X$ (the design matrix of the fixed effects) and $\hat{\beta}^*$ is an estimate of the fixed-effects parameters that you specify in *POM-data-set*.

The SAS data set specified by *POM-data-set* contains the numeric variable *Estimate* (in previous releases, the variable name was required to be EST), and it has at least as many observations as there are fixed-effects parameters. The first $p$ observations of the *Estimate* variable in *POM-data-set* are taken to be the elements of $\hat{\beta}^*$, where $p$ is the number of columns of $X$. You must order these observations according to the non-full-rank parameterization of the MIXED procedure. One easy way to set up *POM-data-set* for a $\hat{\beta}^*$ corresponding to ordinary least squares is illustrated by the following statements:

```plaintext
ods output SolutionF=sf;
proc mixed;
   class a;
   model y = a x / s;
run;

proc mixed;
   class a;
   model y = a x;
   repeated / local=pom(sf);
run;
```

Note that the generalized least squares estimate of the fixed-effects parameters from the second PROC MIXED step usually is not the same as your specified $\hat{\beta}^*$. However, you can iterate the POM fitting until the two estimates agree. Continuing from the previous example, the statements for performing one step of this iteration are as follows:

```plaintext
ods output SolutionF=sf1;
proc mixed;
   class a;
   model y = a x / s;
   repeated / local=pom(sf);
run;

proc compare brief data=sf compare=sf1;
   var estimate;
run;

data sf;
   set sf1;
run;
```

Unfortunately, this iterative process does not always converge. For further details, see the description of pseudo-likelihood in Chapter 3 of Carroll and Ruppert (1988).
LOCALW specifies that only the local effects and no others be weighted. By default, all effects are weighted. The LOCALW option is used in connection with the WEIGHT statement and the LOCAL option in the REPEATED statement.

NONLOCALW specifies that only the nonlocal effects and no others be weighted. By default, all effects are weighted. The NONLOCALW option is used in connection with the WEIGHT statement and the LOCAL option in the REPEATED statement.

\[ R=value-list \]
requests that blocks of the estimated \( R \) matrix be displayed. The first block determined by the SUBJECT= effect is the default displayed block. PROC MIXED displays blanks for value-lists that are 0.

The value-list indicates the subjects for which blocks of \( R \) are to be displayed. For example, the following statement displays block matrices for the first, third, and fifth persons:

\[
\text{repeated} / \text{type}=\text{cs} \ \text{subject}=\text{person} \ r=1,3,5;
\]

See the PARMS statement for the possible forms of value-list. The ODS table name is R.

\[ RC=value-list \]
produces the Cholesky root of blocks of the estimated \( R \) matrix. The value-list specification is the same as with the \( R \) option. The ODS table name is CholR.

\[ RCI=value-list \]
produces the inverse Cholesky root of blocks of the estimated \( R \) matrix. The value-list specification is the same as with the \( R \) option. The ODS table name is InvCholR.

\[ RCORR=value-list \]
produces the correlation matrix corresponding to blocks of the estimated \( R \) matrix. The value-list specification is the same as with the \( R \) option. The ODS table name is RCorr.

\[ RI=value-list \]
produces the inverse of blocks of the estimated \( R \) matrix. The value-list specification is the same as with the \( R \) option. The ODS table name is InvR.

SSCP requests that an unstructured \( R \) matrix be estimated from the sum-of-squares-and-crossproducts matrix of the residuals. It applies only when you specify TYPE=UN and have no RANDOM statements. Also, you must have a sufficient number of subjects for the estimate to be positive definite.

This option is useful when the size of the blocks of \( R \) is large (for example, greater than 10) and you want to use or inspect an unstructured estimate that is much quicker to compute than the default REML estimate. The two estimates will agree for certain balanced data sets when you have a classification fixed effect defined across all time points within a subject.
SUBJECT=effect

SUB=effect identifies the subjects in your mixed model. Complete independence is assumed across subjects; therefore, the SUBJECT= option produces a block-diagonal structure in R with identical blocks. When the SUBJECT= effect consists entirely of classification variables, the blocks of R correspond to observations sharing the same level of that effect. These blocks are sorted according to this effect as well.

Continuous variables are permitted as arguments to the SUBJECT= option. PROC MIXED does not sort by the values of the continuous variable; rather, it considers the data to be from a new subject or group whenever the value of the continuous variable changes from the previous observation. Using a continuous variable decreases execution time for models with a large number of subjects or groups and also prevents the production of a large “Class Level Information” table.

If you want to model nonzero covariance among all of the observations in your SAS data set, specify SUBJECT=INTERCEPT to treat the data as if they are all from one subject. However, be aware that in this case PROC MIXED manipulates an R matrix with dimensions equal to the number of observations.

If no SUBJECT= effect is specified, then every observation is assumed to be from a different subject and R is assumed to be diagonal. For this reason, you usually want to use the SUBJECT= option in the REPEATED statement.

TYPE=covariance-structure specifies the covariance structure of the R matrix. The SUBJECT= option defines the blocks of R, and the TYPE= option specifies the structure of these blocks. Valid values for covariance-structure and their descriptions are provided in Table 63.17 and Table 63.18. The default structure is VC.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Description</th>
<th>Parms</th>
<th>(i, j) element</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTE(1)</td>
<td>Antedependence</td>
<td>2t - 1</td>
<td>$\sigma_i \sigma_j \prod_{k=1}^{j} \rho_k$</td>
</tr>
<tr>
<td>AR(1)</td>
<td>Autoregressive(1)</td>
<td>2</td>
<td>$\sigma^2 \rho^{</td>
</tr>
<tr>
<td>ARH(1)</td>
<td>Heterogeneous AR(1)</td>
<td>t + 1</td>
<td>$\sigma_i \sigma_j \rho^{</td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>ARMA(1,1)</td>
<td>3</td>
<td>$\sigma^2 [\rho^{</td>
</tr>
<tr>
<td>CS</td>
<td>Compound symmetry</td>
<td>2</td>
<td>$\sigma_1 + \sigma^2 1(i = j)$</td>
</tr>
<tr>
<td>CSH</td>
<td>Heterogeneous CS</td>
<td>t + 1</td>
<td>$\sigma_i \sigma_j [\rho 1(i \neq j) + 1(i = j)]$</td>
</tr>
<tr>
<td>FA(q)</td>
<td>Factor analytic</td>
<td>$q \frac{2}{2}(2t - q + 1) + t$</td>
<td>$\sum_{k=1}^{t} \lambda_{i,j} A_{ij} + \sigma^2 1(i = j)$</td>
</tr>
<tr>
<td>FA0(q)</td>
<td>No diagonal FA</td>
<td>$q \frac{2}{2}(2t - q + 1)$</td>
<td>$\sum_{k=1}^{t} \lambda_{i,j} A_{ij}$</td>
</tr>
<tr>
<td>FA1(q)</td>
<td>Equal diagonal FA</td>
<td>$q \frac{2}{2}(2t - q + 1) + 1$</td>
<td>$\sum_{k=1}^{t} \lambda_{i,j} A_{ij}$</td>
</tr>
<tr>
<td>HF</td>
<td>Huynh-Feldt</td>
<td>t + 1</td>
<td>$(\sigma_i^2 + \sigma_j^2)/2 + \lambda 1(i \neq j)$</td>
</tr>
<tr>
<td>LIN(q)</td>
<td>General linear</td>
<td>q</td>
<td>$\sum_{k=1}^{t} \theta_k A_{ij}$</td>
</tr>
<tr>
<td>TOEP</td>
<td>Toeplitz</td>
<td>t</td>
<td>$\sigma_{</td>
</tr>
<tr>
<td>TOEP(q)</td>
<td>Banded Toeplitz</td>
<td>q</td>
<td>$\sigma_{</td>
</tr>
<tr>
<td>TOEPH</td>
<td>Heterogeneous TOEP</td>
<td>2t - 1</td>
<td>$\sigma_i \sigma_j \rho_{</td>
</tr>
<tr>
<td>TOEPH(q)</td>
<td>Banded hetero TOEP</td>
<td>t + q - 1</td>
<td>$\sigma_i \sigma_j \rho_{</td>
</tr>
</tbody>
</table>
In Table 63.17, “Parms” is the number of covariance parameters in the structure, \( t \) is the overall dimension of the covariance matrix, and \( 1(A) = 1 \) when \( A \) is true and 0 otherwise. For example, \( 1(i = j) \) equals 1 when \( i = j \) and 0 otherwise, and \( 1(|i - j| < q) \) equals 1 when \(|i - j| < q\) and 0 otherwise. For the TYPE=TOEPH structures, \( \rho_0 = 1 \), and for the TYPE=UNR structures, \( \rho_i = 1 \) for all \( i \). For the direct product structures, the subscripts “1” and “2” see the first and second structure in the direct product, respectively. And \( i_1 = \text{int}((i + t_2 - 1)/t_2), \) \( j_1 = \text{int}((j + t_2 - 1)/t_2), \) \( i_2 = \text{mod}(i - 1, t_2) + 1, \) and \( j_2 = \text{mod}(j - 1, t_2) + 1. \)

### Table 63.17 continued

<table>
<thead>
<tr>
<th>Structure</th>
<th>Description</th>
<th>Parms</th>
<th>((i, j)) element</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UN</strong></td>
<td>Unstructured</td>
<td>( t(t + 1)/2 )</td>
<td>( \sigma_{ij} )</td>
</tr>
<tr>
<td><strong>UN(q)</strong></td>
<td>Banded</td>
<td>( \frac{q}{2}(2t - q + 1) )</td>
<td>( \sigma_{ij} 1(</td>
</tr>
<tr>
<td><strong>UNR</strong></td>
<td>Unstructured corrs</td>
<td>( t(t + 1)/2 )</td>
<td>( \sigma_i \sigma_j \rho_{\text{max}(i,j)} \min(i,j) )</td>
</tr>
<tr>
<td><strong>UNR(q)</strong></td>
<td>Banded correlations</td>
<td>( \frac{q}{2}(2t - q + 1) )</td>
<td>( \sigma_i \sigma_j \rho_{\text{max}(i,j)} \min(i,j) )</td>
</tr>
<tr>
<td><strong>UN@AR(1)</strong></td>
<td>Direct product AR(1)</td>
<td>( t_1(t_1 + 1)/2 + 1 )</td>
<td>( \sigma_{i_1,j_1} \rho_{</td>
</tr>
<tr>
<td><strong>UN@CS</strong></td>
<td>Direct product CS</td>
<td>( t_1(t_1 + 1)/2 + 1 )</td>
<td>( \sigma_{i_1,j_1} \rho_{i_2 \neq j_2}^{1} )</td>
</tr>
<tr>
<td><strong>UN@UN</strong></td>
<td>Direct product UN</td>
<td>( t_1(t_1 + 1)/2 + \frac{t_2(t_2 + 1)}{2} - 1 )</td>
<td>( \sigma_{i_1,j_1} \sigma_{i_2,j_2} )</td>
</tr>
<tr>
<td><strong>VC</strong></td>
<td>Variance components</td>
<td>( q )</td>
<td>( \sigma_k^2 ) ( i = j )</td>
</tr>
</tbody>
</table>

And \( i \) corresponds to \( k \)th effect.

### Table 63.18 Spatial Covariance Structures

<table>
<thead>
<tr>
<th>Structure</th>
<th>Description</th>
<th>Parms</th>
<th>((i, j)) element</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SP(EXP)(c-list)</strong></td>
<td>Exponential</td>
<td>2</td>
<td>( \sigma^2 \exp{-d_{ij}/\theta} )</td>
</tr>
<tr>
<td><strong>SP(EXPA)(c-list)</strong></td>
<td>Anisotropic exponential</td>
<td>( 2c + 1 )</td>
<td>( \sigma^2 \prod_{k=1}^c \exp{-\theta_k d(i, j, k)^p_k} )</td>
</tr>
<tr>
<td><strong>SP(EXPGA)(c_1 c_2)</strong></td>
<td>2D exponential,</td>
<td>4</td>
<td>( \sigma^2 \exp{-d_{ij}(\theta, \lambda)/\rho} )</td>
</tr>
<tr>
<td></td>
<td>geometrically anisotropic</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SP(GAU)(c-list)</strong></td>
<td>Gaussian</td>
<td>2</td>
<td>( \sigma^2 \exp{-d_{ij}^2/\rho^2} )</td>
</tr>
<tr>
<td><strong>SP(GAUGA)(c_1 c_2)</strong></td>
<td>2D Gaussian,</td>
<td>4</td>
<td>( \sigma^2 \exp{-d_{ij}(\theta, \lambda)^2/\rho^2} )</td>
</tr>
<tr>
<td></td>
<td>geometrically anisotropic</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SP(LIN)(c-list)</strong></td>
<td>Linear</td>
<td>2</td>
<td>( \sigma^2 (1 - \rho d_{ij}) ) ( 1(\rho d_{ij} \leq 1) )</td>
</tr>
<tr>
<td><strong>SP(LINL)(c-list)</strong></td>
<td>Linear log</td>
<td>2</td>
<td>( \sigma^2 (1 - \rho \log(d_{ij})) ) ( \times1(\rho \log(d_{ij}) \leq 1, d_{ij} &gt; 0) )</td>
</tr>
<tr>
<td><strong>SP(MATERN)(c-list)</strong></td>
<td>Matérn</td>
<td>3</td>
<td>( \sigma^2 \frac{1}{\Gamma(v)} \left( \frac{d_{ij}}{2\rho} \right)^v 2K_v(d_{ij}/\rho) )</td>
</tr>
<tr>
<td><strong>SP(MATHSW)(c-list)</strong></td>
<td>Matérn</td>
<td>3</td>
<td>( \sigma^2 \frac{1}{\Gamma(v)} \left( \frac{d_{ij} \sqrt{v}}{\rho} \right)^v 2K_v \left( \frac{2d_{ij} \sqrt{v}}{\rho} \right) )</td>
</tr>
<tr>
<td></td>
<td>(Handcock-Stein-Wallis)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SP(POW)(c-list)</strong></td>
<td>Power</td>
<td>2</td>
<td>( \sigma^2 \rho_{d_{ij}} )</td>
</tr>
<tr>
<td><strong>SP(POWA)(c-list)</strong></td>
<td>Anisotropic power</td>
<td>( c + 1 )</td>
<td>( \sigma^2 \rho_{1} d_{i,j,1} \rho_{2} d_{i,j,2} \ldots \rho_{c} d_{i,j,c} )</td>
</tr>
</tbody>
</table>
In Table 63.18, \(c\)-list contains the names of the numeric variables used as coordinates of the location of the observation in space, and \(d_{ij}\) is the Euclidean distance between the \(i\)th and \(j\)th vectors of these coordinates, which correspond to the \(i\)th and \(j\)th observations in the input data set. For SP(POWA) and SP(EXPA), \(c\) is the number of coordinates, and \(d_{i;j;k}\) is the absolute distance between the \(k\)th coordinate, \(k = 1, \ldots, c\), of the \(i\)th and \(j\)th observations in the input data set. For the geometrically anisotropic structures SP(EXPGA), SP(GAUGA), and SP(SPHGA), exactly two spatial coordinate variables must be specified as \(c_1\) and \(c_2\). Geometric anisotropy is corrected by applying a rotation \(\theta\) and scaling \(\lambda\) to the coordinate system, and \(d_{ij}(\theta, \lambda)\) represents the Euclidean distance between two points in the transformed space. SP(MATERN) and SP(MATHSW) represent covariance structures in a class defined by Matérn (see Matérn 1986; Handcock and Stein 1993; Handcock and Wallis 1994). The function \(K_v\) is the modified Bessel function of the second kind of (real) order \(v > 0\); the parameter \(v\) governs the smoothness of the process (see below for more details).

Table 63.19 lists some examples of the structures in Table 63.17 and Table 63.18.

**Table 63.19** Covariance Structure Examples

<table>
<thead>
<tr>
<th>Description</th>
<th>Structure</th>
<th>Example</th>
</tr>
</thead>
</table>
| Variance components        | VC (default) | \[
|                            |           | \[
|                            |           | \[
| Compound symmetry          | CS        | \[
| Unstructured               | UN        | \[
| Banded main diagonal       | UN(1)     | \[
| First-order autoregressive | AR(1)     | \[

In Table 63.18, \(i, j\) element
<table>
<thead>
<tr>
<th>Description</th>
<th>Structure</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toeplitz</td>
<td>TOEP</td>
<td>$\begin{bmatrix} \sigma^2 &amp; \sigma_1 &amp; \sigma_2 &amp; \sigma_3 \ \sigma_1 &amp; \sigma^2 &amp; \sigma_1 &amp; \sigma_2 \ \sigma_2 &amp; \sigma_1 &amp; \sigma^2 &amp; \sigma_1 \ \sigma_3 &amp; \sigma_2 &amp; \sigma_1 &amp; \sigma^2 \end{bmatrix}$</td>
</tr>
<tr>
<td>Toeplitz with two bands</td>
<td>TOEP(2)</td>
<td>$\begin{bmatrix} \sigma^2 &amp; \sigma_1 &amp; 0 &amp; 0 \ \sigma_1 &amp; \sigma^2 &amp; \sigma_1 &amp; 0 \ 0 &amp; \sigma_1 &amp; \sigma^2 &amp; \sigma_1 \ 0 &amp; 0 &amp; \sigma_1 &amp; \sigma^2 \end{bmatrix}$</td>
</tr>
<tr>
<td>Spatial power</td>
<td>SP(POW)(c)</td>
<td>$\sigma^2 \begin{bmatrix} \rho_{d_1} &amp; 1 &amp; \rho_{d_2} &amp; \rho_{d_3} &amp; \rho_{d_4} \ 1 &amp; \rho_{d_2} &amp; 1 &amp; \rho_{d_3} &amp; \rho_{d_4} \ \rho_{d_2} &amp; 1 &amp; \rho_{d_3} &amp; 1 &amp; \rho_{d_4} \ \rho_{d_3} &amp; 1 &amp; \rho_{d_4} &amp; 1 &amp; \rho_{d_4} \end{bmatrix}$</td>
</tr>
<tr>
<td>Heterogeneous AR(1)</td>
<td>ARH(1)</td>
<td>$\begin{bmatrix} \sigma_1^2 &amp; \sigma_1 \sigma_2 \rho &amp; \sigma_1 \sigma_3 \rho^2 &amp; \sigma_4 \sigma_2 \rho &amp; \sigma_4 \sigma_3 \rho^2 \ \sigma_1 \sigma_2 \rho &amp; \sigma_2^2 &amp; \sigma_2 \sigma_3 \rho &amp; \sigma_2 \sigma_4 \rho^2 &amp; \sigma_2 \sigma_4 \rho^2 \ \sigma_1 \sigma_3 \rho^2 &amp; \sigma_2 \sigma_3 \rho &amp; \sigma_3^2 &amp; \sigma_3 \sigma_4 \rho &amp; \sigma_3 \sigma_4 \rho \ \sigma_4 \sigma_2 \rho &amp; \sigma_4 \sigma_3 \rho &amp; \sigma_4 \sigma_3 \rho &amp; \sigma_4^2 \end{bmatrix}$</td>
</tr>
<tr>
<td>First-order autoregressive moving average</td>
<td>ARMA(1,1)</td>
<td>$\sigma^2 \begin{bmatrix} 1 &amp; \gamma &amp; \gamma \rho &amp; \gamma \rho^2 \ \gamma &amp; 1 &amp; \gamma &amp; \gamma \rho \ \gamma \rho &amp; \gamma &amp; 1 &amp; \gamma \ \gamma \rho^2 &amp; \gamma \rho &amp; \gamma &amp; 1 \end{bmatrix}$</td>
</tr>
<tr>
<td>Heterogeneous CS</td>
<td>CSH</td>
<td>$\begin{bmatrix} \lambda_1^2 + d_1 &amp; \lambda_1 \lambda_2 &amp; \lambda_1 \lambda_3 &amp; \lambda_1 \lambda_4 \ \lambda_2 \lambda_1 &amp; \lambda_2^2 + d_2 &amp; \lambda_2 \lambda_3 &amp; \lambda_2 \lambda_4 \ \lambda_3 \lambda_1 &amp; \lambda_3 \lambda_2 &amp; \lambda_3^2 + d_3 &amp; \lambda_3 \lambda_4 \ \lambda_4 \lambda_1 &amp; \lambda_4 \lambda_2 &amp; \lambda_4 \lambda_3 &amp; \lambda_4^2 + d_4 \end{bmatrix}$</td>
</tr>
<tr>
<td>First-order factor analytic</td>
<td>FA(1)</td>
<td>$\begin{bmatrix} \sigma_1^2 &amp; \frac{\sigma_1^2 + \sigma_2^2}{2} - \lambda &amp; \frac{\sigma_1^2 + \sigma_2^2}{2} - \lambda \ \frac{\sigma_2^2}{2} - \lambda &amp; \frac{\sigma_2^2 + \sigma_3^2}{2} - \lambda &amp; \frac{\sigma_2^2 + \sigma_3^2}{2} - \lambda \ \frac{\sigma_3^2}{2} - \lambda &amp; \frac{\sigma_3^2 + \sigma_4^2}{2} - \lambda &amp; \frac{\sigma_3^2 + \sigma_4^2}{2} - \lambda \ \frac{\sigma_4^2}{2} &amp; \frac{\sigma_4^2 + \sigma_5^2}{2} - \lambda &amp; \frac{\sigma_4^2 + \sigma_5^2}{2} - \lambda \end{bmatrix}$</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>HF</td>
<td>$\begin{bmatrix} \sigma_1^2 &amp; \sigma_2 \sigma_1 \rho_1 &amp; \sigma_2 \sigma_2 \rho_1 &amp; \sigma_2 \sigma_2 \rho_1 \rho_2 \ \sigma_2 \sigma_1 \rho_1 &amp; \sigma_2^2 &amp; \sigma_2 \sigma_3 \rho_2 \ \sigma_2 \sigma_2 \rho_1 &amp; \sigma_2 \sigma_3 \rho_2 &amp; \sigma_2 \sigma_2 \rho_2 \rho_3 \ \sigma_3 \sigma_2 \rho_1 \rho_2 &amp; \sigma_3 \sigma_2 \rho_2 &amp; \sigma_3 \sigma_2 \rho_2 \rho_3 &amp; \sigma_3^2 \end{bmatrix}$</td>
</tr>
<tr>
<td>First-order antedependence</td>
<td>ANTE(1)</td>
<td>$\begin{bmatrix} \sigma_1^2 &amp; \sigma_1 \sigma_2 \rho_1 &amp; \sigma_1 \sigma_3 \rho_2 &amp; \sigma_1 \sigma_4 \rho_3 \ \sigma_2 \sigma_1 \rho_1 &amp; \sigma_2^2 &amp; \sigma_2 \sigma_3 \rho_2 &amp; \sigma_2 \sigma_4 \rho_3 \ \sigma_3 \sigma_2 \rho_1 \rho_2 &amp; \sigma_3 \sigma_3 \rho_1 &amp; \sigma_3^2 &amp; \sigma_3 \sigma_4 \rho_1 \ \sigma_4 \sigma_3 \rho_1 \rho_2 &amp; \sigma_4 \sigma_3 \rho_2 &amp; \sigma_4 \sigma_3 \rho_2 \rho_3 &amp; \sigma_4^2 \end{bmatrix}$</td>
</tr>
<tr>
<td>Heterogeneous Toeplitz</td>
<td>TOEPH</td>
<td>$\begin{bmatrix} \sigma_1^2 &amp; \sigma_2 \sigma_1 \rho_1 &amp; \sigma_2 \sigma_3 \rho_2 &amp; \sigma_2 \sigma_4 \rho_3 \ \sigma_2 \sigma_1 \rho_1 &amp; \sigma_2^2 &amp; \sigma_2 \sigma_3 \rho_2 &amp; \sigma_2 \sigma_4 \rho_3 \ \sigma_3 \sigma_2 \rho_1 \rho_2 &amp; \sigma_3 \sigma_3 \rho_1 &amp; \sigma_3^2 &amp; \sigma_3 \sigma_4 \rho_1 \ \sigma_4 \sigma_3 \rho_1 \rho_2 &amp; \sigma_4 \sigma_3 \rho_2 \rho_3 &amp; \sigma_4 \sigma_3 \rho_2 &amp; \sigma_4^2 \end{bmatrix}$</td>
</tr>
</tbody>
</table>
The following provides some further information about these covariance structures:

**ANTE(1)** specifies the first-order antedependence structure (see Kenward 1987; Patel 1991; Macchiavelli and Arnold 1994). In Table 63.17, $\sigma_i^2$ is the $i$th variance parameter, and $\rho_k$ is the $k$th autocorrelation parameter satisfying $|\rho_k| < 1$.

**AR(1)** specifies a first-order autoregressive structure. PROC MIXED imposes the constraint $|\rho| < 1$ for stationarity.

**ARH(1)** specifies a heterogeneous first-order autoregressive structure. As with TYPE=AR(1), PROC MIXED imposes the constraint $|\rho| < 1$ for stationarity.

**ARMA(1,1)** specifies the first-order autoregressive moving-average structure. In Table 63.17, $\rho$ is the autoregressive parameter, $\gamma$ models a moving-average component, and $\sigma^2$ is the residual variance. In the notation of Fuller (1976, p. 68), $D_1$ and $D_2$.

$$\gamma = \frac{(1 + b_1\theta_1)(\theta_1 + b_1)}{1 + b_1^2 + 2b_1\theta_1}$$

The example in Table 63.19 and $|b_1| < 1$ imply that

$$b_1 = \frac{\beta - \sqrt{\beta^2 - 4\alpha^2}}{2\alpha}$$

where $\alpha = \gamma - \rho$ and $\beta = 1 + \rho^2 - 2\gamma\rho$. PROC MIXED imposes the constraints $|\rho| < 1$ and $|\gamma| < 1$ for stationarity, although for some values of $\rho$ and $\gamma$ in this region the resulting covariance matrix is not positive definite. When the estimated value of $\rho$ becomes negative, the computed covariance is multiplied by $\cos(\pi d_{ij})$ to account for the negativity.

**CS** specifies the compound-symmetry structure, which has constant variance and constant covariance.
Chapter 63: The MIXED Procedure

CSH specifies the heterogeneous compound-symmetry structure. This structure has a different variance parameter for each diagonal element, and it uses the square roots of these parameters in the off-diagonal entries. In Table 63.17, $\sigma_i^2$ is the $i$th variance parameter, and $\rho$ is the correlation parameter satisfying $|\rho| < 1$.

FA($q$) specifies the factor-analytic structure with $q$ factors (Jennrich and Schluchter 1986). This structure is of the form $\mathbf{A}\mathbf{A}^t + \mathbf{D}$, where $\mathbf{A}$ is a $t \times q$ rectangular matrix and $\mathbf{D}$ is a $t \times t$ diagonal matrix with $t$ different parameters. When $q > 1$, the elements of $\mathbf{A}$ in its upper-right corner (that is, the elements in the $i$th row and $j$th column for $j > i$) are set to zero to fix the rotation of the structure.

FA0($q$) is similar to the FA($q$) structure except that no diagonal matrix $\mathbf{D}$ is included. When $q < t$—that is, when the number of factors is less than the dimension of the matrix—this structure is nonnegative definite but not of full rank. In this situation, you can use it for approximating an unstructured $\mathbf{G}$ matrix in the RANDOM statement or for combining with the LOCAL option in the REPEATED statement. When $q = t$, you can use this structure to constrain $\mathbf{G}$ to be nonnegative definite in the RANDOM statement.

FA1($q$) is similar to the TYPE=FA($q$) structure except that all of the elements in $\mathbf{D}$ are constrained to be equal. This offers a useful and more parsimonious alternative to the full factor-analytic structure.

HF specifies the Huynh-Feldt covariance structure (Huynh and Feldt 1970). This structure is similar to the TYPE=CSH structure in that it has the same number of parameters and heterogeneity along the main diagonal. However, it constructs the off-diagonal elements by taking arithmetic rather than geometric means.

You can perform a likelihood ratio test of the Huynh-Feldt conditions by running PROC MIXED twice, once with TYPE=HF and once with TYPE=UN, and then subtracting their respective values of $-2$ times the maximized likelihood.

If PROC MIXED does not converge under your Huynh-Feldt model, you can specify your own starting values with the PARMS statement. The default MIVQUE(0) starting values can sometimes be poor for this structure. A good choice for starting values is often the parameter estimates corresponding to an initial fit that uses TYPE=CS.

LIN($q$) specifies the general linear covariance structure with $q$ parameters. This structure consists of a linear combination of known matrices that are input with the LDATA= option. This structure is very general, and you need to make sure that the variance matrix is positive definite. By default, PROC MIXED sets the initial values of the parameters to 1. You can use the PARMS statement to specify other initial values.

LINEAR($q$) is an alias for TYPE=LIN($q$).

SIMPLE is an alias for TYPE=VC.

SP(EXPA)($c$-list) specifies the spatial anisotropic exponential structure, where $c$-list is a list of variables indicating the coordinates. This structure has $(i, j)$ element equal to

$$
\sigma^2 \prod_{k=1}^{c} \exp\{-\theta_k d(i, j, k)^{p_k}\}
$$
where \( c \) is the number of coordinates and \( d(i, j, k) \) is the absolute distance between the \( k \)th coordinate \((k = 1, \ldots, c)\) of the \( i \)th and \( j \)th observations in the input data set. There are \( 2c + 1 \) parameters to be estimated: \( \theta_k \), \( p_k \) \((k = 1, \ldots, c)\), and \( \sigma^2 \).

You might want to constrain some of the EXPA parameters to known values. For example, suppose you have three coordinate variables \( C1, C2, \) and \( C3 \) and you want to constrain the powers \( p_k \) to equal 2, as in Sacks et al. (1989). Suppose further that you want to model covariance across the entire input data set and you suspect the \( \theta_k \) and \( \sigma^2 \) estimates are close to 3, 4, 5, and 1, respectively. Then specify the following statements:

```plaintext
repeated / type=sp(expa) (c1 c2 c3)
subject=intercept;
pParms (3) (4) (5) (2) (2) (2) (1) / hold=4,5,6;
```

**SP(EXPGA)**\((c1 c2)\) specify modification of the isotropic SP(EXP) covariance structure.

**SP(GAUGA)**\((c1 c2)\) specify modification of the isotropic SP(GAU) covariance structure.

**SP(SPHGA)**\((c1 c2)\) specify modification of the isotropic SP(SPH) covariance structure.

These are structures that allow for geometric anisotropy in two dimensions. The coordinates are specified by the variables \( c1 \) and \( c2 \).

If the spatial process is geometrically anisotropic in \( e = [c_{i1}, c_{i2}] \), then it is isotropic in the coordinate system

\[
A_e = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} e = e^*
\]

for a properly chosen angle \( \theta \) and scaling factor \( \lambda \). Elliptical isocorrelation contours are thereby transformed to spherical contours, adding two parameters to the respective isotropic covariance structures. Euclidean distances (see Table 63.18) are expressed in terms of \( e^* \).

The angle \( \theta \) of the clockwise rotation is reported in radians, \( 0 \leq \theta \leq 2\pi \). The scaling parameter \( \lambda \) represents the ratio of the range parameters in the direction of the major and minor axis of the correlation contours. In other words, following a rotation of the coordinate system by angle \( \theta \), isotropy is achieved by compressing or magnifying distances in one coordinate by the factor \( \lambda \).

Fixing \( \lambda = 1.0 \) reduces the models to isotropic ones for any angle of rotation. If the scaling parameter is held constant at 1.0, you should also hold constant the angle of rotation, as in the following statements:

```plaintext
repeated / type=sp(expga) (gxc gyc)
subject=intercept;
pParms (6) (1.0) (0.0) (1) / hold=2,3;
```

If \( \lambda \) is fixed at any other value than 1.0, the angle of rotation can be estimated. Specifying a starting grid of angles and scaling factors can considerably improve the convergence properties of the optimization algorithm for these models. Only a single random effect with geometrically anisotropic structure is permitted.
SP(MATERN)\((c\text{-list})\) | SP(MATHSW)\((c\text{-list})\) specifies covariance structures in the Matérn class of covariance functions (Matérn 1986). Two observations for the same subject (block of R) that are Euclidean distance \(d_{ij}\) apart have covariance
\[
\sigma^2 \frac{1}{\Gamma(v)} \left( \frac{d_{ij}}{2\rho} \right)^v 2K_v(d_{ij}/\rho) \quad v > 0, \ \rho > 0
\]
where \(K_v\) is the modified Bessel function of the second kind of (real) order \(v > 0\). The smoothness (continuity) of a stochastic process with covariance function in this class increases with \(v\). The Matérn class thus enables data-driven estimation of the smoothness properties. The covariance is identical to the exponential model for \(v = 0.5\) (TYPE=SP(EXP)(c-list)), while for \(v = 1\) the model advocated by Whittle (1954) results. As \(v \to \infty\) the model approaches the gaussian covariance structure (TYPE=SP(GAU)(c-list)).

The MATHSW structure represents the Matérn class in the parameterization of Handcock and Stein (1993) and Handcock and Wallis (1994),
\[
\sigma^2 \frac{1}{\Gamma(v)} \left( \frac{d_{ij} \sqrt{v}}{\rho} \right)^v 2K_v \left( \frac{2d_{ij} \sqrt{v}}{\rho} \right)
\]
Since computation of the function \(K_v\) and its derivatives is numerically very intensive, fitting models with Matérn covariance structures can be more time-consuming than with other spatial covariance structures. Good starting values are essential.

SP(POW)\((c\text{-list})\) | SP(POWA)\((c\text{-list})\) specifies the spatial power structures. When the estimated value of \(\rho\) becomes negative, the computed covariance is multiplied by \(\cos(\pi d_{ij})\) to account for the negativity.

TOEP\(<(q)\rangle\) specifies a banded Toeplitz structure. This can be viewed as a moving-average structure with order equal to \(q - 1\). The TYPE=TOEP option is a full Toeplitz matrix, which can be viewed as an autoregressive structure with order equal to the dimension of the matrix. The specification TYPE=TOEP(1) is the same as \(\sigma^2 I\), where \(I\) is an identity matrix, and it can be useful for specifying the same variance component for several effects.

TOEPH\(<(q)\rangle\) specifies a heterogeneous banded Toeplitz structure. In Table 63.17, \(\sigma_i^2\) is the \(i\)th variance parameter and \(\rho_j\) is the \(j\)th correlation parameter satisfying \(|\rho_j| < 1\). If you specify the order parameter \(q\), then PROC MIXED estimates only the first \(q\) bands of the matrix, setting all higher bands equal to 0. The option TOEPH(1) is equivalent to both the TYPE=UN(1) and TYPE=UNR(1) options.

UN\(<(q)\rangle\) specifies a completely general (unstructured) covariance matrix parameterized directly in terms of variances and covariances. The variances are constrained to be nonnegative, and the covariances are unconstrained. This structure is not constrained to be nonnegative definite in order to avoid nonlinear constraints; however, you can use the TYPE=FA0 structure if you want this constraint to be imposed by a Cholesky factorization. If you specify the order parameter \(q\), then PROC MIXED estimates only the first \(q\) bands of the matrix, setting all higher bands equal to 0.

UNR\(<(q)\rangle\) specifies a completely general (unstructured) covariance matrix parameterized in terms of variances and correlations. This structure fits the same model as
the TYPE=UN\((q)\) option but with a different parameterization. The \(i\)th variance parameter is \(\sigma_i^2\). The parameter \(\rho_{jk}\) is the correlation between the \(j\)th and \(k\)th measurements; it satisfies \(|\rho_{jk}| < 1\). If you specify the order parameter \(r\), then PROC MIXED estimates only the first \(q\) bands of the matrix, setting all higher bands equal to zero.

UN@AR(1) | UN@CS | UN@UN specify direct (Kronecker) product structures designed for multivariate repeated measures (see Galecki 1994). These structures are constructed by taking the Kronecker product of an unstructured matrix (modeling covariance across the multivariate observations) with an additional covariance matrix (modeling covariance across time or another factor). The upper-left value in the second matrix is constrained to equal 1 to identify the model. See the SAS/IML User’s Guide for more details about direct products.

To use these structures in the REPEATED statement, you must specify two distinct REPEATED effects, both of which must be included in the CLASS statement. The first effect indicates the multivariate observations, and the second identifies the levels of time or some additional factor. Note that the input data set must still be constructed in “univariate” format; that is, all dependent observations are still listed observation-wise in one single variable. Although this construction provides for general modeling possibilities, it forces you to construct variables indicating both dimensions of the Kronecker product.

For example, suppose your observed data consist of heights and weights of several children measured over several successive years. Your input data set should then contain variables similar to the following:

- \(Y\), all of the heights and weights, with a separate observation for each
- \(\text{Var}\), indicating whether the measurement is a height or a weight
- \(\text{Year}\), indicating the year of measurement
- \(\text{Child}\), indicating the child on which the measurement was taken

Your PROC MIXED statements for a Kronecker AR(1) structure across years would then be as follows:

```sas
proc mixed;
    class Var Year Child;
    model Y = Var Year Var*Year;
    repeated Var Year / type=un@ar(1)
        subject=Child;
run;
```

You should nearly always want to model different means for the multivariate observations; hence the inclusion of \(\text{Var}\) in the MODEL statement. The preceding mean model consists of cell means for all combinations of \(\text{VAR}\) and \(\text{YEAR}\).

VC specifies standard variance components and is the default structure for both the RANDOM and REPEATED statements. In the RANDOM statement, a distinct variance component is assigned to each effect. In the REPEATED statement, this structure is usually used only with the GROUP= option to specify a heterogeneous variance model.

**SLICE Statement**

```plaintext
SLICE model-effect < / options >;
```

The SLICE statement provides a general mechanism for performing a partitioned analysis of the LS-means for an interaction. This analysis is also known as an analysis of simple effects.

The SLICE statement uses the same `options` as the LSMEANS statement, which are summarized in Table 19.21. For details about the syntax of the SLICE statement, see the section “SLICE Statement” on page 505 in Chapter 19, “Shared Concepts and Topics.”

**NOTE:** Use the section “LSMEANS Statement” on page 460 in Chapter 19, “Shared Concepts and Topics.” only for definitions of the options that you can use with the SLICE statement. PROC MIXED uses a slightly different syntax for the LSMEANS, which is described in the section “LSMEANS Statement” on page 5265.

**STORE Statement**

```plaintext
STORE <OUT= item-store-name < / LABEL='label' > ;
```

The STORE statement requests that the procedure save the context and results of the statistical analysis. The resulting item store has a binary file format that cannot be modified. The contents of the item store can be processed with the PLM procedure.

For details about the syntax of the STORE statement, see the section “STORE Statement” on page 508 in Chapter 19, “Shared Concepts and Topics.”

**WEIGHT Statement**

```plaintext
WEIGHT variable ;
```

If you do not specify a REPEATED statement, the WEIGHT statement operates exactly like the one in PROC GLM. In this case PROC MIXED replaces $X'X$ and $Z'Z$ with $X'WX$ and $Z'WZ$, where $W$ is the diagonal weight matrix. If you specify a REPEATED statement, then the WEIGHT statement replaces $R$ with $LRL$, where $L$ is a diagonal matrix with elements $W^{-1/2}$. Observations with nonpositive or missing weights are not included in the PROC MIXED analysis.

If a computation in PROC MIXED involves $R$, then the WEIGHT statement replaces $R$ with $W^{-1/2}RW^{-1/2}$. For example, the covariance matrix $V$ for the observations usually have the form $V = ZGZ' + R$, which with the WEIGHT statement becomes $V = ZGZ' + W^{-1/2}RW^{-1/2}$.
Details: MIXED Procedure

Mixed Models Theory

This section provides an overview of a likelihood-based approach to general linear mixed models. This approach simplifies and unifies many common statistical analyses, including those involving repeated measures, random effects, and random coefficients. The basic assumption is that the data are linearly related to unobserved multivariate normal random variables. For extensions to nonlinear and nonnormal situations see the documentation of the GLIMMIX and NLMIXED procedures. Additional theory and examples are provided in Littell et al. (2006); Verbeke and Molenberghs (1997, 2000); Brown and Prescott (1999).

Matrix Notation

Suppose that you observe $n$ data points $y_1, \ldots, y_n$ and that you want to explain them by using $n$ values for each of $p$ explanatory variables $x_{11}, \ldots, x_{1p}$, $x_{21}, \ldots, x_{2p}$, $\ldots$, $x_{n1}, \ldots, x_{np}$. The $x_{ij}$ values can be either regression-type continuous variables or dummy variables indicating class membership. The standard linear model for this setup is

$$y_i = \sum_{j=1}^{p} x_{ij} \beta_j + \epsilon_i \quad i = 1, \ldots, n$$

where $\beta_1, \ldots, \beta_p$ are unknown fixed-effects parameters to be estimated and $\epsilon_1, \ldots, \epsilon_n$ are unknown independent and identically distributed normal (Gaussian) random variables with mean 0 and variance $\sigma^2$.

The preceding equations can be written simultaneously by using vectors and a matrix, as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

For convenience, simplicity, and extendability, this entire system is written as

$$y = X\beta + \epsilon$$

where $y$ denotes the vector of observed $y_i$’s, $X$ is the known matrix of $x_{ij}$’s, $\beta$ is the unknown fixed-effects parameter vector, and $\epsilon$ is the unobserved vector of independent and identically distributed Gaussian random errors.

In addition to denoting data, random variables, and explanatory variables in the preceding fashion, the subsequent development makes use of basic matrix operators such as transpose (‘), inverse (^{-1}), generalized inverse (\infty), determinant (\mid \cdot\mid), and matrix multiplication. See Searle (1982) for details about these and other matrix techniques.
Formulation of the Mixed Model

The previous general linear model is certainly a useful one (Searle 1971), and it is the one fitted by the GLM procedure. However, many times the distributional assumption about $\epsilon$ is too restrictive. The mixed model extends the general linear model by allowing a more flexible specification of the covariance matrix of $\epsilon$. In other words, it allows for both correlation and heterogeneous variances, although you still assume normality.

The mixed model is written as

$$y = X\beta + Z\gamma + \epsilon$$

where everything is the same as in the general linear model except for the addition of the known design matrix, $Z$, and the vector of unknown random-effects parameters, $\gamma$. The matrix $Z$ can contain either continuous or dummy variables, just like $X$. The name mixed model comes from the fact that the model contains both fixed-effects parameters, $\beta$, and random-effects parameters, $\gamma$. See Henderson (1990) and Searle, Casella, and McCulloch (1992) for historical developments of the mixed model.

A key assumption in the foregoing analysis is that $\gamma$ and $\epsilon$ are normally distributed with

$$\begin{bmatrix} \gamma \\ \epsilon \end{bmatrix} \sim \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\text{Var} \begin{bmatrix} \gamma \\ \epsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}.$$

The variance of $y$ is, therefore, $V = ZGZ' + R$. You can model $V$ by setting up the random-effects design matrix $Z$ and by specifying covariance structures for $G$ and $R$.

Note that this is a general specification of the mixed model, in contrast to many texts and articles that discuss only simple random effects. Simple random effects are a special case of the general specification with $Z$ containing dummy variables, $G$ containing variance components in a diagonal structure, and $R = \sigma^2 I_n$, where $I_n$ denotes the $n \times n$ identity matrix. The general linear model is a further special case with $Z = 0$ and $R = \sigma^2 I_n$.

The following two examples illustrate the most common formulations of the general linear mixed model.

**Example: Growth Curve with Compound Symmetry**

Suppose that you have three growth curve measurements for $s$ individuals and that you want to fit an overall linear trend in time. Your $X$ matrix is as follows:

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}.$$ 

The first column (coded entirely with 1s) fits an intercept, and the second column (coded with times of 1, 2, 3) fits a slope. Here, $n = 3s$ and $p = 2$.

Suppose further that you want to introduce a common correlation among the observations from a single individual, with correlation being the same for all individuals. One way of setting this up in the general mixed
model is to eliminate the $Z$ and $G$ matrices and let the $R$ matrix be block diagonal with blocks corresponding to the individuals and with each block having the compound-symmetry structure. This structure has two unknown parameters, one modeling a common covariance and the other modeling a residual variance. The form for $R$ would then be as follows:

$$
R = \begin{bmatrix}
\sigma_1^2 + \sigma^2 & \sigma_1^2 & \ldots & \sigma_1^2 \\
\sigma_1^2 & \sigma_1^2 + \sigma^2 & \sigma_1^2 & \sigma_1^2 \\
\sigma_1^2 & \sigma_1^2 & \ldots & \sigma_1^2 \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_1^2 & \sigma_1^2 & \ldots & \sigma_1^2 + \sigma^2 \\
\sigma_1^2 & \sigma_1^2 & \ldots & \sigma_1^2 \\
\end{bmatrix}
$$

where blanks denote zeros. There are $3s$ rows and columns altogether, and the common correlation is $\sigma_1^2/(\sigma_1^2 + \sigma^2)$.

The PROC MIXED statements to fit this model are as follows:

```plaintext
proc mixed;
  class indiv;
  model y = time;
  repeated / type=cs subject=indiv;
run;
```

Here, `indiv` is a classification variable indexing individuals. The `MODEL` statement fits a straight line for `time`; the intercept is fit by default just as in PROC GLM. The `REPEATED` statement models the $R$ matrix: `TYPE=CS` specifies the compound symmetry structure, and `SUBJECT=INDIV` specifies the blocks of $R$.

An alternative way of specifying the common intra-individual correlation is to let

$$
Z = \begin{bmatrix}
1 \\
1 \\
1 \\
\vdots \\
1 \\
1 \\
\end{bmatrix}
$$

$$
G = \begin{bmatrix}
\sigma_1^2 \\
\sigma_1^2 \\
\vdots \\
\sigma_1^2 \\
\end{bmatrix}
$$

and $R = \sigma_1^2 I_n$. The $Z$ matrix has $3s$ rows and $s$ columns, and $G$ is $s \times s$.

You can set up this model in PROC MIXED in two different but equivalent ways:
Both of these specifications fit the same model as the previous one that used the REPEATED statement; however, the RANDOM specifications constrain the correlation to be positive, whereas the REPEATED specification leaves the correlation unconstrained.

**Example: Split-Plot Design**

The split-plot design involves two experimental treatment factors, A and B, and two different sizes of experimental units to which they are applied (see Winer 1971; Snedecor and Cochran 1980; Milliken and Johnson 1992; Steel, Torrie, and Dickey 1997). The levels of A are randomly assigned to the larger-sized experimental unit, called whole plots, whereas the levels of B are assigned to the smaller-sized experimental unit, the subplots. The subplots are assumed to be nested within the whole plots, so that a whole plot consists of a cluster of subplots and a level of A is applied to the entire cluster.

Such an arrangement is often necessary by nature of the experiment, the classical example being the application of fertilizer to large plots of land and different crop varieties planted in subdivisions of the large plots. For this example, fertilizer is the whole-plot factor A and variety is the subplot factor B.

The first example is a split-plot design for which the whole plots are arranged in a randomized block design. The appropriate PROC MIXED statements are as follows:

```sas
proc mixed;
  class a b block;
  model y = a|b;
  random block a*block;
run;
```

Here

\[ R = \sigma^2 I_{24} \]
and $X$, $Z$, and $G$ have the following form:

$$X = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
\end{bmatrix}$$

$$Z = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
\end{bmatrix}$$

$$G = \begin{bmatrix}
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\
\end{bmatrix}$$
where $\sigma_B^2$ is the variance component for Block and $\sigma_{AB}^2$ is the variance component for A*Block. Changing the RANDOM statement as follows fits the same model, but with Z and G sorted differently:

```
random int a / subject=block;
```

Estimating Covariance Parameters in the Mixed Model

Estimation is more difficult in the mixed model than in the general linear model. Not only do you have $\beta$ as in the general linear model, but you have unknown parameters in $\gamma$, $G$, and $R$ as well. Least squares is no longer the best method. Generalized least squares (GLS) is more appropriate, minimizing

$$(y - X\beta)'V^{-1}(y - X\beta)$$
However, it requires knowledge of \( V \) and, therefore, knowledge of \( G \) and \( R \). Lacking such information, one approach is to use estimated GLS, in which you insert some reasonable estimate for \( V \) into the minimization problem. The goal thus becomes finding a reasonable estimate of \( G \) and \( R \).

In many situations, the best approach is to use likelihood-based methods, exploiting the assumption that \( \mathbf{y} \) and \( \varepsilon \) are normally distributed (Hartley and Rao 1967; Patterson and Thompson 1971; Harville 1977; Laird and Ware 1982; Jennrich and Schluchter 1986). PROC MIXED implements two likelihood-based methods: maximum likelihood (ML) and restricted/residual maximum likelihood (REML). A favorable theoretical property of ML and REML is that they accommodate data that are missing at random (Rubin 1976; Little 1995).

PROC MIXED constructs an objective function associated with ML or REML and maximizes it over all unknown parameters. Using calculus, it is possible to reduce this maximization problem to one over only the parameters in \( G \) and \( R \). The corresponding log-likelihood functions are as follows:

\[
\text{ML} : \quad l(G, R) = -\frac{1}{2} \log |V| - \frac{1}{2} \mathbf{r}'V^{-1}\mathbf{r} - \frac{n}{2} \log(2\pi)
\]

\[
\text{REML} : \quad l_R(G, R) = -\frac{1}{2} \log |V| - \frac{1}{2} \log |X'V^{-1}X| - \frac{1}{2} \mathbf{r}'V^{-1}\mathbf{r} - \frac{n - p}{2} \log(2\pi)
\]

where \( \mathbf{r} = \mathbf{y} - X(X'V^{-1}X)^{-1}X'V^{-1}\mathbf{y} \) and \( p \) is the rank of \( X \). PROC MIXED actually minimizes \(-2\) times these functions by using a ridge-stabilized Newton-Raphson algorithm. Lindstrom and Bates (1988) provide reasons for preferring Newton-Raphson to the Expectation-Maximum (EM) algorithm (Dempster, Laird, and Rubin 1977; Laird, Lange, and Stram 1987), as well as analytical details for implementing a QR-decomposition approach to the problem. Wolfinger, Tobias, and Sall (1994) present the sweep-based algorithms that are implemented in PROC MIXED.

One advantage of using the Newton-Raphson algorithm is that the second derivative matrix of the objective function evaluated at the optima is available upon completion. Denoting this matrix \( H \), the asymptotic theory of maximum likelihood (see Serfling 1980) shows that \( 2H^{-1} \) is an asymptotic variance-covariance matrix of the estimated parameters of \( G \) and \( R \). Thus, tests and confidence intervals based on asymptotic normality can be obtained. However, these can be unreliable in small samples, especially for parameters such as variance components that have sampling distributions that tend to be skewed to the right.

If a residual variance \( \sigma^2 \) is a part of your mixed model, it can usually be profiled out of the likelihood. This means solving analytically for the optimal \( \sigma^2 \) and plugging this expression back into the likelihood formula (see Wolfinger, Tobias, and Sall 1994). This reduces the number of optimization parameters by one and can improve convergence properties. PROC MIXED profiles the residual variance out of the log likelihood whenever it appears reasonable to do so. This includes the case when \( R \) equals \( \sigma^2 I \) and when it has blocks with a compound symmetry, time series, or spatial structure. PROC MIXED does not profile the log likelihood when \( R \) has unstructured blocks, when you use the HOLD= or NOITER option in the PARMS statement, or when you use the NOPROFILE option in the PROC MIXED statement.

Instead of ML or REML, you can use the noniterative MIVQUE0 method to estimate \( G \) and \( R \) (Rao 1972; LaMotte 1973; Wolfinger, Tobias, and Sall 1994). In fact, by default PROC MIXED uses MIVQUE0 estimates as starting values for the ML and REML procedures. For variance component models, another estimation method involves equating Type 1, 2, or 3 expected mean squares to their observed values and solving the resulting system. However, Swallow and Monahan (1984) present simulation evidence favoring REML and ML over MIVQUE0 and other method-of-moment estimators.
Estimating Fixed and Random Effects in the Mixed Model

ML, REML, MIVQUE0, or Type1–Type3 provide estimates of $G$ and $R$, which are denoted $\hat{G}$ and $\hat{R}$, respectively. To obtain estimates of $\beta$ and $\gamma$, the standard method is to solve the mixed model equations (Henderson 1984):

$$
\begin{bmatrix}
X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z \\
Z'\hat{R}^{-1}X & Z'\hat{R}^{-1}Z + \hat{G}^{-1}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta} \\
\hat{\gamma}
\end{bmatrix}
= 
\begin{bmatrix}
X'\hat{R}^{-1}y \\
Z'\hat{R}^{-1}y
\end{bmatrix}
$$

The solutions can also be written as

$$
\hat{\beta} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y
$$

$$
\hat{\gamma} = \hat{G}Z'\hat{V}^{-1}(y - X\hat{\beta})
$$

and have connections with empirical Bayes estimators (Laird and Ware 1982; Carlin and Louis 1996).

Note that the mixed model equations are extended normal equations and that the preceding expression assumes that $\hat{G}$ is nonsingular. For the extreme case where the eigenvalues of $\hat{G}$ are very large, $\hat{G}^{-1}$ contributes very little to the equations and $\hat{\gamma}$ is close to what it would be if $\gamma$ actually contained fixed-effects parameters. On the other hand, when the eigenvalues of $\hat{G}$ are very small, $\hat{G}^{-1}$ dominates the equations and $\hat{\gamma}$ is close to 0. For intermediate cases, $\hat{G}^{-1}$ can be viewed as shrinking the fixed-effects estimates of $\gamma$ toward 0 (Robinson 1991).

If $\hat{G}$ is singular, then the mixed model equations are modified (Henderson 1984) as follows:

$$
\begin{bmatrix}
X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z \\
\hat{G}'Z'\hat{R}^{-1}X & \hat{G}'Z'\hat{R}^{-1}Z\hat{G} + \hat{G}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta} \\
\hat{\tau}
\end{bmatrix}
= 
\begin{bmatrix}
X'\hat{R}^{-1}y \\
\hat{G}'Z'\hat{R}^{-1}y
\end{bmatrix}
$$

Denote the generalized inverses of the nonsingular $\hat{G}$ and singular $\hat{G}$ forms of the mixed model equations by $C$ and $M$, respectively. In the nonsingular case, the solution $\hat{\gamma}$ estimates the random effects directly, but in the singular case the estimates of random effects are achieved through a back-transformation $\hat{\gamma} = \hat{G}\hat{\tau}$ where $\hat{\tau}$ is the solution to the modified mixed model equations. Similarly, while in the nonsingular case $C$ itself is the estimated covariance matrix for $(\hat{\beta}, \hat{\gamma})$, in the singular case the covariance estimate for $(\hat{\beta}, \hat{G}\hat{\tau})$ is given by $PMP$ where

$$
P = 
\begin{bmatrix}
I \\
\hat{G}
\end{bmatrix}
$$

An example of when the singular form of the equations is necessary is when a variance component estimate falls on the boundary constraint of 0.

Model Selection

The previous section on estimation assumes the specification of a mixed model in terms of $X$, $Z$, $G$, and $R$. Even though $X$ and $Z$ have known elements, their specific form and construction are flexible, and several possibilities can present themselves for a particular data set. Likewise, several different covariance structures for $G$ and $R$ might be reasonable.
Space does not permit a thorough discussion of model selection, but a few brief comments and references are in order. First, subject matter considerations and objectives are of great importance when selecting a model; see Diggle (1988) and Lindsey (1993).

Second, when the data themselves are looked to for guidance, many of the graphical methods and diagnostics appropriate for the general linear model extend to the mixed model setting as well (Christensen, Pearson, and Johnson 1992).

Finally, a likelihood-based approach to the mixed model provides several statistical measures for model adequacy as well. The most common of these are the likelihood ratio test and Akaike’s and Schwarz’s criteria (Bozdogan 1987; Wolfinger 1993; Keselman et al. 1998, 1999).

### Statistical Properties

If \( G \) and \( R \) are known, \( \hat{\beta} \) is the best linear unbiased estimator (BLUE) of \( \beta \), and \( \hat{\gamma} \) is the best linear unbiased predictor (BLUP) of \( \gamma \) (Searle 1971; Harville 1988, 1990; Robinson 1991; McLean, Sanders, and Stroup 1991). Here, “best” means minimum mean squared error. The covariance matrix of \( (\hat{\beta} - \beta, \hat{\gamma} - \gamma) \) is

\[
\hat{C} = \begin{bmatrix}
X'R^{-1}X & X'R^{-1}Z \\
Z'R^{-1}X & Z'R^{-1}Z + G^{-1}
\end{bmatrix}
\]

where \( \hat{C} \) denotes a generalized inverse (see Searle 1971).

However, \( G \) and \( R \) are usually unknown and are estimated by one of the aforementioned methods. These estimates, \( \hat{G} \) and \( \hat{R} \), are therefore simply substituted into the preceding expression to obtain

\[
\hat{\hat{C}} = \begin{bmatrix}
X\hat{R}^{-1}X & X\hat{R}^{-1}Z \\
Z\hat{R}^{-1}X & Z\hat{R}^{-1}Z + \hat{G}^{-1}
\end{bmatrix}
\]

as the approximate variance-covariance matrix of \( (\hat{\beta} - \beta, \hat{\gamma} - \gamma) \). In this case, the BLUE and BLUP acronyms no longer apply, but the word empirical is often added to indicate such an approximation. The appropriate acronyms thus become EBLUE and EBLUP.

McLean and Sanders (1988) show that \( \hat{\hat{C}} \) can also be written as

\[
\hat{\hat{C}} = \begin{bmatrix}
\hat{C}_{11} & \hat{C}_{21} \\
\hat{C}_{21} & \hat{C}_{22}
\end{bmatrix}
\]

where

\[
\hat{C}_{11} = (X\hat{V}^{-1}X)^{-1} \\
\hat{C}_{21} = -\hat{G}Z\hat{V}^{-1}X\hat{C}_{11} \\
\hat{C}_{22} = (Z\hat{R}^{-1}Z + \hat{G}^{-1})^{-1} - \hat{C}_{21}X\hat{V}^{-1}Z\hat{G}
\]

Note that \( \hat{C}_{11} \) is the familiar estimated generalized least squares formula for the variance-covariance matrix of \( \hat{\beta} \).

As a cautionary note, \( \hat{\hat{C}} \) tends to underestimate the true sampling variability of \( (\hat{\beta}, \hat{\gamma}) \) because no account is made for the uncertainty in estimating \( G \) and \( R \). Although inflation factors have been proposed (Kackar and Harville 1984; Kass and Steffey 1989; Prasad and Rao 1990), they tend to be small for data sets that are fairly well balanced. PROC MIXED does not compute any inflation factors by default, but rather accounts for the downward bias by using the approximate \( t \) and \( F \) statistics described subsequently. The DDFM=KENWARDROGER option in the MODEL statement prompts PROC MIXED to compute a specific inflation factor along with Satterthwaite-based degrees of freedom.
Inference and Test Statistics

For inferences concerning the covariance parameters in your model, you can use likelihood-based statistics. One common likelihood-based statistic is the Wald Z, which is computed as the parameter estimate divided by its asymptotic standard error. The asymptotic standard errors are computed from the inverse of the second derivative matrix of the likelihood with respect to each of the covariance parameters. The Wald Z is valid for large samples, but it can be unreliable for small data sets and for parameters such as variance components, which are known to have a skewed or bounded sampling distribution.

A better alternative is the likelihood ratio $\chi^2$ statistic. This statistic compares two covariance models, one a special case of the other. To compute it, you must run PROC MIXED twice, once for each of the two models, and then subtract the corresponding values of $-2$ times the log likelihoods. You can use either ML or REML to construct this statistic, which tests whether the full model is necessary beyond the reduced model.

As long as the reduced model does not occur on the boundary of the covariance parameter space, the $\chi^2$ statistic computed in this fashion has a large-sample $\chi^2$ distribution that is $\chi^2$ with degrees of freedom equal to the difference in the number of covariance parameters between the two models. If the reduced model does occur on the boundary of the covariance parameter space, the asymptotic distribution becomes a mixture of $\chi^2$ distributions (Self and Liang 1987). A common example of this is when you are testing that a variance component equals its lower boundary constraint of 0.

A final possibility for obtaining inferences concerning the covariance parameters is to simulate or resample data from your model and construct empirical sampling distributions of the parameters. The SAS macro language and the ODS system are useful tools in this regard.

$F$ and $t$ Tests for Fixed- and Random-Effects Parameters

For inferences concerning the fixed- and random-effects parameters in the mixed model, consider estimable linear combinations of the following form:

$$L\begin{bmatrix} \beta \\ \gamma \end{bmatrix}$$

The estimability requirement (Searle 1971) applies only to the $\beta$ portion of $L$, because any linear combination of $\gamma$ is estimable. Such a formulation in terms of a general $L$ matrix encompasses a wide variety of common inferential procedures such as those employed with Type 1–Type 3 tests and LS-means. The CONTRAST and ESTIMATE statements in PROC MIXED enable you to specify your own $L$ matrices. Typically, inference on fixed effects is the focus, and, in this case, the $\gamma$ portion of $L$ is assumed to contain all 0s.

Statistical inferences are obtained by testing the hypothesis

$$H : L\begin{bmatrix} \beta \\ \gamma \end{bmatrix} = 0$$

or by constructing point and interval estimates.

When $L$ consists of a single row, a general $t$ statistic can be constructed as follows (see McLean and Sanders 1988; Stroup 1989a):

$$t = \frac{L\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix}}{\sqrt{LCL'}}$$
Under the assumed normality of $\gamma$ and $\epsilon$, $t$ has an exact $t$ distribution only for data exhibiting certain types of balance and for some special unbalanced cases. In general, $t$ is only approximately $t$-distributed, and its degrees of freedom must be estimated. See the DDFM= option for a description of the various degrees-of-freedom methods available in PROC MIXED.

With $\hat{\nu}$ being the approximate degrees of freedom, the associated confidence interval is

$$L \left[ \begin{array}{c} \hat{\beta} \\ \hat{\gamma} \end{array} \right] \pm t_{\hat{\nu},\alpha/2} \sqrt{L \hat{C} L'}$$

where $t_{\hat{\nu},\alpha/2}$ is the $(1 - \alpha/2)100$ percentile of the $t_{\hat{\nu}}$ distribution.

When the rank of $L$ is greater than 1, PROC MIXED constructs the following general $F$ statistic:

$$F = \left( \begin{array}{c} \hat{\beta} \\ \hat{\gamma} \end{array} \right)' L' (L \hat{C} L')^{-1} L \left( \begin{array}{c} \hat{\beta} \\ \hat{\gamma} \end{array} \right) / r$$

where $r = \text{rank}(L \hat{C} L')$. Analogous to $t$, $F$ in general has an approximate $F$ distribution with $r$ numerator degrees of freedom and $\hat{\nu}$ denominator degrees of freedom.

The $t$ and $F$ statistics enable you to make inferences about your fixed effects, which account for the variance-covariance model you select. An alternative is the $\chi^2$ statistic associated with the likelihood ratio test. This statistic compares two fixed-effects models, one a special case of the other. It is computed just as when comparing different covariance models, although you should use ML and not REML here because the penalty term associated with restricted likelihoods depends upon the fixed-effects specification.

**$F$ Tests With the ANOVAF Option**

The ANOVAF option computes $F$ tests by the following method in models with REPEATED statement and without RANDOM statement. Let $L$ denote the matrix of estimable functions for the hypothesis $H: L\beta = 0$, where $\beta$ are the fixed-effects parameters. Let $M = L' (L \hat{C})^{-1} L$, and suppose that $\hat{C}$ denotes the estimated variance-covariance matrix of $\hat{\beta}$ (see the section “Statistical Properties” for the construction of $\hat{C}$).

The ANOVAF $F$ statistics are computed as

$$F_A = \hat{\beta}' L' (L \hat{C})^{-1} L \hat{\beta} / t_1 = \hat{\beta}' M \hat{\beta} / t_1$$

Notice that this is a modification of the usual $F$ statistic where $(L \hat{C} L')^{-1}$ is replaced with $(L \hat{C})^{-1}$ and $\text{rank}(L)$ is replaced with $t_1 = \text{trace}(M \hat{C})$; see, for example, Brunner, Domhof, and Langer (2002, Sec. 5.4). The $p$-values for this statistic are computed from either an $F_{\nu_1,\nu_2}$ or an $F_{\nu_1,\infty}$ distribution. The respective degrees of freedom are determined by the MIXED procedure as follows:

$$\nu_1 = \frac{t_1^2}{\text{trace}(M \hat{C} M \hat{C})}$$

$$\nu_2 = \frac{2t_1^2}{g' A g}$$

$$\nu_2 = \begin{cases} \min\{\nu_2^*, df_e\}, 1 & g' A g > 1E3 \times \text{MACEPS} \\ 1 & \text{otherwise} \end{cases}$$
The term $g'Ag$ in the term $v_2^*$ for the denominator degrees of freedom is based on approximating $\text{Var}[\text{trace}(M\hat{C})]$ based on a first-order Taylor series about the true covariance parameters. This generalizes results in the appendix of Brunner, Dette, and Munk (1997) to a broader class of models. The vector $g = [g_1, \cdots, g_q]$ contains the partial derivatives

$$\text{trace}\left(L'(L')^{-1}L\frac{\partial \hat{C}}{\partial \theta_i}\right)$$

and A is the asymptotic variance-covariance matrix of the covariance parameter estimates (ASYCOV option).

PROC MIXED reports $v_1$ and $v_2$ as “NumDF” and “DenDF” under the “ANOVA F” heading in the output. The corresponding $p$-values are denoted as “Pr $> F(\text{DDF})$” for $F_{v_1,v_2}$ and “Pr $> F(\text{infty})$” for $F_{v_1,\infty}$, respectively.

$P$-values computed with the ANOVAF option can be identical to the nonparametric tests in Akritas, Arnold, and Brunner (1997) and in Brunner, Domhof, and Langer (2002), provided that the response data consist of properly created (and sorted) ranks and that the covariance parameters are estimated by MIVQUE0 in models with REPEATED statement and properly chosen SUBJECT= and/or GROUP= effects.

If you model an unstructured covariance matrix in a longitudinal model with one or more repeated factors, the ANOVAF results are identical to a multivariate MANOVA where degrees of freedom are corrected with the Greenhouse-Geisser adjustment (Greenhouse and Geisser 1959). For example, suppose that factor A has 2 levels and factor B has 4 levels. The following two sets of statements produce the same $p$-values:

```plaintext
proc mixed data=Mydata anovaf method=mivque0;
  class id A B;
  model score = A | B / chisq;
  repeated / type=un subject=id;
  ods select Tests3;
run;

proc transpose data=MyData out=tdata;
  by id;
  var score;
run;
proc glm data=tdata;
  model col: = / nouni;
  repeated A 2, B 4;
  ods output ModelANOVA=maov epsilons=eps;
run;
proc transpose data=eps(where=(substr(statistic,1,3)='Gre')) out=teps;
  var cvalue1;
run;
data aov; set maov;
  if (_n_ = 1) then merge teps;
  if (Source='A') then do;
    pFddf = ProbF;
    pFinf = 1 - probchi(df*Fvalue,df);
    output;
  end; else if (Source='B') then do;
    pFddf = ProbFGG;
    pFinf = 1 - probchi(df*col1*Fvalue,df*col1);
  end;
```
output;
end; else if (Source='A*B') then do;
    pfddF = ProbFGG;
    pFinf = 1 - probchi(df*col2*Fvalue,df*col2);
output;
end;
run;
proc print data=aov label noobs;
    label Source = 'Effect'
        df = 'NumDF'
    Fvalue = 'Value'
    pFddf = 'Pr > F(DDF)'
    pFinf = 'Pr > F(infty)';
var Source df Fvalue pFddf pFinf;
    format pF: pvalue6.;
run;

The PROC GLM code produces \( p \)-values that correspond to the ANOVAF \( p \)-values shown as \( \text{Pr > F(DDF)} \) in the MIXED output. The subsequent DATA step computes the \( p \)-values that correspond to \( \text{Pr > F(infty)} \) in the PROC MIXED output.

**Parameterization of Mixed Models**

Recall that a mixed model is of the form

\[
y = X\beta + Z\gamma + \epsilon
\]

where \( y \) represents univariate data, \( \beta \) is an unknown vector of fixed effects with known model matrix \( X \), \( \gamma \) is an unknown vector of random effects with known model matrix \( Z \), and \( \epsilon \) is an unknown random error vector.

PROC MIXED constructs a mixed model according to the specifications in the MODEL, RANDOM, and REPEATED statements. Each effect in the MODEL statement generates one or more columns in the model matrix \( X \), and each effect in the RANDOM statement generates one or more columns in the model matrix \( Z \). Effects in the REPEATED statement do not generate model matrices; they serve only to index observations within subjects. This section shows precisely how PROC MIXED builds \( X \) and \( Z \).

**Intercept**

By default, all models automatically include a column of 1s in \( X \) to estimate a fixed-effect intercept parameter \( \mu \). You can use the NOINT option in the MODEL statement to suppress this intercept. The NOINT option is useful when you are specifying a classification effect in the MODEL statement and you want the parameter estimate to be in terms of the mean response for each level of that effect, rather than in terms of a deviation from an overall mean.

By contrast, the intercept is not included by default in \( Z \). To obtain a column of 1s in \( Z \), you must specify in the RANDOM statement either the INTERCEPT effect or some effect that has only one level.
Regression Effects

Numeric variables, or polynomial terms involving them, can be included in the model as regression effects (covariates). The actual values of such terms are included as columns of the model matrices $X$ and $Z$. You can use the bar operator with a regression effect to generate polynomial effects. For instance, $X|X|X$ expands to $X \times X \times X \times X \times X$, a cubic model.

Main Effects

If a classification variable has $m$ levels, PROC MIXED generates $m$ columns in the model matrix for its main effect. Each column is an indicator variable for a given level. The order of the columns is the sort order of the values of their levels and can be controlled with the ORDER= option in the PROC MIXED statement. Table 63.20 is an example.

<table>
<thead>
<tr>
<th>Data</th>
<th>I</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>μ</td>
<td>A1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Typically, there are more columns for these effects than there are degrees of freedom for them. In other words, PROC MIXED uses an overparameterized model.

Interaction Effects

Often a model includes interaction (crossed) effects. With an interaction, PROC MIXED first reorders the terms to correspond to the order of the variables in the CLASS statement. Thus, $B^A$ becomes $A^B$ if $A$ precedes $B$ in the CLASS statement. Then, PROC MIXED generates columns for all combinations of levels that occur in the data. The order of the columns is such that the rightmost variables in the cross index faster than the leftmost variables (Table 63.21). Empty columns (that would contain all 0s) are not generated for $X$, but they are for $Z$.

<table>
<thead>
<tr>
<th>Data</th>
<th>I</th>
<th>A</th>
<th>B</th>
<th>A*B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
In the preceding matrix, main-effects columns are not linearly independent of crossed-effects columns; in
fact, the column space for the crossed effects contains the space of the main effect.

When your model contains many interaction effects, you might be able to code them more parsimoniously by
using the bar operator (|). The bar operator generates all possible interaction effects. For example, A|B|C
expands to A B A*B C A*C B*C A*B*C. To eliminate higher-order interaction effects, use the at sign (@) in
conjunction with the bar operator. For instance, A|B|C|D @ 2 expands to A B A*B C A*C B*C D A*D B*D
C*D.

**Nested Effects**

Nested effects are generated in the same manner as crossed effects. Hence, the design columns generated by
the following two statements are the same (but the ordering of the columns is different):

```plaintext
model Y=A B(A);
model Y=A A*B;
```

The nesting operator in PROC MIXED is more a notational convenience than an operation distinct from
crossing. Nested effects are typically characterized by the property that the nested variables never appear as
main effects. The order of the variables within nesting parentheses is made to correspond to the order of these
variables in the CLASS statement. The order of the columns is such that variables outside the parentheses
index faster than those inside the parentheses, and the rightmost nested variables index faster than the leftmost
variables (Table 63.22).

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>A</th>
<th>B(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Note that nested effects are often distinguished from interaction effects by the implied randomization structure
of the design. That is, they usually indicate random effects within a fixed-effects framework. The fact that
random effects can be modeled directly in the RANDOM statement might make the specification of nested
effects in the MODEL statement unnecessary.

**Continuous-Nesting-Class Effects**

When a continuous variable nests with a classification variable, the design columns are constructed by
multiplying the continuous values into the design columns for the class effect (Table 63.23).
Chapter 63: The MIXED Procedure

Table 63.23  Example of Continuous-Nesting-Class Effects

<table>
<thead>
<tr>
<th>Data</th>
<th>I</th>
<th>A</th>
<th>X(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>A</td>
<td>μ</td>
<td>A1</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

This model estimates a separate slope for X within each level of A.

Continuous-by-Class Effects

Continuous-by-class effects generate the same design columns as continuous-nesting-class effects. The two models are made different by the presence of the continuous variable as a regressor by itself, as well as a contributor to a compound effect. Table 63.24 shows an example.

Table 63.24  Example of Continuous-by-Class Effects

<table>
<thead>
<tr>
<th>Data</th>
<th>I</th>
<th>X</th>
<th>A</th>
<th>X*A</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>A</td>
<td>μ</td>
<td>X</td>
<td>A1</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>1</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>1</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>1</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>1</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>1</td>
<td>23</td>
<td>0</td>
</tr>
</tbody>
</table>

You can use continuous-by-class effects to test for homogeneity of slopes.

General Effects

An example that combines all the effects is X1*X2*A*B*C (D E). The continuous list comes first, followed by the crossed list, followed by the nested list in parentheses. You should be aware of the sequencing of parameters when you use the CONTRAST or ESTIMATE statement to compute some function of the parameter estimates.

Effects might be renamed by PROC MIXED to correspond to ordering rules. For example, B*A(E D) might be renamed A*B(D E) to satisfy the following:

- Classification variables that occur outside parentheses (crossed effects) are sorted in the order in which they appear in the CLASS statement.
- Variables within parentheses (nested effects) are sorted in the order in which they appear in the CLASS statement.
The sequencing of the parameters generated by an effect can be described by which variables have their levels indexed faster:

- Variables in the crossed list index faster than variables in the nested list.
- Within a crossed or nested list, variables to the right index faster than variables to the left.

For example, suppose a model includes four effects—A, B, C, and D—each having two levels, 1 and 2. Suppose the CLASS statement is as follows:

```
class A B C D;
```

Then the order of the parameters for the effect B*A(C D), which is renamed A*B (C D), is

\[
A_1 B_1 C_1 D_1 \rightarrow A_1 B_2 C_1 D_1 \rightarrow A_2 B_1 C_1 D_1 \rightarrow A_2 B_2 C_1 D_1
\]

\[
A_1 B_1 C_2 D_1 \rightarrow A_1 B_2 C_2 D_1 \rightarrow A_2 B_1 C_2 D_1 \rightarrow A_2 B_2 C_2 D_1
\]

\[
A_1 B_1 C_2 D_2 \rightarrow A_1 B_2 C_2 D_2 \rightarrow A_2 B_1 C_2 D_2 \rightarrow A_2 B_2 C_2 D_2
\]

Note that first the crossed effects B and A are sorted in the order in which they appear in the CLASS statement so that A precedes B in the parameter list. Then, for each combination of the nested effects in turn, combinations of A and B appear. The B effect moves fastest because it is rightmost in the cross list. Then A moves next fastest, and D moves next fastest. The C effect is the slowest since it is leftmost in the nested list.

When numeric levels are used, levels are sorted by their character format, which might not correspond to their numeric sort sequence (for example, noninteger levels). Therefore, it is advisable to include a desired format for numeric levels or to use the ORDER=INTERNAL option in the PROC MIXED statement to ensure that levels are sorted by their internal values.

### Implications of the Non-Full-Rank Parameterization

For models with fixed effects involving classification variables, there are more design columns in \( X \) constructed than there are degrees of freedom for the effect. Thus, there are linear dependencies among the columns of \( X \). In this event, all of the parameters are not estimable; there is an infinite number of solutions to the mixed model equations. PROC MIXED uses a generalized inverse (a \( g_2 \)-inverse, Pringle and Rayner 1971) to obtain values for the estimates (Searle 1971). The solution values are not displayed unless you specify the SOLUTION option in the MODEL statement. The solution has the characteristic that estimates are 0 whenever the design column for that parameter is a linear combination of previous columns. With this parameterization, hypothesis tests are constructed to test linear functions of the parameters that are estimable.

Some procedures (such as the CATMOD procedure) reparameterize models to full rank by using restrictions on the parameters. PROC GLM and PROC MIXED do not reparameterize, making the hypotheses that are commonly tested more understandable. See Goodnight (1978) for additional reasons for not reparameterizing.

### Missing Level Combinations

PROC MIXED handles missing level combinations of classification variables similarly to the way PROC GLM does. Both procedures delete fixed-effects parameters corresponding to missing levels in order to preserve estimability. However, PROC MIXED does not delete missing level combinations for random-effects parameters because linear combinations of the random-effects parameters are always estimable. These conventions can affect the way you specify your CONTRAST and ESTIMATE coefficients.
Residuals and Influence Diagnostics

Residual Diagnostics

Consider a residual vector of the form $\bar{e} = PY$, where $P$ is a projection matrix, possibly an oblique projector. A typical element $\bar{e}_i$ with variance $\bar{v}_i$ and estimated variance $\hat{v}_i$ is said to be standardized as

$$\frac{\bar{e}_i}{\sqrt{\text{Var}[\bar{e}_i]}} = \frac{\bar{e}_i}{\sqrt{\hat{v}_i}}$$

and studentized as

$$\frac{\bar{e}_i}{\sqrt{\hat{v}_i}}$$

External studentization uses an estimate of $\text{Var}[\bar{e}_i]$ that does not involve the $i$th observation. Externally studentized residuals are often preferred over internally studentized residuals because they have well-known distributional properties in standard linear models for independent data.

Residuals that are scaled by the estimated variance of the response, i.e., $\bar{e}_i / \sqrt{\text{Var}[Y]}$, are referred to as Pearson-type residuals.

Marginal and Conditional Residuals

The marginal and conditional means in the linear mixed model are $E[Y] = X\beta$ and $E[Y|y] = X\beta + Z\gamma$, respectively. Accordingly, the vector $r_m$ of marginal residuals is defined as

$$r_m = Y - X\hat{\beta}$$

and the vector $r_c$ of conditional residuals is

$$r_c = Y - X\hat{\beta} - Z\hat{\gamma} = r_m - Z\hat{\gamma}$$

Following Gregoire, Schabenberger, and Barrett (1995), let $Q = X(X'\hat{V}^{-1}X)^{-1}X'$ and $K = I - ZGZ'\hat{V}^{-1}$. Then

$$\tilde{\text{Var}}[r_m] = \hat{V} - Q$$

$$\tilde{\text{Var}}[r_c] = K(\hat{V} - Q)K'$$

For an individual observation the raw, studentized, and Pearson-type residuals computed by the MIXED procedure are given in Table 63.25.

<table>
<thead>
<tr>
<th>Type of Residual</th>
<th>Marginal</th>
<th>Conditional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>$r_{mi} = Y_i - X_i\hat{\beta}$</td>
<td>$r_{ci} = r_{mi} - Z_i\hat{\gamma}$</td>
</tr>
<tr>
<td>Studentized</td>
<td>$r_{mi}^{\text{student}} = \frac{r_{mi}}{\sqrt{\text{Var}[r_{mi}]}$</td>
<td>$r_{ci}^{\text{student}} = \frac{r_{ci}}{\sqrt{\text{Var}[r_{ci}]}$</td>
</tr>
<tr>
<td>Pearson</td>
<td>$r_{mi}^{\text{Pearson}} = \frac{r_{mi}}{\sqrt{\text{Var}[Y]}}$</td>
<td>$r_{ci}^{\text{Pearson}} = \frac{r_{ci}}{\sqrt{\text{Var}[Y</td>
</tr>
</tbody>
</table>
When the OUTPM= option is specified in addition to the RESIDUAL option in the MODEL statement, $r_{mi}^{\text{student}}$ and $r_{mi}^{\text{pearson}}$ are added to the data set as variables Resid, StudentResid, and PearsonResid, respectively. When the OUTP= option is specified, $r_{ci}^{\text{student}}$ and $r_{ci}^{\text{pearson}}$ are added to the data set. Raw residuals are part of the OUTPM= and OUTP= data sets without the RESIDUAL option.

**Scaled Residuals**

For correlated data, a set of scaled quantities can be defined through the Cholesky decomposition of the variance-covariance matrix. Since fitted residuals in linear models are rank-deficient, it is customary to draw on the variance-covariance matrix of the data. If $\text{Var}[Y] = V$ and $C' C = V$, then $C'^{-1} Y$ has uniform dispersion and its elements are uncorrelated.

Scaled residuals in a mixed model are meaningful for quantities based on the marginal distribution of the data. Let $\tilde{C}$ denote the Cholesky root of $\tilde{V}$, so that $\tilde{C}' \tilde{C} = \tilde{V}$, and define

$$Y_c = \tilde{C}'^{-1} Y$$

$$r_{m(c)} = \tilde{C}'^{-1} r_m$$

By analogy with other scalings, the inverse Cholesky decomposition can also be applied to the residual vector, $\tilde{C}'^{-1} r_m$, although $V$ is not the variance-covariance matrix of $r_m$.

To diagnose whether the covariance structure of the model has been specified correctly can be difficult based on $Y_c$, since the inverse Cholesky transformation affects the expected value of $Y_c$. You can draw on $r_{m(c)}$ as a vector of (approximately) uncorrelated data with constant mean.

When the OUTPM= option in the MODEL statement is specified in addition to the VCIRY option, $Y_c$ is added as variable ScaledDep and $r_{m(c)}$ is added as ScaledResid to the data set.

**Influence Diagnostics**

**Basic Idea and Statistics**

The general idea of quantifying the influence of one or more observations relies on computing parameter estimates based on all data points, removing the cases in question from the data, refitting the model, and computing statistics based on the change between full-data and reduced-data estimation. Influence statistics can be coarsely grouped by the aspect of estimation that is their primary target:

- overall measures compare changes in objective functions: (restricted) likelihood distance (Cook and Weisberg 1982, Ch. 5.2)
- influence on parameter estimates: Cook’s $D$ (Cook 1977, 1979), MDFFITS (Belsley, Kuh, and Welsch 1980, p. 32)
- influence on precision of estimates: CovRatio and CovTrace
- influence on fitted and predicted values: PRESS residual, PRESS statistic (Allen 1974), DFFITS (Belsley, Kuh, and Welsch 1980, p. 15)
- outlier properties: internally and externally studentized residuals, leverage

For linear models for uncorrelated data, it is not necessary to refit the model after removing a data point in order to measure the impact of an observation on the model. The change in fixed effect estimates, residuals, residual sums of squares, and the variance-covariance matrix of the fixed effects can be computed based on
the fit to the full data alone. By contrast, in mixed models several important complications arise. Data points can affect not only the fixed effects but also the covariance parameter estimates on which the fixed-effects estimates depend. Furthermore, closed-form expressions for computing the change in important model quantities might not be available.

This section provides background material for the various influence diagnostics available with the MIXED procedure. See the section “Mixed Models Theory” on page 5313 for relevant expressions and definitions. The parameter vector \( \theta \) denotes all unknown parameters in the \( R \) and \( G \) matrix.

The observations whose influence is being ascertained are represented by the set \( U \) and referred to simply as “the observations in \( U \)”. The estimate of a parameter vector, such as \( \beta \), obtained from all observations except those in the set \( U \) is denoted \( \hat{\beta}(U) \). In case of a matrix \( A \), the notation \( A(U) \) represents the matrix with the rows in \( U \) removed; these rows are collected in \( A_U \). If \( A \) is symmetric, then notation \( A(U) \) implies removal of rows and columns. The vector \( Y_U \) comprises the responses of the data points being removed, and \( V(U) \) is the variance-covariance matrix of the remaining observations. When \( k = 1 \), lowercase notation emphasizes that single points are removed, such as \( A(u) \).

Managing the Covariance Parameters
An important component of influence diagnostics in the mixed model is the estimated variance-covariance matrix \( V = ZGZ' + R \). To make the dependence on the vector of covariance parameters explicit, write it as \( V(\theta) \). If one parameter, \( \sigma^2 \), is profiled or factored out of \( V \), the remaining parameters are denoted as \( \theta^* \). Notice that in a model where \( G \) is diagonal and \( R = \sigma^2I \), the parameter vector \( \theta^* \) contains the ratios of each variance component and \( \sigma^2 \) (see Wolfinger, Tobias, and Sall 1994). When \( \text{ITER}=0 \), two scenarios are distinguished:

1. If the residual variance is not profiled, either because the model does not contain a residual variance or because it is part of the Newton-Raphson iterations, then \( \hat{\theta}(U) = \hat{\theta} \).

2. If the residual variance is profiled, then \( \hat{\theta}^*(U) = \hat{\theta}^* \) and \( \hat{\sigma}^2(U) \neq \hat{\sigma}^2 \). Influence statistics such as Cook’s \( D \) and internally studentized residuals are based on \( V(\hat{\theta}) \), whereas externally studentized residuals and the DFFITS statistic are based on \( V(\hat{\theta}(U)) = \sigma^2(U)V(\hat{\theta}(U)) \). In a random components model with uncorrelated errors, for example, the computation of \( V(\hat{\theta}(U)) \) involves scaling of \( G \) and \( \hat{R} \) by the full-data estimate \( \tilde{\sigma}^2 \) and multiplying the result with the reduced-data estimate \( \hat{\sigma}^2(U) \).

Certain statistics, such as MDFFITS, CovRatio, and CovTrace, require an estimate of the variance of the fixed effects that is based on the reduced number of observations. For example, \( V(\hat{\theta}(U)) \) is evaluated at the reduced-data parameter estimates but computed for the entire data set. The matrix \( V(U)(\hat{\theta}(U)) \), on the other hand, has rows and columns corresponding to the points in \( U \) removed. The resulting matrix is evaluated at the delete-case estimates.

When influence analysis is iterative, the entire vector \( \theta \) is updated, whether the residual variance is profiled or not. The matrices to be distinguished here are \( V(\hat{\theta}) \), \( V(\hat{\theta}(U)) \), and \( V(U)(\hat{\theta}(U)) \), with unambiguous notation.

Predicted Values, PRESS Residual, and PRESS Statistic
An unconditional predicted value is \( \hat{y}_i = x_i'\hat{\beta} \), where the vector \( x_i \) is the \( i \)th row of \( X \). The (raw) residual is given as \( \hat{\varepsilon}_i = y_i - \hat{y}_i \), and the PRESS residual is

\[
\hat{\varepsilon}_{i(U)} = y_i - x_i'\hat{\beta}(U)
\]
The PRESS statistic is the sum of the squared PRESS residuals,

\[ \text{PRESS} = \sum_{i \in U} \hat{e}_i^2(U) \]

where the sum is over the observations in \( U \).

If \( \text{EFFECT=} \), \( \text{SIZE=} \), or \( \text{KEEP=} \) is not specified, PROC MIXED computes the PRESS residual for each observation selected through \( \text{SELECT=} \) (or all observations if \( \text{SELECT=} \) is not given). If \( \text{EFFECT=} \), \( \text{SIZE=} \), or \( \text{KEEP=} \) is specified, the procedure computes PRESS.

**Leverage**

For the general mixed model, leverage can be defined through the projection matrix that results from a transformation of the model with the inverse of the Cholesky decomposition of \( V \), or through an oblique projector. The MIXED procedure follows the latter path in the computation of influence diagnostics. The leverage value reported for the \( i \)th observation is the \( i \)th diagonal entry of the matrix

\[ H = X(X'V(\hat{\theta})^{-1}X)^{-1}X'V(\hat{\theta})^{-1} \]

which is the weight of the observation in contributing to its own predicted value, \( H = d\hat{Y}/dY \).

While \( H \) is idempotent, it is generally not symmetric and thus not a projection matrix in the narrow sense.

The properties of these leverages are generalizations of the properties in models with diagonal variance-covariance matrices. For example, \( \hat{Y} = HY \), and in a model with intercept and \( V = \sigma^2I \), the leverage values

\[ h_{ii} = x'_i(X'X)^{-1}x_i \]

are \( h^1_{ii} = 1/n \leq h_{ii} \leq 1 = h^0_{ii} \) and \( \sum_{i=1}^{n} h_{ii} = \text{rank}(X) \). The lower bound for \( h_{ii} \) is achieved in an intercept-only model, and the upper bound is achieved in a saturated model. The trace of \( H \) equals the rank of \( X \).

If \( v_{ij} \) denotes the element in row \( i \), column \( j \) of \( V^{-1} \), then for a model containing only an intercept the diagonal elements of \( H \) are

\[ h_{ii} = \frac{\sum_{j=1}^{n} v_{ij}}{\sum_{j=1}^{n} \sum_{j=1}^{n} v_{ij}} \]

Because \( \sum_{j=1}^{n} v_{ij} \) is a sum of elements in the \( i \)th row of the inverse variance-covariance matrix, \( h_{ii} \) can be negative, even if the correlations among data points are nonnegative. In case of a saturated model with \( X = I \), \( h_{ii} = 1.0 \).

**Internally andExternally Studentized Residuals**

See the section “Residual Diagnostics” on page 5330 for the distinction between standardization, studentization, and scaling of residuals. Internally studentized marginal and conditional residuals are computed with the RESIDUAL option of the MODEL statement. The INFLUENCE option computes internally and externally studentized marginal residuals.

The computation of internally studentized residuals relies on the diagonal entries of \( V(\hat{\theta}) - Q(\hat{\theta}) \), where \( Q(\hat{\theta}) = X(X'V(\hat{\theta})^{-1}X)^{-1}X' \). Externally studentized residuals require iterative influence analysis or a profiled residual variance. In the former case the studentization is based on \( V(\hat{\theta}_U) \); in the latter case it is based on \( \sigma^2(U) V(\hat{\theta}^*) \).
**Cook’s D**

Cook’s D statistic is an invariant norm that measures the influence of observations in \( U \) on a vector of parameter estimates (Cook 1977). In case of the fixed-effects coefficients, let

\[
\delta(U) = \hat{\beta} - \hat{\beta}(U)
\]

Then the MIXED procedure computes

\[
D(\beta) = \delta'(U)^T \text{Var}[\hat{\beta}]^{-1} \delta(U) / \text{rank}(X)
\]

where \( \text{Var}[\hat{\beta}]^{-1} \) is the matrix that results from sweeping \( (X'V(\hat{\theta}^{-1})X)^{-1} \).

If \( V \) is known, Cook’s D can be calibrated according to a chi-square distribution with degrees of freedom equal to the rank of \( X \) (Christensen, Pearson, and Johnson 1992). For estimated \( V \) the calibration can be carried out according to an \( F(\text{rank}(X), n - \text{rank}(X)) \) distribution. To interpret \( D \) on a familiar scale, Cook (1979) and Cook and Weisberg (1982, p. 116) refer to the 50th percentile of the reference distribution. If \( D \) is equal to that percentile, then removing the points in \( U \) moves the fixed-effects coefficient vector from the center of the confidence region to the 50% confidence ellipsoid (Myers 1990, p. 262).

In the case of iterative influence analysis, the MIXED procedure also computes a \( D \)-type statistic for the covariance parameters. If \( \Gamma \) is the asymptotic variance-covariance matrix of \( \hat{\theta} \), then MIXED computes

\[
D_\theta = (\hat{\theta} - \hat{\theta}(U))^T \Gamma^{-1} (\hat{\theta} - \hat{\theta}(U))
\]

**DFFITS and MDFFITS**

A DFFIT measures the change in predicted values due to removal of data points. If this change is standardized by the externally estimated standard error of the predicted value in the full data, the DFFITS statistic of Belsley, Kuh, and Welsch (1980, p. 15) results:

\[
\text{DFFITS}_i = (\hat{y}_i - \hat{y}_{i(u)}) / \text{ese}(\hat{y}_i)
\]

The MIXED procedure computes DFFITS when the **EFFECT=** or **SIZE=** modifier of the **INFLUENCE** option is not in effect. In general, an external estimate of the estimated standard error is used. When \( \text{ITER} > 0 \), the estimate is

\[
\text{ese}(\hat{y}_i) = \sqrt{x_i'(X'V(\hat{\theta}(u))^{-1}X)^{-1}x_i}
\]

When \( \text{ITER}=0 \) and \( \sigma^2 \) is profiled, then

\[
\text{ese}(\hat{y}_i) = \hat{\sigma}(u) \sqrt{x_i'(X'V(\hat{\theta}^*)^{-1}X)^{-1}x_i}
\]

When the **EFFECT=**, **SIZE=**, or **KEEP=** modifier is specified, the MIXED procedure computes a multivariate version suitable for the deletion of multiple data points. The statistic, termed MDFFITS after the MDDFIT statistic of Belsley, Kuh, and Welsch (1980, p. 32), is closely related to Cook’s D. Consider the case \( V = \sigma^2V(\theta^*) \) so that

\[
\text{Var}[\hat{\beta}] = \sigma^2(X'V(\theta^*)^{-1}X)^{-1}
\]
and let \( \hat{\text{Var}}[\hat{\beta}(U)] \) be an estimate of \( \text{Var}[\hat{\beta}(U)] \) that does not use the observations in \( U \). The MDFFITS statistic is then computed as

\[
\text{MDFFITS}(\beta) = \delta'(U) \hat{\text{Var}}[\hat{\beta}(U)]^{-1} \delta(U) / \text{rank}(X)
\]

If \( \text{ITER} = 0 \) and \( \sigma^2 \) is profiled, then \( \hat{\text{Var}}[\hat{\beta}(U)]^{-1} \) is obtained by sweeping

\[
\hat{\sigma}^2(U) (X'(U) V(U)(\theta)^{-1} X(U))^{-1}
\]

The underlying idea is that if \( \theta^* \) were known, then

\[
(X'(U) V(U)(\theta)^{-1} X(U))^{-1}
\]

would be \( \text{Var}[\beta] / \sigma^2 \) in a generalized least squares regression with all but the data in \( U \).

In the case of iterative influence analysis, \( \hat{\text{Var}}[\hat{\beta}(U)] \) is evaluated at \( \hat{\theta}(U) \). Furthermore, a MDFFITS-type statistic is then computed for the covariance parameters:

\[
\text{MDFFITS}(\theta) = (\hat{\theta} - \hat{\theta}(U)) \hat{\text{Var}}[\hat{\theta}(U)]^{-1} (\hat{\theta} - \hat{\theta}(U))
\]

**Covariance Ratio and Trace**

These statistics depend on the availability of an external estimate of \( V \), or at least of \( \sigma^2 \). Whereas Cook’s \( D \) and MDFFITS measure the impact of data points on a vector of parameter estimates, the covariance-based statistics measure impact on their precision. Following Christensen, Pearson, and Johnson (1992), the MIXED procedure computes

\[
\text{CovTrace}(\beta) = |\text{trace}(\hat{\text{Var}}[\hat{\beta}]^{-1} \hat{\text{Var}}[\hat{\beta}(U)]) - \text{rank}(X)|
\]

\[
\text{CovRatio}(\beta) = \frac{\text{det}_{ns}(\hat{\text{Var}}[\hat{\beta}(U)])}{\text{det}_{ns}(\hat{\text{Var}}[\hat{\beta}])}
\]

where \( \text{det}_{ns}(M) \) denotes the determinant of the nonsingular part of matrix \( M \).

In the case of iterative influence analysis these statistics are also computed for the covariance parameter estimates. If \( q \) denotes the rank of \( \hat{\text{Var}}[\hat{\theta}] \), then

\[
\text{CovTrace}(\theta) = |\text{trace}(\hat{\text{Var}}[\hat{\theta}]^{-1} \hat{\text{Var}}[\hat{\theta}(U)]) - q|
\]

\[
\text{CovRatio}(\theta) = \frac{\text{det}_{ns}(\hat{\text{Var}}[\hat{\theta}(U)])}{\text{det}_{ns}(\hat{\text{Var}}[\hat{\theta}])}
\]

**Likelihood Distances**

The log-likelihood function \( l \) and restricted log-likelihood function \( l_R \) of the linear mixed model are given in the section “Estimating Covariance Parameters in the Mixed Model” on page 5318. Denote as \( \psi \) the collection of all parameters, i.e., the fixed effects \( \beta \) and the covariance parameters \( \theta \). Twice the difference between the (restricted) log-likelihood evaluated at the full-data estimates \( \hat{\psi} \) and at the reduced-data estimates \( \hat{\psi}(U) \) is known as the (restricted) likelihood distance:

\[
\text{RLD}(U) = 2(l_R(\hat{\psi}) - l_R(\hat{\psi}(U)))
\]

\[
\text{LD}(U) = 2(l(\hat{\psi}) - l(\hat{\psi}(U)))
\]
Cook and Weisberg (1982, Ch. 5.2) refer to these differences as *likelihood distances*, Beckman, Nachtsheim, and Cook (1987) call the measures *likelihood displacements*. If the number of elements in $\psi$ that are subject to updating following point removal is $q$, then likelihood displacements can be compared against cutoffs from a chi-square distribution with $q$ degrees of freedom. Notice that this reference distribution does not depend on the number of observations removed from the analysis, but rather on the number of model parameters that are updated. The likelihood displacement gives twice the amount by which the log likelihood of the full data changes if one were to use an estimate based on fewer data points. It is thus a global, summary measure of the influence of the observations in $U$ jointly on all parameters.

Unless METHOD=ML, the MIXED procedure computes the likelihood displacement based on the residual (=restricted) log likelihood, even if METHOD=MIVQUE0 or METHOD=TYPE1, TYPE2, or TYPE3.

**Noniterative Update Formulas**

Update formulas that do not require refitting of the model are available for the cases where $V = \sigma^2 I$, $V$ is known, or $V^*$ is known. When ITER=0 and these update formulas can be invoked, the MIXED procedure uses the computational devices that are outlined in the following paragraphs. It is then assumed that the variance-covariance matrix of the fixed effects has the form $(X'V^{-1}X)^{-}$. When DDFM=KENWARDROGER, this is not the case; the estimated variance-covariance matrix is then inflated to better represent the uncertainty in the estimated covariance parameters. Influence statistics when DDFM=KENWARDROGER should iteratively update the covariance parameters (ITER > 0). The dependence of $V$ on $\theta$ is suppressed in the sequel for brevity.

**Updating the Fixed Effects**

Denote by $U$ the $(n \times k)$ matrix that is assembled from $k$ columns of the identity matrix. Each column of $U$ corresponds to the removal of one data point. The point being targeted by the $i$th column of $U$ corresponds to the row in which a 1 appears. Furthermore, define

$$
\Omega = (X'V^{-1}X)^{-}
$$

$$
Q = X\Omega X'
$$

$$
P = V^{-1}(V - Q)V^{-1}
$$

The change in the fixed-effects estimates following removal of the observations in $U$ is

$$
\hat{\beta} - \hat{\beta}_{(U)} = \Omega X'V^{-1}U(U'PU)^{-1}U'V^{-1}(y - X\hat{\beta})
$$

Using results in Cook and Weisberg (1982, A2) you can further compute

$$
\tilde{\Omega} = (X'(U)V_{(U)}^{-1}X(U))^{-} = \Omega + \Omega X'V^{-1}U(U'PU)^{-1}U'V^{-1}X\Omega
$$

If $X$ is $(n \times p)$ of rank $m < p$, then $\Omega$ is deficient in rank and the MIXED procedure computes needed quantities in $\tilde{\Omega}$ by sweeping (Goodnight 1979). If the rank of the $(k \times k)$ matrix $U'PU$ is less than $k$, the removal of the observations introduces a new singularity, whether $X$ is of full rank or not. The solution vectors $\hat{\beta}$ and $\hat{\beta}_{(U)}$ then do not have the same expected values and should not be compared. When the MIXED procedure encounters this situation, influence diagnostics that depend on the choice of generalized inverse are not computed. The procedure also monitors the singularity criteria when sweeping the rows of $(X'V^{-1}X)^{-}$ and of $(X'(U)V_{(U)}^{-1}X(U))^{-}$. If a new singularity is encountered or a former singularity disappears, no influence statistics are computed.
**Residual Variance**  When $\sigma^2$ is profiled out of the marginal variance-covariance matrix, a closed-form estimate of $\sigma^2$ that is based on only the remaining observations can be computed provided $V^* = V(\hat{\theta}^*)$ is known. Hurtado (1993, Thm. 5.2) shows that

$$(n - q - r)\hat{\sigma}^2_{(U)} = (n - q)\hat{\sigma}^2 - \epsilon'_U (\hat{\sigma}^2 U'PU)^{-1}\epsilon_U$$

and $\epsilon_U = U'V^{-1}(y - X\hat{\beta})$. In the case of maximum likelihood estimation $q = 0$ and for REML estimation $q = \text{rank}(X)$. The constant $r$ equals the rank of $(U'PU)$ for REML estimation and the number of effective observations that are removed if METHOD=ML.

**Likelihood Distances**  For noniterative methods the following computational devices are used to compute (restricted) likelihood distances provided that the residual variance $\sigma^2$ is profiled.

The log likelihood function $l(\hat{\theta})$ evaluated at the full-data and reduced-data estimates can be written as

$$l(\hat{\theta}) = -\frac{n}{2} \log(\hat{\sigma}^2) - \frac{1}{2} \log |V^*| - \frac{1}{2} (y - X\hat{\beta})'V^{-1}(y - X\hat{\beta})/\hat{\sigma}^2 - \frac{n}{2} \log(2\pi)$$

$$l(\hat{\theta}_{(U)}) = -\frac{n}{2} \log(\hat{\sigma}^2_{(U)}) - \frac{1}{2} \log |V^*| - \frac{1}{2} (y - X\hat{\beta}_{(U)})'V^{-1}(y - X\hat{\beta}_{(U)})/\hat{\sigma}^2_{(U)} - \frac{n}{2} \log(2\pi)$$

Notice that $l(\hat{\theta}_{(U)})$ evaluates the log likelihood for $n$ data points at the reduced-data estimates. It is not the log likelihood obtained by fitting the model to the reduced data. The likelihood distance is then

$$LD_{(U)} = n \log \left( \frac{\hat{\sigma}^2_{(U)}}{\hat{\sigma}^2} \right) - n + (y - X\hat{\beta}_{(U)})'V^{-1}(y - X\hat{\beta}_{(U)})/\hat{\sigma}^2_{(U)}$$

Expressions for $RLD_{(U)}$ in noniterative influence analysis are derived along the same lines.

---

**Default Output**

The following sections describe the output PROC MIXED produces by default. This output is organized into various tables, and they are discussed in order of appearance.

**Model Information**

The “Model Information” table describes the model, some of the variables it involves, and the method used in fitting it. It also lists the method (profile, factor, parameter, or none) for handling the residual variance in the model. The profile method concentrates the residual variance out of the optimization problem, whereas the parameter method retains it as a parameter in the optimization. The factor method keeps the residual fixed, and none is displayed when a residual variance is not part of the model.

The “Model Information” table also has a row labeled Fixed Effects SE Method. This row describes the method used to compute the approximate standard errors for the fixed-effects parameter estimates and related functions of them. The two possibilities for this row are Model-Based, which is the default method, and Empirical, which results from using the EMPIRICAL option in the PROC MIXED statement.

The ODS name of the “Model Information” table is ModelInfo.
Class Level Information

The “Class Level Information” table lists the levels of every variable specified in the CLASS statement. You should check this information to make sure the data are correct. You can adjust the order of the CLASS variable levels with the ORDER= option in the PROC MIXED statement. The ODS name of the “Class Level Information” table is ClassLevels.

Dimensions

The “Dimensions” table lists the sizes of relevant matrices. This table can be useful in determining CPU time and memory requirements. The ODS name of the “Dimensions” table is Dimensions.

Number of Observations

The “Number of Observations” table shows the number of observations read from the data set and the number of observations used in fitting the model.

Iteration History

The “Iteration History” table describes the optimization of the residual log likelihood or log likelihood. The function to be minimized (the objective function) is $-2l$ for ML and $-2l_R$ for REML; the column name of the objective function in the “Iteration History” table is “-2 Log Like” for ML and “-2 Res Log Like” for REML. The minimization is performed by using a ridge-stabilized Newton-Raphson algorithm, and the rows of this table describe the iterations that this algorithm takes in order to minimize the objective function.

The Evaluations column of the “Iteration History” table tells how many times the objective function is evaluated during each iteration.

The Criterion column of the “Iteration History” table is, by default, a relative Hessian convergence quantity given by

$$
\frac{g_k^T H_k^{-1} g_k}{|f_k|}
$$

where $f_k$ is the value of the objective function at iteration $k$, $g_k$ is the gradient (first derivative) of $f_k$, and $H_k$ is the Hessian (second derivative) of $f_k$. If $H_k$ is singular, then PROC MIXED uses the following relative quantity:

$$
\frac{g_k^T g_k}{|f_k|}
$$

To prevent the division by $|f_k|$, use the ABSOLUTE option in the PROC MIXED statement. To use a relative function or gradient criterion, use the CONVF or CONVG option, respectively.

The Hessian criterion is considered superior to function and gradient criteria because it measures orthogonality rather than lack of progress (Bates and Watts 1988). Provided the initial estimate is feasible and the maximum number of iterations is not exceeded, the Newton-Raphson algorithm is considered to have converged when the criterion is less than the tolerance specified with the CONVF, CONVG, or CONVH option in the PROC MIXED statement. The default tolerance is $1E-8$. If convergence is not achieved, PROC MIXED displays the estimates of the parameters at the last iteration.
A convergence criterion that is missing indicates that a boundary constraint has been dropped; it is usually not a cause for concern.

If you specify the ITDETAILS option in the PROC MIXED statement, then the covariance parameter estimates at each iteration are included as additional columns in the “Iteration History” table.

The ODS name of the “Iteration History” table is IterHistory.

Convergence Status

The “Convergence Status” table informs about the status of the iterative estimation process at the end of the Newton-Raphson optimization. It appears as a message in the listing, and this message is repeated in the log. The ODS object ConvergenceStatus also contains several nonprinting columns that can be helpful in checking the success of the iterative process, in particular during batch processing or when analyzing BY groups. The Status variable takes on the value 0 for a successful convergence (even if the Hessian matrix might not be positive definite). The values 1 and 2 of the Status variable indicate lack of convergence and infeasible initial parameter values, respectively. The variables pdG and pdH can be used to check whether the G and H (Hessian) matrices are positive definite.

For models that are not fit iteratively, such as models without random effects or when the NOITER option is in effect, the “Convergence Status” is not produced.

Covariance Parameter Estimates

The “Covariance Parameter Estimates” table contains the estimates of the parameters in \( G \) and \( R \) (see the section “Estimating Covariance Parameters in the Mixed Model” on page 5318). Their values are labeled in the table along with Subject and Group information if applicable. The estimates are displayed in the Estimate column and are the results of one of the following estimation methods: REML, ML, MIVQUE0, SSCP, Type1, Type2, or Type3.

If you specify the RATIO option in the PROC MIXED statement, the Ratio column is added to the table listing the ratio of each parameter estimate to that of the residual variance.

Specifying the COVTEST option in the PROC MIXED statement produces the “Std Error,” “Z Value,” and “Pr Z” columns. The “Std Error” column contains the approximate standard errors of the covariance parameter estimates. These are the square roots of the diagonal elements of the observed inverse Fisher information matrix, which equals \( 2H^{-1} \), where \( H \) is the Hessian matrix. The \( H \) matrix consists of the second derivatives of the objective function with respect to the covariance parameters; see Wolfinger, Tobias, and Sall (1994) for formulas. When you use the SCORING= option and PROC MIXED converges without stopping the scoring algorithm, PROC MIXED uses the expected Hessian matrix to compute the covariance matrix instead of the observed Hessian. The observed or expected inverse Fisher information matrix can be viewed as an asymptotic covariance matrix of the estimates.

The “Z Value” column is the estimate divided by its approximate standard error, and the “Pr Z” column is the one- or two-tailed area of the standard Gaussian density outside of the Z-value. The MIXED procedure computes one-sided \( p \)-values for the residual variance and for covariance parameters with a lower bound of 0. The procedure computes two-sided \( p \)-values otherwise. These statistics constitute Wald tests of the covariance parameters, and they are valid only asymptotically.

CAUTION: Wald tests can be unreliable in small samples.

The ODS name of the “Covariance Parameter Estimates” table is CovParms.
Fit Statistics

The “Fit Statistics” table provides some statistics about the estimated mixed model. Expressions for the $-2$ times the log likelihood are provided in the section “Estimating Covariance Parameters in the Mixed Model” on page 5318. If the log likelihood is an extremely large number, then PROC MIXED has deemed the estimated $V$ matrix to be singular. In this case, all subsequent results should be viewed with caution.

In addition, the “Fit Statistics” table lists three information criteria: AIC, AICC, and BIC, all in smaller-is-better form. Expressions for these criteria are described under the IC option.

The ODS name of the “Model Fitting Information” table is FitStatistics.

Null Model Likelihood Ratio Test

If one covariance model is a submodel of another, you can carry out a likelihood ratio test for the significance of the more general model by computing $-2$ times the difference between their log likelihoods. Then compare this statistic to the $\chi^2$ distribution with degrees of freedom equal to the difference in the number of parameters for the two models.

This test is reported in the “Null Model Likelihood Ratio Test” table to determine whether it is necessary to model the covariance structure of the data at all. The “Chi-Square” value is $-2$ times the log likelihood from the null model minus $-2$ times the log likelihood from the fitted model, where the null model is the one with only the fixed effects listed in the MODEL statement and $R = \sigma^2 I$. This statistic has an asymptotic $\chi^2$ distribution with $q - 1$ degrees of freedom, where $q$ is the effective number of covariance parameters (those not estimated to be on a boundary constraint). The “Pr > ChiSq” column contains the upper-tail area from this distribution. This $p$-value can be used to assess the significance of the model fit.

This test is not produced for cases where the null hypothesis lies on the boundary of the parameter space, which is typically for variance component models. This is because the standard asymptotic theory does not apply in this case (Self and Liang 1987, Case 5).

If you specify a PARMS statement, PROC MIXED constructs a likelihood ratio test between the best model from the grid search and the final fitted model and reports the results in the “Parameter Search” table.

The ODS name of the “Null Model Likelihood Ratio Test” table is LRT.

Type 3 Tests of Fixed Effects

The “Type 3 Tests of Fixed Effects” table contains hypothesis tests for the significance of each of the fixed effects—that is, those effects you specify in the MODEL statement. By default, PROC MIXED computes these tests by first constructing a Type 3 $L$ matrix (see Chapter 15, “The Four Types of Estimable Functions”) for each effect. This $L$ matrix is then used to compute the following $F$ statistic:

$$F = \frac{\hat{\beta}'[L(X'\hat{V}^{-1}X)^{-1}L']^{-1}L\hat{\beta}}{r}$$

where $r = \text{rank}(L(X'\hat{V}^{-1}X)^{-1}L')$. A $p$-value for the test is computed as the tail area beyond this statistic from an $F$ distribution with NDF and DDF degrees of freedom. The numerator degrees of freedom (NDF) are the row rank of $L$, and the denominator degrees of freedom are computed by using one of the methods described under the DDFM= option. Small values of the $p$-value (typically less than 0.05 or 0.01) indicate a significant effect.

You can use the HTYPE= option in the MODEL statement to obtain tables of Type 1 (sequential) tests and Type 2 (adjusted) tests in addition to or instead of the table of Type 3 (partial) tests.
You can use the CHISQ option in the MODEL statement to obtain Wald $\chi^2$ tests of the fixed effects. These are carried out by using the numerator of the $F$ statistic and comparing it with the $\chi^2$ distribution with NDF degrees of freedom. It is more liberal than the $F$ test because it effectively assumes infinite denominator degrees of freedom.

The ODS names of the “Type 1 Tests of Fixed Effects” through the “Type 3 Tests of Fixed Effects” tables are Tests1 through Tests3, respectively.

**ODS Table Names**

Each table created by PROC MIXED has a name associated with it, and you must use this name to reference the table when using ODS statements. These names are listed in Table 63.26.

<table>
<thead>
<tr>
<th>Table Name</th>
<th>Description</th>
<th>Required Statement / Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>AccRates</td>
<td>Acceptance rates for posterior sampling</td>
<td>PRIOR</td>
</tr>
<tr>
<td>AsyCorr</td>
<td>Asymptotic correlation matrix of covariance parameters</td>
<td>PROC MIXED ASYCORR</td>
</tr>
<tr>
<td>AsyCov</td>
<td>Asymptotic covariance matrix of covariance parameters</td>
<td>PROC MIXED ASYCOV</td>
</tr>
<tr>
<td>Base</td>
<td>Base densities used for posterior sampling</td>
<td>PRIOR</td>
</tr>
<tr>
<td>Bound</td>
<td>Computed bound for posterior rejection sampling</td>
<td>PRIOR</td>
</tr>
<tr>
<td>CholG</td>
<td>Cholesky root of the estimated G matrix</td>
<td>RANDOM / GC</td>
</tr>
<tr>
<td>CholR</td>
<td>Cholesky root of blocks of the estimated R matrix</td>
<td>REPEATED / RC</td>
</tr>
<tr>
<td>CholV</td>
<td>Cholesky root of blocks of the estimated V matrix</td>
<td>RANDOM / VC</td>
</tr>
<tr>
<td>ClassLevels</td>
<td>Level information from the CLASS statement</td>
<td>Default output</td>
</tr>
<tr>
<td>Coef</td>
<td>$L$ matrix coefficients</td>
<td>E option in MODEL, CONTRAST, ESTIMATE, or LSMEANS</td>
</tr>
<tr>
<td></td>
<td>Results from the CONTRAST statements</td>
<td>CONTRAST</td>
</tr>
<tr>
<td>ConvergenceStatus</td>
<td>Convergence status</td>
<td>Default</td>
</tr>
<tr>
<td>CorrB</td>
<td>Approximate correlation matrix of fixed-effects parameter estimates</td>
<td>MODEL / CORRB</td>
</tr>
<tr>
<td>CovB</td>
<td>Approximate covariance matrix of fixed-effects parameter estimates</td>
<td>MODEL / COVB</td>
</tr>
<tr>
<td>CovParms</td>
<td>Estimated covariance parameters</td>
<td>Default output</td>
</tr>
<tr>
<td>Diffs</td>
<td>Differences of LS-means</td>
<td>LSMEANS / DIFF (or PDIFF)</td>
</tr>
<tr>
<td>Dimensions</td>
<td>Dimensions of the model</td>
<td>Default output</td>
</tr>
</tbody>
</table>
Table 63.26  continued

<table>
<thead>
<tr>
<th>Table Name</th>
<th>Description</th>
<th>Required Statement / Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates</td>
<td>Results from ESTIMATE statements</td>
<td>ESTIMATE</td>
</tr>
<tr>
<td>FitStatistics</td>
<td>Fit statistics</td>
<td>Default</td>
</tr>
<tr>
<td>G</td>
<td>Estimated G matrix</td>
<td>RANDOM / G</td>
</tr>
<tr>
<td>GCorr</td>
<td>Correlation matrix from the estimated G matrix</td>
<td>RANDOM / GCORR</td>
</tr>
<tr>
<td>HLM1</td>
<td>Type 1 Hotelling-Lawley-McKeon tests of fixed effects</td>
<td>MODEL / HTYPE=1 and REPEATED / HLM TYPE=UN</td>
</tr>
<tr>
<td>HLM2</td>
<td>Type 2 Hotelling-Lawley-McKeon tests of fixed effects</td>
<td>MODEL / HTYPE=2 and REPEATED / HLM TYPE=UN</td>
</tr>
<tr>
<td>HLM3</td>
<td>Type 3 Hotelling-Lawley-McKeon tests of fixed effects</td>
<td>REPEATED / HLM TYPE=UN</td>
</tr>
<tr>
<td>HLPS1</td>
<td>Type 1 Hotelling-Lawley-Pillai-Samson tests of fixed effects</td>
<td>MODEL / HTYPE=1 and REPEATED / HLPS TYPE=UN</td>
</tr>
<tr>
<td>HLPS2</td>
<td>Type 2 Hotelling-Lawley-Pillai-Samson tests of fixed effects</td>
<td>MODEL / HTYPE=1 and REPEATED / HLPS TYPE=UN</td>
</tr>
<tr>
<td>HLPS3</td>
<td>Type 3 Hotelling-Lawley-Pillai-Samson tests of fixed effects</td>
<td>REPEATED / HLPS TYPE=UN</td>
</tr>
<tr>
<td>Influence</td>
<td>Influence diagnostics</td>
<td>MODEL / INFLUENCE</td>
</tr>
<tr>
<td>InfoCrit</td>
<td>Information criteria</td>
<td>PROC MIXED IC</td>
</tr>
<tr>
<td>InvCholG</td>
<td>Inverse Cholesky root of the estimated G matrix</td>
<td>RANDOM / GCI</td>
</tr>
<tr>
<td>InvCholR</td>
<td>Inverse Cholesky root of blocks of the estimated R matrix</td>
<td>REPEATED / RCI</td>
</tr>
<tr>
<td>InvCholV</td>
<td>Inverse Cholesky root of blocks of the estimated V matrix</td>
<td>RANDOM / VCI</td>
</tr>
<tr>
<td>InvCovB</td>
<td>Inverse of approximate covariance matrix of fixed-effects parameter estima-</td>
<td>MODEL / COVBI</td>
</tr>
<tr>
<td>InvG</td>
<td>Inverse of the estimated G matrix</td>
<td>RANDOM / GI</td>
</tr>
<tr>
<td>InvR</td>
<td>Inverse of blocks of the estimated R matrix</td>
<td>REPEATED / RI</td>
</tr>
<tr>
<td>InvV</td>
<td>Inverse of blocks of the estimated V matrix</td>
<td>RANDOM / VI</td>
</tr>
<tr>
<td>IterHistory</td>
<td>Iteration history</td>
<td>Default output</td>
</tr>
<tr>
<td>LComponents</td>
<td>Single-degree-of-freedom estimates that correspond to rows of the L matrix</td>
<td>MODEL / LCOMPONENTS</td>
</tr>
<tr>
<td>LRT</td>
<td>Likelihood ratio test</td>
<td>Default output</td>
</tr>
<tr>
<td>LSMeans</td>
<td>LS-means</td>
<td>LSMEANS</td>
</tr>
<tr>
<td>MMEq</td>
<td>Mixed model equations</td>
<td>PROC MIXED MMEQ</td>
</tr>
<tr>
<td>MMEqSol</td>
<td>Mixed model equations solution</td>
<td>PROC MIXED MMEQSOL</td>
</tr>
<tr>
<td>ModelInfo</td>
<td>Model information</td>
<td>Default output</td>
</tr>
<tr>
<td>NObs</td>
<td>Number of observations read and used</td>
<td>Default output</td>
</tr>
</tbody>
</table>
In Table 63.26, “Coef” refers to multiple tables produced by the E, E1, E2, or E3 option in the MODEL statement and the E option in the CONTRAST, ESTIMATE, and LSMEANS statements. You can create one large data set of these tables with a statement similar to the following:

```ods output Coef=c;```

To create separate data sets, use the following statement:

```ods output Coef(match_all)=c;```

Here the resulting data sets are named C, C1, C2, etc. The same principles apply to data sets created from the R, CholR, InvCholR, RCorr, InvR, CholV, InvCholV, VCorr, and InvV tables.

In Table 63.26, the following changes have occurred from SAS 6. The Predicted, PredMeans, and Sample tables from SAS 6 no longer exist and have been replaced by output data sets; see descriptions of the MODEL statement options OUTP= and OUTPM= and the PRIOR statement option OUT= for more details. The ML and REML tables from SAS 6 have been replaced by the IterHistory table. The Tests, HLM, and HLPS tables from SAS 6 have been renamed Tests3, HLM3, and HLPS3, respectively.

Table 63.27 lists the variable names associated with the data sets created when you use the ODS OUTPUT option in conjunction with the preceding tables. In Table 63.27, n is used to denote a generic number that depends on the particular data set and model you select, and it can assume a different value each time it is used (even within the same table). The phrase model specific appears in rows of the affected tables to indicate that columns in these tables depend on the variables you specify in the model.
**CAUTION:** There is a danger of name collisions with the variables in the *model specific* tables in Table 63.27 and variables in your input data set. You should avoid using input variables with the same names as the variables in these tables.

Table 63.27 Variable Names for the ODS Tables Produced in PROC MIXED

<table>
<thead>
<tr>
<th>Table Name</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>AsyCorr</td>
<td>Row, CovParm, CovP1–CovPn</td>
</tr>
<tr>
<td>AsyCov</td>
<td>Row, CovParm, CovP1–CovPn</td>
</tr>
<tr>
<td>Base</td>
<td>Type, Parm1–Parmmn</td>
</tr>
<tr>
<td>Bound</td>
<td>Technique, Converge, Iterations, Evaluations, LogBound, CovP1–CovPn, TCovP1–TCovPn</td>
</tr>
<tr>
<td>CholG</td>
<td><em>Model specific</em>, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Row, Col1–Coln</td>
</tr>
<tr>
<td>CholR</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>CholV</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>ClassLevels</td>
<td>Class, Levels, Values</td>
</tr>
<tr>
<td>Coef</td>
<td><em>Model specific</em>, LMatrix, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Row1–Rown</td>
</tr>
<tr>
<td>Contrasts</td>
<td>Label, NumDF, DenDF, ChiSquare, FValue, ProbChiSq, ProbF</td>
</tr>
<tr>
<td>CorrB</td>
<td><em>Model specific</em>, Effect, Row, Col1–Coln</td>
</tr>
<tr>
<td>CovB</td>
<td><em>Model specific</em>, Effect, Row, Col1–Coln</td>
</tr>
<tr>
<td>CovParms</td>
<td>CovParm, Subject, Group, Estimate, StandardError, ZValue, ProbZ, Alpha, Lower, Upper</td>
</tr>
<tr>
<td>Diffs</td>
<td><em>Model specific</em>, Effect, Margins, ByLevel, AT variables, Diff, StandardError, DF, tValue, Tails, Probt, Adjustment, Adjp, Alpha, Lower, Upper, AdjLow, AdjUpp</td>
</tr>
<tr>
<td>Dimensions</td>
<td>Descr, Value</td>
</tr>
<tr>
<td>Estimates</td>
<td>Label, Estimate, StandardError, DF, tValue, Tails, Probt, Alpha, Lower, Upper</td>
</tr>
<tr>
<td>FitStatistics</td>
<td>Descr, Value</td>
</tr>
<tr>
<td>G</td>
<td><em>Model specific</em>, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Row, Col1–Coln</td>
</tr>
<tr>
<td>GCorr</td>
<td><em>Model specific</em>, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Row, Col1–Coln</td>
</tr>
<tr>
<td>HLM1</td>
<td>Effect, NumDF, DenDF, FValue, ProbF</td>
</tr>
<tr>
<td>HLM2</td>
<td>Effect, NumDF, DenDF, FValue, ProbF</td>
</tr>
<tr>
<td>HLM3</td>
<td>Effect, NumDF, DenDF, FValue, ProbF</td>
</tr>
<tr>
<td>HLPS1</td>
<td>Effect, NumDF, DenDF, FValue, ProbF</td>
</tr>
<tr>
<td>HLPS2</td>
<td>Effect, NumDF, DenDF, FValue, ProbF</td>
</tr>
<tr>
<td>HLPS3</td>
<td>Effect, NumDF, DenDF, FValue, ProbF</td>
</tr>
<tr>
<td>Influence</td>
<td><em>Dependent on option modifiers</em>, Effect, Tuple, Obs1–Obsk, Level, Iter, Index, Predicted, Residual, Leverage, PressRes, PRESS, Student, RMSE, RStudent, CookD, DFFITS, MDFFITS, CovRatio, CovTrace, CookDCP, MDFFITSCP, CovRatioCP, CovTraceCP, LD, RLD, Parm1–Parmnp, CovP1–CovPq, Notes</td>
</tr>
<tr>
<td>InfoCrit</td>
<td>Neg2LogLike,Parms,AIC,AICC,HQIC,BIC,CAIC</td>
</tr>
<tr>
<td>Table Name</td>
<td>Variables</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td>InvCholG</td>
<td><em>Model specific</em>, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Row, Col1–Coln</td>
</tr>
<tr>
<td>InvCholR</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>InvCholV</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>InvCovB</td>
<td><em>Model specific</em>, Effect, Row, Col1–Coln</td>
</tr>
<tr>
<td>InvG</td>
<td><em>Model specific</em>, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Row, Col1–Coln</td>
</tr>
<tr>
<td>InvR</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>InvV</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>IterHistory</td>
<td>CovP1–CovPn, Iteration, Evaluations, M2ResLogLike, M2LogLike, Criterion</td>
</tr>
<tr>
<td>LComponents</td>
<td>Effect, TestType, LIndex, Estimate, StdErr, DF, tValue, Probt</td>
</tr>
<tr>
<td>LRT</td>
<td>DF, ChiSquare, ProbChiSq</td>
</tr>
<tr>
<td>LSMeans</td>
<td><em>Model specific</em>, Effect, Margins, ByLevel, AT variables, Estimate, StandardError, DF, tValue, Probt, Alpha, Lower, Upper, Cov1–Covn, Corr1–Corrn</td>
</tr>
<tr>
<td>MMEq</td>
<td><em>Model specific</em>, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Row, Col1–Coln</td>
</tr>
<tr>
<td>MMEqSol</td>
<td><em>Model specific</em>, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Row, Col1–Coln</td>
</tr>
<tr>
<td>ModelInfo</td>
<td>Descr, Value</td>
</tr>
<tr>
<td>Nobs</td>
<td>Label, N, NObsRead, NObsUsed, SumFreqsRead, SumFreqsUsed</td>
</tr>
<tr>
<td>ParmSearch</td>
<td>CovP1–CovPn, Var, ResLogLike, M2ResLogLike2, LogLike, M2LogLike, LogDetH</td>
</tr>
<tr>
<td>Posterior</td>
<td>Descr, Value</td>
</tr>
<tr>
<td>R</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>RCorr</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>Search</td>
<td>Parm, TCovP1–TCovPn, Posterior</td>
</tr>
<tr>
<td>Slices</td>
<td><em>Model specific</em>, Effect, Margins, ByLevel, AT variables, NumDF, DenDF, FValue, ProbF</td>
</tr>
<tr>
<td>SolutionF</td>
<td><em>Model specific</em>, Effect, Estimate, StandardError, DF, tValue, Probt, Alpha, Lower, Upper</td>
</tr>
<tr>
<td>SolutionR</td>
<td><em>Model specific</em>, Effect, Subject, Sub1–Subn, Group, Group1–Groupn, Estimate, StdErrPred, DF, tValue, Probt, Alpha, Lower, Upper</td>
</tr>
<tr>
<td>Tests1</td>
<td>Effect, NumDF, DenDF, ChiSquare, FValue, ProbChiSq, ProbF</td>
</tr>
<tr>
<td>Tests2</td>
<td>Effect, NumDF, DenDF, ChiSquare, FValue, ProbChiSq, ProbF</td>
</tr>
<tr>
<td>Tests3</td>
<td>Effect, NumDF, DenDF, ChiSquare, FValue, ProbChiSq, ProbF</td>
</tr>
<tr>
<td>Type1</td>
<td>Source, DF, SS, MS, EMS, ErrorTerm, ErrorDF, FValue, ProbF</td>
</tr>
<tr>
<td>Type2</td>
<td>Source, DF, SS, MS, EMS, ErrorTerm, ErrorDF, FValue, ProbF</td>
</tr>
<tr>
<td>Type3</td>
<td>Source, DF, SS, MS, EMS, ErrorTerm, ErrorDF, FValue, ProbF</td>
</tr>
<tr>
<td>Trans</td>
<td>Prior, TCovP, CovP1–CovPn</td>
</tr>
<tr>
<td>V</td>
<td>Index, Row, Col1–Coln</td>
</tr>
<tr>
<td>VCorr</td>
<td>Index, Row, Col1–Coln</td>
</tr>
</tbody>
</table>
Some of the variables listed in Table 63.27 are created only when you specify certain options in the relevant PROC MIXED statements.

### ODS Graphics

Statistical procedures use ODS Graphics to create graphs as part of their output. ODS Graphics is described in detail in Chapter 21, “Statistical Graphics Using ODS.”

Before you create graphs, ODS Graphics must be enabled (for example, by specifying the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section “Enabling and Disabling ODS Graphics” on page 606 in Chapter 21, “Statistical Graphics Using ODS.”

The overall appearance of graphs is controlled by ODS styles. Styles and other aspects of using ODS Graphics are discussed in the section “A Primer on ODS Statistical Graphics” on page 605 in Chapter 21, “Statistical Graphics Using ODS.”

Some graphs are produced by default; other graphs are produced by using statements and options.

### ODS Graph Names

You can reference every graph produced through ODS Graphics with a name. The names of the graphs that PROC MIXED generates are listed in Table 63.28, along with the required statements and options.

<table>
<thead>
<tr>
<th>ODS Graph Name</th>
<th>Plot Description</th>
<th>Statement or Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boxplot</td>
<td>Box plots</td>
<td>PLOTS=BOXPLOT</td>
</tr>
<tr>
<td>CovRatioPlot</td>
<td>CovRatio statistics for fixed effects or covariance parameters</td>
<td>PLOTS=INFLUENCESTATPANEL(UNPACK) and MODEL / INFLUENCE</td>
</tr>
<tr>
<td>CooksDPlot</td>
<td>Cook’s D for fixed effects or covariance parameters</td>
<td>PLOTS=INFLUENCESTATPANEL(UNPACK) and MODEL / INFLUENCE</td>
</tr>
<tr>
<td>DistancePlot</td>
<td>Likelihood or restricted likelihood distance</td>
<td>MODEL / INFLUENCE</td>
</tr>
<tr>
<td>InfluenceEstPlot</td>
<td>Panel of deletion estimates</td>
<td>MODEL / INFLUENCE(EST) or PLOTS=INFLUENCEESTPLOT(UNPACK) and MODEL / INFLUENCE</td>
</tr>
<tr>
<td>InfluenceEstPlot</td>
<td>Parameter estimates after removing observation or sets of observations</td>
<td>PLOTS=INFLUENCEESTPLOT(UNPACK) and MODEL / INFLUENCE</td>
</tr>
<tr>
<td>InfluenceStatPanel</td>
<td>Panel of influence statistics</td>
<td>MODEL / INFLUENCE</td>
</tr>
<tr>
<td>PearsonBoxPlot</td>
<td>Box plot of Pearson residuals</td>
<td>PLOTS=PEARSONPANEL(UNPACK BOX)</td>
</tr>
<tr>
<td>PearsonByPredicted</td>
<td>Pearson residuals vs. predicted</td>
<td>PLOTS=PEARSONPANEL(UNPACK)</td>
</tr>
<tr>
<td>PearsonHistogram</td>
<td>Histogram of Pearson residuals</td>
<td>PLOTS=PEARSONPANEL(UNPACK)</td>
</tr>
<tr>
<td>PearsonPanel</td>
<td>Panel of Pearson residuals</td>
<td>MODEL / RESIDUAL</td>
</tr>
</tbody>
</table>
When ODS Graphics is enabled, the LSMESTIMATE and SLICE statements can produce plots that are associated with their analyses. For information about these plots, see the sections “LSMESTIMATE Statement” on page 476 and “SLICE Statement” on page 505 in Chapter 19, “Shared Concepts and Topics.”

Residual Plots

The MIXED procedure can generate panels of residual diagnostics. Each panel consists of a plot of residuals versus predicted values, a histogram with normal density overlaid, a Q-Q plot, and summary residual and fit statistics (Figure 63.15). The plots are produced even if the OUTP= and OUTPM= options in the MODEL statement are not specified. Residual panels can be generated for marginal and conditional raw, studentized, and Pearson residuals as well as for scaled residuals (see the section “Residual Diagnostics” on page 5330).

Recall the example in the section “Getting Started: MIXED Procedure” on page 5238. The following statements generate several $2 \times 2$ panels of residual graphs:

```
ods graphics on;
proc mixed data=heights plots=studentpanel(marginal conditional);
   class Family Gender;
   model Height = Gender / residual;
```
The graphs are created when ODS Graphics is enabled. The panel of the studentized marginal residuals is shown in Figure 63.15, and the panel of the studentized conditional residuals is shown in Figure 63.16.

**Figure 63.15** Panel of the Studentized (Marginal) Residuals

Since the fixed-effects part of the model comprises only an intercept and the gender effect, the marginal mean takes on only two values, one for each gender. The “Residual Statistics” inset in the lower-right corner provides descriptive statistics for the set of residuals that is displayed. Note that residuals in a mixed model do not necessarily sum to zero, even if the model contains an intercept.
**Figure 63.16** Panel of the Conditional Studentized Residuals

Influence Plots

The graphical features of the MIXED procedure enable you to generate plots of influence diagnostics and of deletion estimates. The type and number of plots produced depend on your modifiers of the INFLUENCE option in the MODEL statement and on the PLOTS= option in the PROC MIXED statement. Plots related to covariance parameters are produced only when diagnostics are computed by iterative methods (ITER=). The estimates of the fixed effects—and covariance parameters when updates are iterative—are plotted when you specify the ESTIMATES modifier or when you request PLOTS=INFLUENCEESTPLOT.
Two basic types of influence panels are shown in Figure 63.17 and Figure 63.18. The diagnostics panel shows Cook’s $D$ and CovRatio statistics for the fixed effects and the covariance parameters. For the SAS statements that produce these influence panels, see Example 63.8. In this example, the impact of subjects (Person) on the analysis is assessed. The Cook’s $D$ statistic measures a subject’s impact on the estimates, and the CovRatio statistic measures a subject’s impact on the precision of the estimates. Separate statistics are computed for the fixed effects and the covariance parameters. The CovRatio statistic has a threshold of 1.0. Values larger than 1.0 indicate that precision of the estimates is lost by exclusion of the observations in question. Values smaller than 1.0 indicate that precision is gained by exclusion of the observations from the analysis. For example, it is evident from Output 63.17 that person 20 has considerable impact on the covariance parameter estimates and moderate influence on the fixed-effects estimates. Furthermore, exclusion of this subject from the analysis increases the precision of the covariance parameters, whereas the effect on the precision of the fixed effects is minor.

Output 63.18 shows another type of influence plot, a panel of the deletion estimates. Each plot within the panel corresponds to one of the model parameters. A reference line is drawn at the estimate based on the full data.

**Figure 63.17** Influence Diagnostics
**Computational Issues**

**Computational Method**

In addition to numerous matrix-multiplication routines, PROC MIXED frequently uses the sweep operator (Goodnight 1979) and the Cholesky root (Golub and Van Loan 1989). The routines perform a modified W transformation (Goodnight and Hemmerle 1979) for G-side likelihood calculations and a direct method for R-side likelihood calculations. For the Type 3 $F$ tests, PROC MIXED uses the algorithm described in Chapter 44, “The GLM Procedure.”

PROC MIXED uses a ridge-stabilized Newton-Raphson algorithm to optimize either a full (ML) or residual (REML) likelihood function. The Newton-Raphson algorithm is preferred to the EM algorithm (Lindstrom and Bates 1988). PROC MIXED profiles the likelihood with respect to the fixed effects and also with respect to the residual variance whenever it appears reasonable to do so. The residual profiling can be avoided by using the NOPROFILE option of the PROC MIXED statement. PROC MIXED uses the MIVQUE0 method (Rao 1972; Giesbrecht 1989) to compute initial values.
The likelihoods that PROC MIXED optimizes are usually well-defined continuous functions with a single optimum. The Newton-Raphson algorithm typically performs well and finds the optimum in a few iterations. It is a quadratically converging algorithm, meaning that the error of the approximation near the optimum is squared at each iteration. The quadratic convergence property is evident when the convergence criterion drops to zero by factors of 10 or more.

Using the notation from Table 63.29, the following are estimates of the computational speed of the algorithms used in PROC MIXED. For likelihood calculations, the crossproducts matrix construction is of order $N(p + g)^2$ and the sweep operations are of order $(p + g)^3$. The first derivative calculations for parameters in $G$ are of order $qg^3$ for ML and $q(g^3 + pg^2 + p^2g)$ for REML. If you specify a subject effect in the RANDOM statement and if you are not using the REPEATED statement, then replace $g$ with $g/S$ and $q$ with $qS$ in these calculations. The first derivative calculations for parameters in $R$ are of order $qS(t^3 + gt^2 + g^2t)$ for ML and $qS(t^3 + (p + g)t^2 + (p^2 + g^2)t)$ for REML. For the second derivatives, replace $q$ with $q(q + 1)/2$ in the first derivative expressions. When you specify both $G$- and $R$-side parameters (that is, when you use both the RANDOM and REPEATED statements), then additional calculations are required of an order equal to the sum of the orders for $G$ and $R$. Considerable execution times can result in this case.

For further details about the computational techniques used in PROC MIXED, see Wolfinger, Tobias, and Sall (1994).

### Parameter Constraints

By default, some covariance parameters are assumed to satisfy certain boundary constraints during the Newton-Raphson algorithm. For example, variance components are constrained to be nonnegative, and autoregressive parameters are constrained to be between –1 and 1. You can remove these constraints with the NOBOUND option in the PARMS statement (or with the NOBOUND option in the PROC MIXED statement), but this can lead to estimates that produce an infinite likelihood. You can also introduce or change boundary constraints with the LOWERB= and UPPERB= options in the PARMS statement.

During the Newton-Raphson algorithm, a parameter might be set equal to one of its boundary constraints for a few iterations and then it might move away from the boundary. You see a missing value in the Criterion column of the “Iteration History” table whenever a boundary constraint is dropped.

For some data sets the final estimate of a parameter might equal one of its boundary constraints. This is usually not a cause for concern, but it might lead you to consider a different model. For instance, a variance component estimate can equal zero; in this case, you might want to drop the corresponding random effect from the model. However, be aware that changing the model in this fashion can affect degrees-of-freedom calculations.
Convergence Problems

For some data sets, the Newton-Raphson algorithm can fail to converge. Nonconvergence can result from a number of causes, including flat or ridged likelihood surfaces and ill-conditioned data.

It is also possible for PROC MIXED to converge to a point that is not the global optimum of the likelihood, although this usually occurs only with the spatial covariance structures.

If you experience convergence problems, the following points might be helpful:

- One useful tool is the PARMS statement, which lets you input initial values for the covariance parameters and performs a grid search over the likelihood surface.

- Sometimes the Newton-Raphson algorithm does not perform well when two of the covariance parameters are on a different scale—that is, when they are several orders of magnitude apart. This is because the Hessian matrix is processed jointly for the two parameters, and elements of it corresponding to one of the parameters can become close to internal tolerances in PROC MIXED. In this case, you can improve stability by rescaling the effects in the model so that the covariance parameters are on the same scale.

- Data that are extremely large or extremely small can adversely affect results because of the internal tolerances in PROC MIXED. Rescaling it can improve stability.

- For stubborn problems, you might want to specify ODS OUTPUT COVPARMS=data-set-name to output the “Covariance Parameter Estimates” table as a precautionary measure. That way, if the problem does not converge, you can read the final parameter values back into a new run with the PARMSDATA= option in the PARMS statement.

- Fisher scoring can be more robust than Newton-Raphson with poor MIVQUE(0) starting values. Specifying a SCORING= value of 5 or so might help to recover from poor starting values.

- Tuning the singularity options SINGULAR=, SINGCHOL=, and SINGRES= in the MODEL statement can improve the stability of the optimization process.

- Tuning the MAXITER= and MAXFUNC= options in the PROC MIXED statement can save resources. Also, the ITDETAILS option displays the values of all the parameters at each iteration.

- Using the NOPROFILE and NOBOUND options in the PROC MIXED statement might help convergence, although they can produce unusual results.

- Although the CONVH convergence criterion usually gives the best results, you might want to try CONVF or CONVG, possibly along with the ABSOLUTE option.

- If the convergence criterion reaches a relatively small value such as 1E–7 but never gets lower than 1E–8, you might want to specify CONVH=1E–6 in the PROC MIXED statement to get results; however, interpret the results with caution.

- An infinite likelihood during the iteration process means that the Newton-Raphson algorithm has stepped into a region where either the R or V matrix is nonpositive definite. This is usually no cause for concern as long as iterations continue. If PROC MIXED stops because of an infinite likelihood, recheck your model to make sure that no observations from the same subject are producing identical rows in R or V and that you have enough data to estimate the particular covariance structure you have selected. Any time that the final estimated likelihood is infinite, subsequent results should be interpreted with caution.
• A nonpositive definite Hessian matrix can indicate a surface saddlepoint or linear dependencies among the parameters.

• A warning message about the singularities of $X$ changing indicates that there is some linear dependency in the estimate of $X'V^{-1}X$ that is not found in $X'X$. This can adversely affect the likelihood calculations and optimization process. If you encounter this problem, make sure that your model specification is reasonable and that you have enough data to estimate the particular covariance structure you have selected. Rearranging effects in the \texttt{MODEL} statement so that the most significant ones are first can help, because \texttt{PROC MIXED} sweeps the estimate of $X'V^{-1}X$ in the order of the \texttt{MODEL} effects and the sweep is more stable if larger pivots are dealt with first. If this does not help, specifying starting values with the \texttt{PARMS} statement can place the optimization on a different and possibly more stable path.

• Lack of convergence can indicate model misspecification or a violation of the normality assumption.

**Memory**

Let $p$ be the number of columns in $X$, and let $g$ be the number of columns in $Z$. For large models, most of the memory resources are required for holding symmetric matrices of order $p$, $g$, and $p + g$. The approximate memory requirement in bytes is

\[
40(p^2 + g^2) + 32(p + g)^2
\]

If you have a large model that exceeds the memory capacity of your computer, see the suggestions listed under “Computing Time.”

**Computing Time**

\texttt{PROC MIXED} is computationally intensive, and execution times can be long. In addition to the CPU time used in collecting sums and crossproducts and in solving the mixed model equations (as in \texttt{PROC GLM}), considerable CPU time is often required to compute the likelihood function and its derivatives. These latter computations are performed for every Newton-Raphson iteration.

If you have a model that takes too long to run, the following suggestions can be helpful:

• Examine the “Model Information” table to find out the number of columns in the $X$ and $Z$ matrices. A large number of columns in either matrix can greatly increase computing time. You might want to eliminate some higher-order effects if they are too large.

• If you have a $Z$ matrix with a lot of columns, use the \texttt{DDFM=BW} option in the \texttt{MODEL} statement to eliminate the time required for the containment method.

• If possible, “factor out” a common effect from the effects in the \texttt{RANDOM} statement and make it the \texttt{SUBJECT=} effect. This creates a block-diagonal $G$ matrix and can often speed calculations.

• If possible, use the same or nested \texttt{SUBJECT=} effects in all \texttt{RANDOM} and \texttt{REPEATED} statements.

• If your data set is very large, you might want to analyze it in pieces. The \texttt{BY} statement can help implement this strategy.
• In general, specify random effects with a lot of levels in the REPEATED statement and those with a few levels in the RANDOM statement.

• The METHOD=MIVQUE0 option runs faster than either the METHOD=REML or METHOD=ML option because it is noniterative.

• You can specify known values for the covariance parameters by using the HOLD= or NOITER option in the PARMS statement or the GDATA= option in the RANDOM statement. This eliminates the need for iteration.

• The LOGNOTE option in the PROC MIXED statement writes periodic messages to the SAS log concerning the status of the calculations. It can help you diagnose where the slowdown is occurring.

Examples: MIXED Procedure

The following are basic examples of the use of PROC MIXED. More examples and details can be found in Littell et al. (2006); Wolfinger (1997); Verbeke and Molenberghs (1997, 2000); Murray (1998); Singer (1998); Sullivan, Dukes, and Losina (1999), and Brown and Prescott (1999).

Example 63.1: Split-Plot Design

PROC MIXED can fit a variety of mixed models. One of the most common mixed models is the split-plot design. The split-plot design involves two experimental factors, A and B. Levels of A are randomly assigned to whole plots (main plots), and levels of B are randomly assigned to split plots (subplots) within each whole plot. The design provides more precise information about B than about A, and it often arises when A can be applied only to large experimental units. An example is where A represents irrigation levels for large plots of land and B represents different crop varieties planted in each large plot.

Consider the following data from Stroup (1989a), which arise from a balanced split-plot design with the whole plots arranged in a randomized complete-block design. The variable A is the whole-plot factor, and the variable B is the subplot factor. A traditional analysis of these data involves the construction of the whole-plot error (A*Block) to test A and the pooled residual error (B*Block and A*B*Block) to test B and A*B. To carry out this analysis with PROC GLM, you must use a TEST statement to obtain the correct F test for A.

Performing a mixed model analysis with PROC MIXED eliminates the need for the error term construction. PROC MIXED estimates variance components for Block, A*Block, and the residual, and it automatically incorporates the correct error terms into test statistics.

The following statements create a DATA set for a split-plot design with four blocks, three whole-plot levels, and two subplot levels:

```plaintext
data sp;
  input Block A B Y @@;
datalines;
1 1 1 56 1 1 2 41
1 2 1 50 1 2 2 36
1 3 1 39 1 3 2 35
```
The following statements fit the split-plot model assuming random block effects:

```sas
proc mixed;
  class A B Block;
  model Y = A B A*B;
  random Block A*Block;
run;
```

The variables A, B, and Block are listed as classification variables in the `CLASS` statement. The columns of model matrix \( X \) consist of indicator variables corresponding to the levels of the fixed effects A, B, and A*B listed on the right side of the `MODEL` statement. The dependent variable Y is listed on the left side of the `MODEL` statement.

The columns of the model matrix \( Z \) consist of indicator variables corresponding to the levels of the random effects Block and A*Block. The \( G \) matrix is diagonal and contains the variance components of Block and A*Block. The \( R \) matrix is also diagonal and contains the residual variance.

The SAS statements produce Output 63.1.1–Output 63.1.8.

The “Model Information” table in Output 63.1.1 lists basic information about the split-plot model. REML is used to estimate the variance components, and the residual variance is profiled from the optimization.

**Output 63.1.1** Results for Split-Plot Analysis

<table>
<thead>
<tr>
<th>The Mixed Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model Information</strong></td>
</tr>
<tr>
<td>Data Set</td>
</tr>
<tr>
<td>Dependent Variable</td>
</tr>
<tr>
<td>Covariance Structure</td>
</tr>
<tr>
<td>Estimation Method</td>
</tr>
<tr>
<td>Residual Variance Method</td>
</tr>
<tr>
<td>Fixed Effects SE Method</td>
</tr>
<tr>
<td>Degrees of Freedom Method</td>
</tr>
</tbody>
</table>

The “Class Level Information” table in Output 63.1.2 lists the levels of all variables specified in the `CLASS` statement. You can check this table to make sure that the data are correct.
Example 63.1: Split-Plot Design

**Output 63.1.2** Split-Plot Example (continued)

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>1 2 3</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>1 2</td>
</tr>
<tr>
<td>Block</td>
<td>4</td>
<td>1 2 3 4</td>
</tr>
</tbody>
</table>

The “Dimensions” table in **Output 63.1.3** lists the magnitudes of various vectors and matrices. The $X$ matrix is seen to be $24 \times 12$, and the $Z$ matrix is $24 \times 16$.

**Output 63.1.3** Split-Plot Example (continued)

<table>
<thead>
<tr>
<th>Dimensions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariance Parameters</td>
<td>3</td>
</tr>
<tr>
<td>Columns in $X$</td>
<td>12</td>
</tr>
<tr>
<td>Columns in $Z$</td>
<td>16</td>
</tr>
<tr>
<td>Subjects</td>
<td>1</td>
</tr>
<tr>
<td>Max Obs per Subject</td>
<td>24</td>
</tr>
</tbody>
</table>

The “Number of Observations” table in **Output 63.1.4** shows that all observations read from the data set are used in the analysis.

**Output 63.1.4** Split-Plot Example (continued)

<table>
<thead>
<tr>
<th>Number of Observations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations Read</td>
<td>24</td>
</tr>
<tr>
<td>Number of Observations Used</td>
<td>24</td>
</tr>
<tr>
<td>Number of Observations Not Used</td>
<td>0</td>
</tr>
</tbody>
</table>

PROC MIXED estimates the variance components for Block, A*Block, and the residual by REML. The REML estimates are the values that maximize the likelihood of a set of linearly independent error contrasts, and they provide a correction for the downward bias found in the usual maximum likelihood estimates. The objective function is $-2$ times the logarithm of the restricted likelihood, and PROC MIXED minimizes this objective function to obtain the estimates.

The minimization method is the Newton-Raphson algorithm, which uses the first and second derivatives of the objective function to iteratively find its minimum. The “Iteration History” table in **Output 63.1.5** records the steps of that optimization process. For this example, only one iteration is required to obtain the estimates. The Evaluations column reveals that the restricted likelihood is evaluated once for each of the iterations. A criterion of 0 indicates that the Newton-Raphson algorithm has converged.
Output 63.1.5  Split-Plot Analysis (continued)

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Evaluations</th>
<th>-2 Res Log Like</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>139.81461222</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>119.76184570</td>
<td>0.00000000</td>
</tr>
</tbody>
</table>

Convergence criteria met.

The REML estimates for the variance components of Block, A*Block, and the residual are 62.40, 15.38, and 9.36, respectively, as listed in the Estimate column of the “Covariance Parameter Estimates” table in Output 63.1.6.

Output 63.1.6  Split-Plot Analysis (continued)

<table>
<thead>
<tr>
<th>Covariance Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cov Parm</td>
</tr>
<tr>
<td>Block</td>
</tr>
<tr>
<td>A*Block</td>
</tr>
<tr>
<td>Residual</td>
</tr>
</tbody>
</table>

The “Fit Statistics” table in Output 63.1.7 lists several pieces of information about the fitted mixed model, including the residual log likelihood. The Akaike (AIC) and Bayesian (BIC) information criteria can be used to compare different models; the ones with smaller values are preferred. The AICC information criteria is a small-sample bias-adjusted form of the Akaike criterion (Hurvich and Tsai 1989).

Output 63.1.7  Split-Plot Analysis (continued)

<table>
<thead>
<tr>
<th>Fit Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 Res Log Likelihood</td>
</tr>
<tr>
<td>AIC (Smaller is Better)</td>
</tr>
<tr>
<td>AICC (Smaller is Better)</td>
</tr>
<tr>
<td>BIC (Smaller is Better)</td>
</tr>
</tbody>
</table>

Finally, the fixed effects are tested by using Type 3 estimable functions (Output 63.1.8).
Example 63.1: Split-Plot Design

Output 63.1.8  Split-Plot Analysis (continued)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>6</td>
<td>4.07</td>
<td>0.0764</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>9</td>
<td>19.39</td>
<td>0.0017</td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>9</td>
<td>4.02</td>
<td>0.0566</td>
</tr>
</tbody>
</table>

The tests match the one obtained from the following PROC GLM statements:

```plaintext
proc glm data=sp;
  class A B Block;
  model Y = A B A*B Block A*Block;
  test h=A e=A*Block;
run;
```

You can continue this analysis by producing solutions for the fixed and random effects and then testing various linear combinations of them by using the CONTRAST and ESTIMATE statements. If you use the same CONTRAST and ESTIMATE statements with PROC GLM, the test statistics correspond to the fixed-effects-only model. The test statistics from PROC MIXED incorporate the random effects.

The various “inference space” contrasts given by Stroup (1989a) can be implemented via the ESTIMATE statement. Consider the following examples:

```plaintext
proc mixed data=sp;
  class A B Block;
  model Y = A B A*B;
  random Block A*Block;
  estimate 'a1 mean narrow'
    intercept 1 A 1 B .5 .5 A*B .5 .5 |
               Block .25 .25 .25 .25
           A*Block .25 .25 .25 .25 0 0 0 0 0 0 0 0 0;
  estimate 'a1 mean intermed'
    intercept 1 A 1 B .5 .5 A*B .5 .5 |
               Block .25 .25 .25 .25 0 0 0 0 0 0 0 0 0;
  estimate 'a1 mean broad'
    intercept 1 a 1 b .5 .5 A*B .5 .5;
run;
```

These statements result in Output 63.1.9.
**Output 63.1.9** Inference Space Results

| Label            | Estimate | Error  | DF | t Value | Pr > |t| |
|------------------|----------|--------|----|---------|-------|---|
| a1 mean narrow   | 32.8750  | 1.0817 | 9  | 30.39   | <.0001|
| a1 mean intermed| 32.8750  | 2.2396 | 9  | 14.68   | <.0001|
| a1 mean broad    | 32.8750  | 4.5403 | 9  | 7.24    | <.0001|

Note that all the estimates are equal, but their standard errors increase with the size of the inference space. The narrow inference space consists of the observed levels of Block and A*Block, and the \(t\)-statistic value of 30.39 applies only to these levels. This is the same \(t\) statistic computed by PROC GLM, because it computes standard errors from the narrow inference space. The intermediate inference space consists of the observed levels of Block and the entire population of levels from which A*Block are sampled. The \(t\)-statistic value of 14.68 applies to this intermediate space. The broad inference space consists of arbitrary random levels of both Block and A*Block, and the \(t\)-statistic value of 7.24 is appropriate. Note that the larger the inference space, the weaker the conclusion. However, the broad inference space is usually the one of interest, and even in this space conclusive results are common. The highly significant \(p\)-value for 'a1 mean broad' is an example. You can also obtain the 'a1 mean broad' result by specifying A in an LSMEANS statement. For more discussion of the inference space concept, see McLean, Sanders, and Stroup (1991).

The following statements illustrate another feature of the RANDOM statement. Recall that the basic statements for a split-plot design with whole plots arranged in randomized blocks are as follows.

```plaintext
proc mixed;
  class A B Block;
  model Y = A B A*B;
  random Block A*Block;
run;
```

An equivalent way of specifying this model is as follows:

```plaintext
/* equivalent model */
proc mixed data=sp;
  class A B Block;
  model Y = A B A*B;
  random intercept A / subject=Block;
run;
```

In general, if all of the effects in the RANDOM statement can be nested within one effect, you can specify that one effect by using the SUBJECT= option. The subject effect is, in a sense, “factored out” of the random effects. The specification that uses the SUBJECT= effect can result in faster execution times for large problems because PROC MIXED is able to perform the likelihood calculations separately for each subject.
Example 63.2: Repeated Measures

The following data are from Pothoff and Roy (1964) and consist of growth measurements for 11 girls and 16 boys at ages 8, 10, 12, and 14. Some of the observations are suspect (for example, the third observation for person 20); however, all of the data are used here for comparison purposes.

The analysis strategy employs a linear growth curve model for the boys and girls as well as a variance-covariance model that incorporates correlations for all of the observations arising from the same person. The data are assumed to be Gaussian, and their likelihood is maximized to estimate the model parameters. For overviews of this approach to repeated measures, see Jennrich and Schluchter (1986); Louis (1988); Crowder and Hand (1990); Diggle, Liang, and Zeger (1994); Everitt (1995). Jennrich and Schluchter present results for the Pothoff and Roy data from various covariance structures. The PROC MIXED statements to fit an unstructured variance matrix (their Model 2) are as follows:

data pr;
  input Person Gender $ y1 y2 y3 y4;
  y=y1; Age=8; output;
  y=y2; Age=10; output;
  y=y3; Age=12; output;
  y=y4; Age=14; output;
drop y1-y4;
datalines;
1  F  21.0  20.0  21.5  23.0
2  F  21.0  21.5  24.0  25.5
3  F  20.5  24.0  24.5  26.0
4  F  23.5  24.5  25.0  26.5
5  F  21.5  23.0  22.5  23.5
6  F  20.0  21.0  21.0  22.5
7  F  21.5  22.5  23.0  25.0
8  F  23.0  23.0  23.5  24.0
9  F  20.0  21.0  22.0  21.5
10 F  16.5  19.0  19.0  19.5
11 F  24.5  25.0  28.0  28.0
12 M  26.0  25.0  29.0  31.0
13 M  21.5  22.5  23.0  26.5
14 M  23.0  22.5  24.0  27.5
15 M  25.5  27.5  26.5  27.0
16 M  20.0  23.5  22.5  26.0
17 M  24.5  25.5  27.0  28.5
18 M  22.0  22.0  24.5  26.5
19 M  24.0  21.5  24.5  25.5
20 M  23.0  20.5  31.0  26.0
21 M  27.5  28.0  31.0  31.5
22 M  23.0  23.0  23.5  25.0
23 M  21.5  23.5  24.0  28.0
24 M  17.0  24.5  26.0  29.5
25 M  22.5  25.5  25.5  26.0
26 M  23.0  24.5  26.0  30.0
27 M  22.0  21.5  23.5  25.0
;
To follow Jennrich and Schluchter, this example uses maximum likelihood (METHOD=ML) instead of the default REML to estimate the unknown covariance parameters. The COVTEST option requests asymptotic tests of all the covariance parameters.

The MODEL statement first lists the dependent variable Y. The fixed effects are then listed after the equal sign. The variable Gender requests a different intercept for the girls and boys, Age models an overall linear growth trend, and Gender*Age makes the slopes different over time. It is actually not necessary to specify Age separately, but doing so enables PROC MIXED to carry out a test for heterogeneous slopes. The SOLUTION option requests the display of the fixed-effects solution vector.

The REPEATED statement contains no effects, taking advantage of the default assumption that the observations are ordered similarly for each subject. The TYPE=UN option requests an unstructured block for each SUBJECT=Person. The R matrix is, therefore, block diagonal with 27 blocks, each block consisting of identical 4x4 unstructured matrices. The 10 parameters of these unstructured blocks make up the covariance parameters estimated by maximum likelihood. The R option requests that the first block of R be displayed.

The results from this analysis are shown in Output 63.2.1–Output 63.2.9.

### Output 63.2.1 Repeated Measures Analysis with Unstructured Covariance Matrix

<table>
<thead>
<tr>
<th>The Mixed Procedure</th>
<th>Model Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Set</td>
<td>WORK.PR</td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>y</td>
</tr>
<tr>
<td>Covariance Structure</td>
<td>Unstructured</td>
</tr>
<tr>
<td>Subject Effect</td>
<td>Person</td>
</tr>
<tr>
<td>Estimation Method</td>
<td>ML</td>
</tr>
<tr>
<td>Residual Variance Method</td>
<td>None</td>
</tr>
<tr>
<td>Fixed Effects SE Method</td>
<td>Model-Based</td>
</tr>
<tr>
<td>Degrees of Freedom Method</td>
<td>Between-Within</td>
</tr>
</tbody>
</table>

In Output 63.2.1, the covariance structure is listed as “Unstructured,” and no residual variance is used with this structure. The default degrees-of-freedom method here is “Between-Within.”
**Output 63.2.2** Repeated Measures Analysis (continued)

<table>
<thead>
<tr>
<th>Class Level Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
</tr>
<tr>
<td>Person</td>
</tr>
<tr>
<td>Gender</td>
</tr>
</tbody>
</table>

In **Output 63.2.2**, note that **Person** has 27 levels and **Gender** has 2.

**Output 63.2.3** Repeated Measures Analysis (continued)

<table>
<thead>
<tr>
<th>Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariance Parameters</td>
</tr>
<tr>
<td>Columns in X</td>
</tr>
<tr>
<td>Columns in Z</td>
</tr>
<tr>
<td>Subjects</td>
</tr>
<tr>
<td>Max Obs per Subject</td>
</tr>
</tbody>
</table>

In **Output 63.2.3**, the 10 covariance parameters result from the $4 \times 4$ unstructured blocks of $R$. There is no $Z$ matrix for this model, and each of the 27 subjects has a maximum of 4 observations.

**Output 63.2.4** Repeated Measures Analysis (continued)

<table>
<thead>
<tr>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations Read</td>
</tr>
<tr>
<td>Number of Observations Used</td>
</tr>
<tr>
<td>Number of Observations Not Used</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Iteration History</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iteration</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

Convergence criteria met.

Three Newton-Raphson iterations are required to find the maximum likelihood estimates (**Output 63.2.4**). The default relative Hessian criterion has a final value less than 1E–8, indicating the convergence of the Newton-Raphson algorithm and the attainment of an optimum.
Output 63.2.5 Repeated Measures Analysis (continued)

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.1192</td>
<td>2.4409</td>
<td>3.6105</td>
<td>2.5222</td>
</tr>
<tr>
<td>2</td>
<td>2.4409</td>
<td>3.9279</td>
<td>2.7175</td>
<td>3.0624</td>
</tr>
<tr>
<td>3</td>
<td>3.6105</td>
<td>2.7175</td>
<td>5.9798</td>
<td>3.8235</td>
</tr>
<tr>
<td>4</td>
<td>2.5222</td>
<td>3.0624</td>
<td>3.8235</td>
<td>4.6180</td>
</tr>
</tbody>
</table>

The 4×4 matrix in Output 63.2.5 is the estimated unstructured covariance matrix. It is the estimate of the first block of R, and the other 26 blocks all have the same estimate.

Output 63.2.6 Repeated Measures Analysis (continued)

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Z</th>
<th>Pr Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN(1, 1)</td>
<td>Person</td>
<td>5.1192</td>
<td>1.4169</td>
<td>3.61</td>
<td>0.0002</td>
</tr>
<tr>
<td>UN(2, 1)</td>
<td>Person</td>
<td>2.4409</td>
<td>0.9835</td>
<td>2.48</td>
<td>0.0131</td>
</tr>
<tr>
<td>UN(2, 2)</td>
<td>Person</td>
<td>3.9279</td>
<td>1.0824</td>
<td>3.63</td>
<td>0.0001</td>
</tr>
<tr>
<td>UN(3, 1)</td>
<td>Person</td>
<td>3.6105</td>
<td>1.2767</td>
<td>2.83</td>
<td>0.0047</td>
</tr>
<tr>
<td>UN(3, 2)</td>
<td>Person</td>
<td>2.7175</td>
<td>1.0740</td>
<td>2.53</td>
<td>0.0114</td>
</tr>
<tr>
<td>UN(3, 3)</td>
<td>Person</td>
<td>5.9798</td>
<td>1.6279</td>
<td>3.67</td>
<td>0.0001</td>
</tr>
<tr>
<td>UN(4, 1)</td>
<td>Person</td>
<td>2.5222</td>
<td>1.0649</td>
<td>2.37</td>
<td>0.0179</td>
</tr>
<tr>
<td>UN(4, 2)</td>
<td>Person</td>
<td>3.0624</td>
<td>1.0135</td>
<td>3.02</td>
<td>0.0025</td>
</tr>
<tr>
<td>UN(4, 3)</td>
<td>Person</td>
<td>3.8235</td>
<td>1.2508</td>
<td>3.06</td>
<td>0.0022</td>
</tr>
<tr>
<td>UN(4, 4)</td>
<td>Person</td>
<td>4.6180</td>
<td>1.2573</td>
<td>3.67</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

The “Covariance Parameter Estimates” table in Output 63.2.6 lists the 10 estimated covariance parameters in order; note their correspondence to the first block of R displayed in Output 63.2.5. The parameter estimates are labeled according to their location in the block in the Cov Parm column, and all of these estimates are associated with Person as the subject effect. The Std Error column lists approximate standard errors of the covariance parameters obtained from the inverse Hessian matrix. These standard errors lead to approximate Wald Z statistics, which are compared with the standard normal distribution. The results of these tests indicate that all the parameters are significantly different from 0; however, the Wald test can be unreliable in small samples.

To carry out Wald tests of various linear combinations of these parameters, use the following procedure. First, run the statements again, adding the ASYCOV option and an ODS statement:

```plaintext
ods output CovParms=cp AsyCov=asy;
proc mixed data=pr method=ml covtest asy cov;
   class Person Gender;
   model y = Gender Age Gender*Age / s;
   repeated / type=un subject=Person r;
run;
```
Example 63.2: Repeated Measures

This creates two data sets, cp and asy, which contain the covariance parameter estimates and their asymptotic variance covariance matrix, respectively. Then read these data sets into the SAS/IML matrix programming language as follows:

```
proc iml;
  use cp;
  read all var (Estimate) into est;
  use asy;
  read all var ('CovP1': 'CovP10') into asy;
quit;
```

You can then construct your desired linear combinations and corresponding quadratic forms with the asy matrix.

**Output 63.2.7 Repeated Measures Analysis (continued)**

<table>
<thead>
<tr>
<th>Fit Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 Log Likelihood</td>
</tr>
<tr>
<td>AIC (Smaller is Better)</td>
</tr>
<tr>
<td>AICC (Smaller is Better)</td>
</tr>
<tr>
<td>BIC (Smaller is Better)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Null Model Likelihood Ratio Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>9</td>
</tr>
</tbody>
</table>

The null model likelihood ratio test (LRT) in Output 63.2.7 is highly significant for this model, indicating that the unstructured covariance matrix is preferred to the diagonal matrix of the ordinary least squares null model. The degrees of freedom for this test is 9, which is the difference between 10 and the 1 parameter for the null model’s diagonal matrix.

**Output 63.2.8 Repeated Measures Analysis (continued)**

<table>
<thead>
<tr>
<th>Solution for Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

The “Solution for Fixed Effects” table in Output 63.2.8 lists the solution vector for the fixed effects. The estimate of the boys’ intercept is 15.8423, while that for the girls is $15.8423 + 1.5831 = 17.0654$. Similarly,
the estimate for the boys’ slope is 0.8268, while that for the girls is $0.8268 - 0.3504 = 0.4764$. Thus the girls’ starting point is larger than that for the boys, but their growth rate is about half that of the boys.

Note that two of the estimates equal 0; this is a result of the overparameterized model used by PROC MIXED. You can obtain a full-rank parameterization by using the following MODEL statement:

```
model y = Gender Gender*Age / noint s;
```

Here, the NOINT option causes the different intercepts to be fit directly as the two levels of Gender. However, this alternative specification results in different tests for these effects.

**Output 63.2.9** Repeated Measures Analysis (continued)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>25</td>
<td>1.17</td>
<td>0.2904</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>25</td>
<td>110.54</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Age*Gender</td>
<td>1</td>
<td>25</td>
<td>7.99</td>
<td>0.0091</td>
</tr>
</tbody>
</table>

The “Type 3 Tests of Fixed Effects” table in Output 63.2.9 displays Type 3 tests for all of the fixed effects. These tests are partial in the sense that they account for all of the other fixed effects in the model. In addition, you can use the HTYPE= option in the MODEL statement to obtain Type 1 (sequential) or Type 2 (also partial) tests of effects.

It is usually best to consider higher-order terms first, and in this case the Age*Gender test reveals a difference between the slopes that is statistically significant at the 1% level. Note that the p-value for this test (0.0091) is the same as the p-value in the “Age*Gender F” row in the “Solution for Fixed Effects” table (Output 63.2.8) and that the $F$ statistic (7.99) is the square of the $t$ statistic (–2.83), ignoring rounding error. Similar connections are evident among the other rows in these two tables.

The Age test is one for an overall growth curve accounting for possible heterogeneous slopes, and it is highly significant. Finally, the Gender row tests the null hypothesis of a common intercept, and this hypothesis cannot be rejected from these data.

As an alternative to the $F$ tests shown here, you can carry out likelihood ratio tests of various hypotheses by fitting the reduced models, subtracting $–2 \log$ likelihoods, and comparing the resulting statistics with $\chi^2$ distributions.

Since the different levels of the repeated effect represent different years, it is natural to try fitting a time series model to the data within each subject. To obtain time series structures in R, you can replace TYPE=UN with TYPE=AR(1) or TYPE=TOEP to obtain the first- or nth-order autoregressive covariance matrices, respectively. For example, the statements to fit an AR(1) structure are as follows:

```r
/* first-order autoregressive */
proc mixed data=pr method=ml;
   class Person Gender;
   model y = Gender Age Gender*Age / s;
   repeated / type=ar(1) sub=Person r;
run;
```
Example 63.2: Repeated Measures

To fit a random coefficients model, use the following statements:

```sas
/* random coefficients model */
proc mixed data=pr method=ml;
  class Person Gender;
  model y = Gender Age Gender*Age / s;
  random intercept Age / type=un sub=Person g;
run;
```

This specifies an unstructured covariance matrix for the random intercept and slope. In mixed model notation, \( G \) is block diagonal with identical 2x2 unstructured blocks for each person. By default, \( R \) becomes \( \sigma^2 I \). See Example 63.5 for further information about this model.

Finally, you can fit a compound symmetry structure by using `TYPE=CS`, as follows:

```sas
proc mixed data=pr method=ml covtest;
  class Person Gender;
  model y = Gender Age Gender*Age / s;
  repeated / type=cs subject=Person r;
run;
```

The results from this analysis are shown in Output 63.2.10–Output 63.2.17.

The “Model Information” table in Output 63.2.10 is the same as before except for the change in “Covariance Structure.”

**Output 63.2.10** Repeated Measures Analysis with Compound Symmetry Structure

```
The Mixed Procedure

Model Information

Data Set WORK.PR
Dependent Variable y
Covariance Structure Compound Symmetry
Subject Effect Person
Estimation Method ML
Residual Variance Method Profile
Fixed Effects SE Method Model-Based
Degrees of Freedom Method Between-Within
```

The “Dimensions” table in Output 63.2.11 shows that there are only two covariance parameters in the compound symmetry model; this covariance structure has common variance and common covariance.
Output 63.2.11  Analysis with Compound Symmetry (continued)

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person</td>
<td>27</td>
<td>1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27</td>
</tr>
<tr>
<td>Gender</td>
<td>2</td>
<td>F M</td>
</tr>
</tbody>
</table>

Dimensions

- Covariance Parameters: 2
- Columns in X: 6
- Columns in Z: 0
- Subjects: 27
- Max Obs per Subject: 4

Number of Observations

- Number of Observations Read: 108
- Number of Observations Used: 108
- Number of Observations Not Used: 0

Since the data are balanced, only one step is required to find the estimates (Output 63.2.12).

Output 63.2.12  Analysis with Compound Symmetry (continued)

Iteration History

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Evaluations</th>
<th>-2 Log Like</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>478.24175986</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>428.63905802</td>
<td>0.00000000</td>
</tr>
</tbody>
</table>

Convergence criteria met.

Output 63.2.13 displays the estimated R matrix for the first subject. Note the compound symmetry structure here, which consists of a common covariance with a diagonal enhancement.
Example 63.2: Repeated Measures

Output 63.2.13 Analysis with Compound Symmetry (continued)

| Estimated R Matrix for Person 1 |
|-----------------|----------------|----------------|----------------|
| Row             | Col1           | Col2           | Col3           | Col4           |
| 1               | 4.9052         | 3.0306         | 3.0306         | 3.0306         |
| 2               | 3.0306         | 4.9052         | 3.0306         | 3.0306         |
| 3               | 3.0306         | 3.0306         | 4.9052         | 3.0306         |
| 4               | 3.0306         | 3.0306         | 3.0306         | 4.9052         |

The common covariance is estimated to be 3.0306, as listed in the CS row of the “Covariance Parameter Estimates” table in Output 63.2.14, and the residual variance is estimated to be 1.8746, as listed in the Residual row. You can use these two numbers to estimate the intraclass correlation coefficient (ICC) for this model. Here, the ICC estimate equals $3.0306 / (3.0306 + 1.8746) = 0.6178$. You can also obtain this number by adding the RCORR option to the REPEATED statement.

Output 63.2.14 Analysis with Compound Symmetry (continued)

<table>
<thead>
<tr>
<th>Covariance Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cov Parm</td>
</tr>
<tr>
<td>CS</td>
</tr>
<tr>
<td>Residual</td>
</tr>
</tbody>
</table>

In the case shown in Output 63.2.15, the null model LRT has only one degree of freedom, corresponding to the common covariance parameter. The test indicates that modeling this extra covariance is superior to fitting the simple null model.

Output 63.2.15 Analysis with Compound Symmetry (continued)

<table>
<thead>
<tr>
<th>Fit Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 Log Likelihood</td>
</tr>
<tr>
<td>AIC (Smaller is Better)</td>
</tr>
<tr>
<td>AICC (Smaller is Better)</td>
</tr>
<tr>
<td>BIC (Smaller is Better)</td>
</tr>
</tbody>
</table>

Null Model Likelihood Ratio Test

<table>
<thead>
<tr>
<th>DF</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49.60</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Note that the fixed-effects estimates and their standard errors (Output 63.2.16) are not very different from those in the preceding unstructured example (Output 63.2.8).
Output 63.2.16 Analysis with Compound Symmetry (continued)

| Effect          | Gender | Estimate | Error | DF  | t Value | Pr > |t| |
|-----------------|--------|----------|-------|-----|---------|------|---|
| Intercept       |        | 16.3406  | 0.9631| 25  | 16.97   | <.0001 |
| Gender F        |        | 1.0321   | 1.5089| 25  | 0.68    | 0.5003 |
| Gender M        |        |          |       |     |         |       |
| Age             |        | 0.7844   | 0.0765| 79  | 10.25   | <.0001 |
| Age*Gender F    |        | -0.3048  | 0.1199| 79  | -2.54   | 0.0130 |
| Age*Gender M    |        |          |       |     |         |       |

The $F$ tests shown in Output 63.2.17 are also similar to those from the preceding unstructured example (Output 63.2.9). Again, the slopes are significantly different but the intercepts are not.

Output 63.2.17 Analysis with Compound Symmetry (continued)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>25</td>
<td>0.47</td>
<td>0.5003</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>79</td>
<td>111.10</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Age*Gender</td>
<td>1</td>
<td>79</td>
<td>6.46</td>
<td>0.0130</td>
</tr>
</tbody>
</table>

You can fit the same compound symmetry model with the following specification by using the RANDOM statement:

```
proc mixed data=pr method=ml;
  class Person Gender;
  model y = Gender Age Gender*Age / s;
  random Person;
run;
```

Compound symmetry is the structure that Jennrich and Schluchter deemed best among the ones they fit. To carry the analysis one step further, you can use the GROUP= option as follows to specify heterogeneity of this structure across girls and boys:

```
proc mixed data=pr method=ml;
  class Person Gender;
  model y = Gender Age Gender*Age / s;
  repeated / type=cs subject=Person group=Gender;
run;
```

The results from this analysis are shown in Output 63.2.18–Output 63.2.24. Note that in Output 63.2.18 Gender is listed as a “Group Effect.”
Output 63.2.18  Repeated Measures Analysis with Heterogeneous Structures

The Mixed Procedure
Model Information

Data Set WORK.PR
Dependent Variable y
Covariance Structure Compound Symmetry
Subject Effect Person
Group Effect Gender
Estimation Method ML
Residual Variance Method None
Fixed Effects SE Method Model-Based
Degrees of Freedom Method Between-Within

The four covariance parameters listed in Output 63.2.19 result from the two compound symmetry structures corresponding to the two levels of Gender.

Output 63.2.19  Analysis with Heterogeneous Structures (continued)

Class Level Information

Class Levels Values
Person 27 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Gender 2 F M

Dimensions

Covariance Parameters 4
Columns in X 6
Columns in Z 0
Subjects 27
Max Obs per Subject 4

Number of Observations

Number of Observations Read 108
Number of Observations Used 108
Number of Observations Not Used 0

As Output 63.2.20 shows, even with the heterogeneity, only one iteration is required for convergence.
Output 63.2.20  Analysis with Heterogeneous Structures (continued)

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Evaluations</th>
<th>-2 Log Like</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>478.24175986</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>408.81297228</td>
<td>0.0000000</td>
</tr>
</tbody>
</table>

Convergence criteria met.

The “Covariance Parameter Estimates” table in Output 63.2.21 lists the heterogeneous estimates. Note that both the common covariance and the diagonal enhancement differ between girls and boys.

Output 63.2.21  Analysis with Heterogeneous Structures (continued)

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Group</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance</td>
<td>Person</td>
<td>Gender F</td>
<td>0.5900</td>
</tr>
<tr>
<td>CS</td>
<td>Person</td>
<td>Gender F</td>
<td>3.8804</td>
</tr>
<tr>
<td>Variance</td>
<td>Person</td>
<td>Gender M</td>
<td>2.7577</td>
</tr>
<tr>
<td>CS</td>
<td>Person</td>
<td>Gender M</td>
<td>2.4463</td>
</tr>
</tbody>
</table>

As Output 63.2.22 shows, both Akaike’s information criterion (424.8) and Schwarz’s Bayesian information criterion (435.2) are smaller for this model than for the homogeneous compound symmetry model (440.6 and 448.4, respectively). This indicates that the heterogeneous model is more appropriate. To construct the likelihood ratio test between the two models, subtract the –2 log likelihood values: 428.6 – 408.8 = 19.8. Comparing this value with the $\chi^2$ distribution with two degrees of freedom yields a $p$-value less than 0.0001, again favoring the heterogeneous model.

Output 63.2.22  Analysis with Heterogeneous Structures (continued)

<table>
<thead>
<tr>
<th>Fit Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 Log Likelihood</td>
</tr>
<tr>
<td>AIC (Smaller is Better)</td>
</tr>
<tr>
<td>AICC (Smaller is Better)</td>
</tr>
<tr>
<td>BIC (Smaller is Better)</td>
</tr>
</tbody>
</table>

Null Model Likelihood Ratio Test

<table>
<thead>
<tr>
<th>DF</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>69.43</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Note that the fixed-effects estimates shown in Output 63.2.23 are the same as in the homogeneous case, but the standard errors are different.

**Output 63.2.23** Analysis with Heterogeneous Structures (*continued*)

| Effect    | Gender | Estimate | Error | DF   | t Value | Pr > |t| |
|-----------|--------|----------|-------|------|---------|------|---|
| Intercept |        | 16.3406  | 1.1130| 25   | 14.68   | <.0001|
| Gender    | F      | 1.0321   | 1.3890| 25   | 0.74    | 0.4644|
| Gender    | M      | 0        |       |      |         |      |  
| Age       |        | 0.7844   | 0.09283| 79   | 8.45    | <.0001|
| Age*Gender| F      | -0.3048  | 0.1063| 79   | -2.87   | 0.0053|
| Age*Gender| M      | 0        |       |      |         |      |  

The fixed-effects tests shown in Output 63.2.24 are similar to those from previous models, although the *p*-values do change as a result of specifying a different covariance structure. It is important for you to select a reasonable covariance structure in order to obtain valid inferences for your fixed effects.

**Output 63.2.24** Analysis with Heterogeneous Structures (*continued*)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>25</td>
<td>0.55</td>
<td>0.4644</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>79</td>
<td>141.37</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Age*Gender</td>
<td>1</td>
<td>79</td>
<td>8.22</td>
<td>0.0053</td>
</tr>
</tbody>
</table>

**Example 63.3: Plotting the Likelihood**

The data for this example are from Hemmerle and Hartley (1973) and are also used for an example in the VARCOMP procedure. The response variable consists of measurements from an oven experiment, and the model contains a fixed effect *A* and random effects *B* and *A*B.

The SAS statements are as follows:

```sas
data hh;
  input a b y @@;
datalines;
  1 1 237 1 1 254 1 1 246
  1 2 178 1 2 179
  2 1 208 2 1 178 2 1 187
  2 2 146 2 2 145 2 2 141
  3 1 186 3 1 183
  3 2 142 3 2 125 3 2 136
;```
ods output ParmSearch=parms;
proc mixed data=hh asycov mmeq mmeqsol covtest;
   class a b;
   model y = a / outp=predicted;
   random b a*b;
   lsmeans a;
   parms (17 to 20 by .1) (.3 to .4 by .005) (1.0);
run;
proc print data=predicted;
run;

The ASYCOV option in the PROC MIXED statement requests the asymptotic variance matrix of the covariance parameter estimates. This matrix is the observed inverse Fisher information matrix, which equals $2H^{-1}$, where $H$ is the Hessian matrix of the objective function evaluated at the final covariance parameter estimates. The MMEQ and MMEQSOL options in the PROC MIXED statement request that the mixed model equations and their solution be displayed.

The OUTP= option in the MODEL statement produces the data set predicted, containing the predicted values. Least squares means (LSMEANS) are requested for A. The PARMS and ODS statements are used to construct a data set containing the likelihood surface.

The results from this analysis are shown in Output 63.3.1–Output 63.3.13.

The “Model Information” table in Output 63.3.1 lists details about this variance components model.

Output 63.3.1 Model Information

<table>
<thead>
<tr>
<th>The Mixed Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Information</td>
</tr>
<tr>
<td>Data Set</td>
</tr>
<tr>
<td>Dependent Variable</td>
</tr>
<tr>
<td>Covariance Structure</td>
</tr>
<tr>
<td>Estimation Method</td>
</tr>
<tr>
<td>Residual Variance Method</td>
</tr>
<tr>
<td>Fixed Effects SE Method</td>
</tr>
<tr>
<td>Degrees of Freedom Method</td>
</tr>
</tbody>
</table>

The “Class Level Information” table in Output 63.3.2 lists the levels for A and B.

Output 63.3.2 Class Level Information

<table>
<thead>
<tr>
<th>Class Level Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
</tr>
<tr>
<td>a</td>
</tr>
<tr>
<td>b</td>
</tr>
</tbody>
</table>
Example 63.3: Plotting the Likelihood

The “Dimensions” table in Output 63.3.3 reveals that X is $16 \times 4$ and Z is $16 \times 8$. Since there are no SUBJECT= effects, PROC MIXED considers the data to be effectively from one subject with 16 observations.

### Output 63.3.3 Model Dimensions and Number of Observations

<table>
<thead>
<tr>
<th>Dimensions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariance Parameters</td>
<td>3</td>
</tr>
<tr>
<td>Columns in X</td>
<td>4</td>
</tr>
<tr>
<td>Columns in Z</td>
<td>8</td>
</tr>
<tr>
<td>Subjects</td>
<td>1</td>
</tr>
<tr>
<td>Max Obs per Subject</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Observations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations Read</td>
<td>16</td>
</tr>
<tr>
<td>Number of Observations Used</td>
<td>16</td>
</tr>
<tr>
<td>Number of Observations Not Used</td>
<td>0</td>
</tr>
</tbody>
</table>

Only a portion of the “Parameter Search” table is shown in Output 63.3.4 because the full listing has 651 rows.

### Output 63.3.4 Selected Results of Parameter Search

<table>
<thead>
<tr>
<th>The Mixed Procedure</th>
<th>-2 Res Log</th>
<th>Like</th>
</tr>
</thead>
<tbody>
<tr>
<td>CovP1</td>
<td>CovP2</td>
<td>CovP3</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3000</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3050</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3100</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3150</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3200</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3250</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3300</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3350</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3400</td>
<td>1.0000</td>
</tr>
<tr>
<td>17.0000</td>
<td>0.3450</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3500</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3600</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3650</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3700</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3750</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3800</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3850</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3900</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.3950</td>
<td>1.0000</td>
</tr>
<tr>
<td>20.0000</td>
<td>0.4000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
As Output 63.3.5 shows, convergence occurs quickly because PROC MIXED starts from the best value from the grid search.

**Output 63.3.5** Iteration History and Convergence Status

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Evaluations</th>
<th>-2 Res Log Like</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>104.93416367</td>
<td>0.00000000</td>
</tr>
</tbody>
</table>

Convergence criteria met.

The “Covariance Parameter Estimates” table in Output 63.3.6 lists the variance components estimates. Note that B is much more variable than A*B.

**Output 63.3.6** Estimated Covariance Parameters

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Estimate</th>
<th>Error</th>
<th>Value</th>
<th>Pr &gt; Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>1464.36</td>
<td>2098.01</td>
<td>0.70</td>
<td>0.2426</td>
</tr>
<tr>
<td>a*b</td>
<td>26.9581</td>
<td>59.6570</td>
<td>0.45</td>
<td>0.3257</td>
</tr>
<tr>
<td>Residual</td>
<td>78.8426</td>
<td>35.3512</td>
<td>2.23</td>
<td>0.0129</td>
</tr>
</tbody>
</table>

The asymptotic covariance matrix in Output 63.3.7 also reflects the large variability of B relative to A*B.

**Output 63.3.7** Asymptotic Covariance Matrix of Covariance Parameters

<table>
<thead>
<tr>
<th>Row</th>
<th>Cov Parm</th>
<th>CovP1</th>
<th>CovP2</th>
<th>CovP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>b</td>
<td>4401640</td>
<td>1.2831</td>
<td>-273.32</td>
</tr>
<tr>
<td>2</td>
<td>a*b</td>
<td>1.2831</td>
<td>3558.96</td>
<td>-502.84</td>
</tr>
<tr>
<td>3</td>
<td>Residual</td>
<td>-273.32</td>
<td>-502.84</td>
<td>1249.71</td>
</tr>
</tbody>
</table>
As Output 63.3.8 shows, the PARMS likelihood ratio test (LRT) compares the best model from the grid search with the final fitted model. Since these models are nearly the same, the LRT is not significant.

**Output 63.3.8** Fit Statistics and Likelihood Ratio Test

<table>
<thead>
<tr>
<th>Fit Statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 Res Log Likelihood</td>
<td>104.9</td>
</tr>
<tr>
<td>AIC (Smaller is Better)</td>
<td>110.9</td>
</tr>
<tr>
<td>AICC (Smaller is Better)</td>
<td>113.6</td>
</tr>
<tr>
<td>BIC (Smaller is Better)</td>
<td>107.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PARMS Model Likelihood Ratio Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

The mixed model equations are analogous to the normal equations in the standard linear model. As Output 63.3.9 shows, for this example, rows 1–4 correspond to the fixed effects, rows 5–12 correspond to the random effects, and Col13 corresponds to the dependent variable.
The solution matrix in Output 63.3.10 results from sweeping all but the last row of the mixed model equations matrix. The final column contains a solution vector for the fixed and random effects. The first four rows correspond to fixed effects and the last eight correspond to random effects.
The A factor is significant at the 5% level (Output 63.3.11).

Output 63.3.11 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>2</td>
<td>2</td>
<td>28.00</td>
<td>0.0345</td>
</tr>
</tbody>
</table>

Output 63.3.12 shows that the significance of A appears to be from the difference between its first level and its other two levels.
Output 63.3.12  Least Squares Means for A Effect

| Effect | a  | Estimate | Error   | DF | t Value | Pr > |t|   |
|--------|----|----------|---------|----|---------|------|----|
| a      | 1  | 212.82   | 27.6014 | 2  | 7.71    | 0.0164 |
| a      | 2  | 167.50   | 27.5463 | 2  | 6.08    | 0.0260 |
| a      | 3  | 159.61   | 27.6014 | 2  | 5.78    | 0.0286 |

Output 63.3.13 lists the predicted values from the model. These values are the sum of the fixed-effects estimates and the empirical best linear unbiased predictors (EBLUPs) of the random effects.

Output 63.3.13  Predicted Values

<table>
<thead>
<tr>
<th>Obs</th>
<th>a</th>
<th>b</th>
<th>y</th>
<th>Pred</th>
<th>Pred DF</th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
<th>Resid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>237</td>
<td>242.723</td>
<td>4.72563</td>
<td>10</td>
<td>0.05</td>
<td>232.193</td>
<td>253.252</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>254</td>
<td>242.723</td>
<td>4.72563</td>
<td>10</td>
<td>0.05</td>
<td>232.193</td>
<td>253.252</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>246</td>
<td>242.723</td>
<td>4.72563</td>
<td>10</td>
<td>0.05</td>
<td>232.193</td>
<td>253.252</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>178</td>
<td>182.916</td>
<td>5.52589</td>
<td>10</td>
<td>0.05</td>
<td>170.603</td>
<td>195.228</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>179</td>
<td>182.916</td>
<td>5.52589</td>
<td>10</td>
<td>0.05</td>
<td>170.603</td>
<td>195.228</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
<td>208</td>
<td>192.670</td>
<td>4.70076</td>
<td>10</td>
<td>0.05</td>
<td>182.196</td>
<td>203.144</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
<td>178</td>
<td>192.670</td>
<td>4.70076</td>
<td>10</td>
<td>0.05</td>
<td>182.196</td>
<td>203.144</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>1</td>
<td>187</td>
<td>192.670</td>
<td>4.70076</td>
<td>10</td>
<td>0.05</td>
<td>182.196</td>
<td>203.144</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>2</td>
<td>146</td>
<td>142.330</td>
<td>4.70076</td>
<td>10</td>
<td>0.05</td>
<td>131.856</td>
<td>152.804</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>2</td>
<td>145</td>
<td>142.330</td>
<td>4.70076</td>
<td>10</td>
<td>0.05</td>
<td>131.856</td>
<td>152.804</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>2</td>
<td>141</td>
<td>142.330</td>
<td>4.70076</td>
<td>10</td>
<td>0.05</td>
<td>131.856</td>
<td>152.804</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>1</td>
<td>186</td>
<td>185.687</td>
<td>5.52589</td>
<td>10</td>
<td>0.05</td>
<td>173.374</td>
<td>197.999</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>1</td>
<td>183</td>
<td>185.687</td>
<td>5.52589</td>
<td>10</td>
<td>0.05</td>
<td>173.374</td>
<td>197.999</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>2</td>
<td>142</td>
<td>133.542</td>
<td>4.72563</td>
<td>10</td>
<td>0.05</td>
<td>123.013</td>
<td>144.072</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>2</td>
<td>125</td>
<td>133.542</td>
<td>4.72563</td>
<td>10</td>
<td>0.05</td>
<td>123.013</td>
<td>144.072</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>2</td>
<td>136</td>
<td>133.542</td>
<td>4.72563</td>
<td>10</td>
<td>0.05</td>
<td>123.013</td>
<td>144.072</td>
</tr>
</tbody>
</table>

To plot the likelihood surface by using ODS Graphics, use the following statements:

```plaintext
proc template;
  define statgraph surface;
  begingraph;
    layout overlay3d;
      surfaceplotparm x=CovP1 y=CovP2 z=ResLogLike;
    endlayout;
  endgraph;
end;
run;
proc sgrender data=parms template=surface;
run;
```

The results from this plot are shown in Output 63.3.14. The peak of the surface is the REML estimates for the B and A*B variance components.
Example 63.4: Known G and R

This animal breeding example from Henderson (1984, p. 48) considers multiple traits. The data are artificial and consist of measurements of two traits on three animals, but the second trait of the third animal is missing. Assuming an additive genetic model, you can use PROC MIXED to predict the breeding value of both traits on all three animals and also to predict the second trait of the third animal. The data are as follows:

```plaintext
data h;
    input Trait Animal Y;
datalines;
1 1 6
1 2 8
1 3 7
2 1 9
2 2 5
2 3 .
;
```

Output 63.3.14 Plot of Likelihood Surface

![Plot of Likelihood Surface](image-url)
Both G and R are known.

\[
G = \begin{bmatrix}
2 & 1 & 1 & 2 & 1 & 1 \\
1 & 2 & .5 & 1 & 2 & .5 \\
1 & .5 & 2 & 1 & .5 & 2 \\
2 & 1 & 1 & 3 & 1.5 & 1.5 \\
1 & 2 & .5 & 1.5 & 3 & .75 \\
1 & .5 & 2 & 1.5 & .75 & 3
\end{bmatrix}
\]

\[
R = \begin{bmatrix}
4 & 0 & 0 & 1 & 0 & 0 \\
0 & 4 & 0 & 0 & 1 & 0 \\
0 & 0 & 4 & 0 & 0 & 1 \\
1 & 0 & 0 & 5 & 0 & 0 \\
0 & 1 & 0 & 0 & 5 & 0 \\
0 & 0 & 1 & 0 & 0 & 5
\end{bmatrix}
\]

In order to read G into PROC MIXED by using the GDATA= option in the RANDOM statement, perform the following DATA step:

```r
data g;
  input Row Col1-Col6;
  datalines;
  1 2 1 1 2 1 1
  2 1 2 .5 1 2 .5
  3 1 .5 2 1 .5 2
  4 2 1 1 3 1.5 1.5
  5 1 2 .5 1.5 3 .75
  6 1 .5 2 1.5 .75 3
;```

The preceding data are in the dense representation for a GDATA= data set. You can also construct a data set with the sparse representation by using Row, Col, and Value variables, although this would require 21 observations instead of 6 for this example.

The PROC MIXED statements are as follows:

```r
proc mixed data=h mmeq mmeqsol;
  class Trait Animal;
  model Y = Trait / noint s outp=predicted;
  random Trait*Animal / type=un gdata=g g gi s;
  repeated / type=un sub=Animal r ri;
  parms (4) (1) (5) / noiter;
run;
proc print data=predicted;
run;
```

The MMEQ and MMEQSOL options request the mixed model equations and their solution. The variables Trait and Animal are classification variables, and Trait defines the entire X matrix for the fixed-effects portion of the model, since the intercept is omitted with the NOINT option. The fixed-effects solution vector and predicted values are also requested by using the S and OUTP= options, respectively.
The random effect Trait*Animal leads to a $Z$ matrix with six columns, the first five corresponding to the identity matrix and the last consisting of 0s. An unstructured $G$ matrix is specified by using the TYPE=UN option, and it is read into PROC MIXED from a SAS data set by using the GDATA=G specification. The $G$ and GI options request the display of $G$ and $G^{-1}$, respectively. The S option requests that the random-effects solution vector be displayed.

Note that the preceding $R$ matrix is block diagonal if the data are sorted by animals. The REPEATED statement exploits this fact by requesting $R$ to have unstructured $2 \times 2$ blocks corresponding to animals, which are the subjects. The $R$ and RI options request that the estimated $2 \times 2$ blocks for the first animal and its inverse be displayed. The PARMS statement lists the parameters of this $2 \times 2$ matrix. Note that the parameters from $G$ are not specified in the PARMS statement because they have already been assigned by using the GDATA= option in the RANDOM statement. The NOITER option prevents PROC MIXED from computing residual (restricted) maximum likelihood estimates; instead, the known values are used for inferences.

The results from this analysis are shown in Output 63.4.1–Output 63.4.12.

The “Unstructured” covariance structure (Output 63.4.1) applies to both $G$ and $R$ here. The levels of Trait and Animal have been specified correctly.

**Output 63.4.1** Model and Class Level Information

<table>
<thead>
<tr>
<th>The Mixed Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Information</td>
</tr>
<tr>
<td>Data Set</td>
</tr>
<tr>
<td>Dependent Variable</td>
</tr>
<tr>
<td>Covariance Structure</td>
</tr>
<tr>
<td>Subject Effect</td>
</tr>
<tr>
<td>Estimation Method</td>
</tr>
<tr>
<td>Residual Variance Method</td>
</tr>
<tr>
<td>Fixed Effects SE Method</td>
</tr>
<tr>
<td>Degrees of Freedom Method</td>
</tr>
</tbody>
</table>

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait</td>
<td>2</td>
<td>1 2</td>
</tr>
<tr>
<td>Animal</td>
<td>3</td>
<td>1 2 3</td>
</tr>
</tbody>
</table>

The three covariance parameters indicated in Output 63.4.2 correspond to those from the $R$ matrix. Those from $G$ are considered fixed and known because of the GDATA= option.
### Output 63.4.2 Model Dimensions and Number of Observations

<table>
<thead>
<tr>
<th>Dimensions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariance Parameters</td>
<td>3</td>
</tr>
<tr>
<td>Columns in X</td>
<td>2</td>
</tr>
<tr>
<td>Columns in Z</td>
<td>6</td>
</tr>
<tr>
<td>Subjects</td>
<td>1</td>
</tr>
<tr>
<td>Max Obs per Subject</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Observations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations Read</td>
<td>6</td>
</tr>
<tr>
<td>Number of Observations Used</td>
<td>5</td>
</tr>
<tr>
<td>Number of Observations Not Used</td>
<td>1</td>
</tr>
</tbody>
</table>

Because starting values for the covariance parameters are specified in the PARMS statement, the MIXED procedure prints the residual (restricted) log likelihood at the starting values. Because of the NOITER option in the PARMS statement, this is also the final log likelihood in this analysis (Output 63.4.3).

### Output 63.4.3 REML Log Likelihood

<table>
<thead>
<tr>
<th>Parameter Search</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CovP1</td>
<td>4.0000</td>
</tr>
<tr>
<td>CovP2</td>
<td>1.0000</td>
</tr>
<tr>
<td>CovP3</td>
<td>5.0000</td>
</tr>
<tr>
<td>Res Log Like</td>
<td>-7.3731</td>
</tr>
<tr>
<td>-2 Res Log Like</td>
<td>14.7463</td>
</tr>
</tbody>
</table>

The block of R corresponding to the first animal and the inverse of this block are shown in Output 63.4.4.

### Output 63.4.4 Inverse R Matrix

<table>
<thead>
<tr>
<th>Estimated R Matrix for Animal 1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Row</td>
<td>Col1</td>
</tr>
<tr>
<td>1</td>
<td>4.0000</td>
</tr>
<tr>
<td>2</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimated Inv(R) Matrix for Animal 1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Row</td>
<td>Col1</td>
</tr>
<tr>
<td>1</td>
<td>0.2632</td>
</tr>
<tr>
<td>2</td>
<td>-0.05263</td>
</tr>
</tbody>
</table>
The \( G \) matrix as specified in the GDATA= data set and its inverse are shown in Output 63.4.5 and Output 63.4.6.

**Output 63.4.5 G Matrix**

<table>
<thead>
<tr>
<th>Row</th>
<th>Effect</th>
<th>Trait</th>
<th>Animal</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Trait*Animal</td>
<td>1</td>
<td>1</td>
<td>2.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>2.0000</td>
</tr>
<tr>
<td>2</td>
<td>Trait*Animal</td>
<td>1</td>
<td>2</td>
<td>1.0000</td>
<td>2.0000</td>
<td>0.5000</td>
<td>1.0000</td>
</tr>
<tr>
<td>3</td>
<td>Trait*Animal</td>
<td>1</td>
<td>3</td>
<td>1.0000</td>
<td>0.5000</td>
<td>2.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>4</td>
<td>Trait*Animal</td>
<td>2</td>
<td>1</td>
<td>2.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>3.0000</td>
</tr>
<tr>
<td>5</td>
<td>Trait*Animal</td>
<td>2</td>
<td>2</td>
<td>1.0000</td>
<td>2.0000</td>
<td>0.5000</td>
<td>1.5000</td>
</tr>
<tr>
<td>6</td>
<td>Trait*Animal</td>
<td>2</td>
<td>3</td>
<td>1.0000</td>
<td>0.5000</td>
<td>2.0000</td>
<td>1.5000</td>
</tr>
</tbody>
</table>

**Estimated G Matrix**

<table>
<thead>
<tr>
<th>Row</th>
<th>Col5</th>
<th>Col6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>2</td>
<td>2.0000</td>
<td>0.5000</td>
</tr>
<tr>
<td>3</td>
<td>0.5000</td>
<td>2.0000</td>
</tr>
<tr>
<td>4</td>
<td>1.5000</td>
<td>1.5000</td>
</tr>
<tr>
<td>5</td>
<td>3.0000</td>
<td>0.7500</td>
</tr>
<tr>
<td>6</td>
<td>0.7500</td>
<td>3.0000</td>
</tr>
</tbody>
</table>

**Output 63.4.6 Inverse G Matrix**

<table>
<thead>
<tr>
<th>Row</th>
<th>Effect</th>
<th>Trait</th>
<th>Animal</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Trait*Animal</td>
<td>1</td>
<td>1</td>
<td>2.5000</td>
<td>-1.0000</td>
<td>-1.0000</td>
<td>-1.6667</td>
</tr>
<tr>
<td>2</td>
<td>Trait*Animal</td>
<td>1</td>
<td>2</td>
<td>-1.0000</td>
<td>2.0000</td>
<td>0.6667</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Trait*Animal</td>
<td>1</td>
<td>3</td>
<td>-1.0000</td>
<td>2.0000</td>
<td>0.6667</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Trait*Animal</td>
<td>2</td>
<td>1</td>
<td>-1.6667</td>
<td>0.6667</td>
<td>0.6667</td>
<td>1.6667</td>
</tr>
<tr>
<td>5</td>
<td>Trait*Animal</td>
<td>2</td>
<td>2</td>
<td>0.6667</td>
<td>-1.3333</td>
<td>-0.6667</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Trait*Animal</td>
<td>2</td>
<td>3</td>
<td>0.6667</td>
<td>-1.3333</td>
<td>-0.6667</td>
<td></td>
</tr>
</tbody>
</table>

**Estimated Inv(G) Matrix**

<table>
<thead>
<tr>
<th>Row</th>
<th>Col5</th>
<th>Col6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.6667</td>
<td>0.6667</td>
</tr>
<tr>
<td>2</td>
<td>-1.3333</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-1.3333</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.6667</td>
<td>-0.6667</td>
</tr>
<tr>
<td>5</td>
<td>1.3333</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.3333</td>
<td></td>
</tr>
</tbody>
</table>
The table of covariance parameter estimates in Output 63.4.7 displays only the parameters in $R$. Because of the GDATA= option in the RANDOM statement, the G-side parameters do not participate in the parameter estimation process. Because of the NOITER option in the PARMS statement, however, the R-side parameters in this output are identical to their starting values.

**Output 63.4.7**  R-Side Covariance Parameters

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN(1,1)</td>
<td>Animal</td>
<td>4.0000</td>
</tr>
<tr>
<td>UN(2,1)</td>
<td>Animal</td>
<td>1.0000</td>
</tr>
<tr>
<td>UN(2,2)</td>
<td>Animal</td>
<td>5.0000</td>
</tr>
</tbody>
</table>

The coefficients of the mixed model equations in Output 63.4.8 agree with Henderson (1984, p. 55). Recall from Output 63.4.1 that there are 2 columns in $X$ and 6 columns in $Z$. The first 8 columns of the mixed model equations correspond to the $X$ and $Z$ components. Column 9 represents the Y border.

**Output 63.4.8**  Mixed Model Equations with Y Border

<table>
<thead>
<tr>
<th>Row</th>
<th>Effect</th>
<th>Trait</th>
<th>Animal</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Trait</td>
<td>1</td>
<td></td>
<td>0.7763</td>
<td>-0.1053</td>
<td>0.2632</td>
<td>0.2632</td>
</tr>
<tr>
<td>2</td>
<td>Trait</td>
<td>2</td>
<td></td>
<td>-0.1053</td>
<td>0.4211</td>
<td>-0.05263</td>
<td>-0.05263</td>
</tr>
<tr>
<td>3</td>
<td>Trait*Animal</td>
<td>1</td>
<td>1</td>
<td>0.2632</td>
<td>-0.05263</td>
<td>2.7632</td>
<td>-1.0000</td>
</tr>
<tr>
<td>4</td>
<td>Trait*Animal</td>
<td>1</td>
<td>2</td>
<td>0.2632</td>
<td>-0.05263</td>
<td>-1.0000</td>
<td>2.2632</td>
</tr>
<tr>
<td>5</td>
<td>Trait*Animal</td>
<td>1</td>
<td>3</td>
<td>0.2500</td>
<td>0.2105</td>
<td>-1.7193</td>
<td>0.6667</td>
</tr>
<tr>
<td>6</td>
<td>Trait*Animal</td>
<td>2</td>
<td>1</td>
<td>-0.05263</td>
<td>0.2105</td>
<td>-1.7193</td>
<td>0.6667</td>
</tr>
<tr>
<td>7</td>
<td>Trait*Animal</td>
<td>2</td>
<td>2</td>
<td>-0.05263</td>
<td>0.2105</td>
<td>0.6667</td>
<td>-1.3860</td>
</tr>
<tr>
<td>8</td>
<td>Trait*Animal</td>
<td>2</td>
<td>3</td>
<td>0.6667</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Mixed Model Equations**

<table>
<thead>
<tr>
<th>Row</th>
<th>Col5</th>
<th>Col6</th>
<th>Col7</th>
<th>Col8</th>
<th>Col9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2500</td>
<td>-0.05263</td>
<td>-0.05263</td>
<td>4.6974</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.2105</td>
<td>0.2105</td>
<td>2.2105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-1.0000</td>
<td>-1.7193</td>
<td>0.6667</td>
<td>0.6667</td>
<td>1.1053</td>
</tr>
<tr>
<td>4</td>
<td>0.6667</td>
<td>-1.3860</td>
<td>1.8421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.2500</td>
<td>0.6667</td>
<td>0.6667</td>
<td>-1.3333</td>
<td>1.7500</td>
</tr>
<tr>
<td>6</td>
<td>0.6667</td>
<td>1.8772</td>
<td>-0.6667</td>
<td>-0.6667</td>
<td>1.5789</td>
</tr>
<tr>
<td>7</td>
<td>-0.6667</td>
<td>1.5439</td>
<td>0.6316</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-1.3333</td>
<td>-0.6667</td>
<td>1.3333</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The solution to the mixed model equations also matches that given by Henderson (1984, p. 55). After solving the augmented mixed model equations, you can find the solutions for fixed and random effects in the last column (Output 63.4.9).
Example 63.4: Known \( G \) and \( R \)

Output 63.4.9  Solutions of the Mixed Model Equations with Y Border

<table>
<thead>
<tr>
<th>Mixed Model Equations Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

Mixed Model Equations Solution

<table>
<thead>
<tr>
<th>Row</th>
<th>Col15</th>
<th>Col16</th>
<th>Col17</th>
<th>Col18</th>
<th>Col19</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.170</td>
<td>-1.300</td>
<td>-1.182</td>
<td>-1.167</td>
<td>6.9909</td>
</tr>
<tr>
<td>2</td>
<td>-0.941</td>
<td>-2.159</td>
<td>-2.105</td>
<td>-1.314</td>
<td>6.9959</td>
</tr>
<tr>
<td>3</td>
<td>1.020</td>
<td>1.801</td>
<td>1.092</td>
<td>1.007</td>
<td>0.05450</td>
</tr>
<tr>
<td>4</td>
<td>0.708</td>
<td>1.090</td>
<td>1.734</td>
<td>0.720</td>
<td>-0.04955</td>
</tr>
<tr>
<td>5</td>
<td>1.7812</td>
<td>1.0095</td>
<td>0.7197</td>
<td>1.7756</td>
<td>0.02230</td>
</tr>
<tr>
<td>6</td>
<td>1.0095</td>
<td>2.7518</td>
<td>1.6392</td>
<td>1.4849</td>
<td>0.2651</td>
</tr>
<tr>
<td>7</td>
<td>0.7197</td>
<td>1.6392</td>
<td>2.6874</td>
<td>0.9930</td>
<td>-0.2601</td>
</tr>
<tr>
<td>8</td>
<td>1.7756</td>
<td>1.4849</td>
<td>0.9930</td>
<td>2.7645</td>
<td>0.1276</td>
</tr>
</tbody>
</table>

The solutions for the fixed and random effects in Output 63.4.10 correspond to the last column in Output 63.4.9. Note that the standard errors for the fixed effects and the prediction standard errors for the random effects are the square root values of the diagonal entries in the solution of the mixed model equations (Output 63.4.9).

Output 63.4.10  Solutions for Fixed and Random Effects

Solution for Fixed Effects

<table>
<thead>
<tr>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>Trait 1</td>
</tr>
<tr>
<td>Trait 2</td>
</tr>
</tbody>
</table>

Solution for Random Effects

<table>
<thead>
<tr>
<th>Std Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>Trait*Animal 1</td>
</tr>
<tr>
<td>Trait*Animal 1</td>
</tr>
<tr>
<td>Trait*Animal 1</td>
</tr>
<tr>
<td>Trait*Animal 2</td>
</tr>
<tr>
<td>Trait*Animal 2</td>
</tr>
<tr>
<td>Trait*Animal 2</td>
</tr>
</tbody>
</table>
The estimates for the two traits are nearly identical, but the standard error of the second trait is larger because of the missing observation.

The Estimate column in the “Solution for Random Effects” table lists the best linear unbiased predictions (BLUPs) of the breeding values of both traits for all three animals. The \( p \)-values are missing because the default containment method for computing degrees of freedom results in zero degrees of freedom for the random effects parameter tests.

**Output 63.4.11** Significance Test Comparing Traits

<table>
<thead>
<tr>
<th>Type 3 Tests of Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effect</strong></td>
</tr>
<tr>
<td>Trait</td>
</tr>
</tbody>
</table>

The two estimated traits are significantly different from zero at the 5% level (Output 63.4.11).

**Output 63.4.12** Predicted Observations

<table>
<thead>
<tr>
<th>Obs</th>
<th>Trait</th>
<th>Animal</th>
<th>Y</th>
<th>Pred</th>
<th>Pred DF</th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
<th>Resid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>7.04542</td>
<td>1.33027</td>
<td>0</td>
<td></td>
<td></td>
<td>-1.04542</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>6.94137</td>
<td>1.39806</td>
<td>0</td>
<td></td>
<td></td>
<td>1.05863</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>7.01321</td>
<td>1.41129</td>
<td>0</td>
<td></td>
<td></td>
<td>-0.01321</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>7.26094</td>
<td>1.72839</td>
<td>0</td>
<td></td>
<td></td>
<td>1.73906</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>6.73576</td>
<td>1.74077</td>
<td>0</td>
<td></td>
<td></td>
<td>-1.73576</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3</td>
<td></td>
<td>7.12015</td>
<td>2.99088</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The predicted values are not the predictions of future records in the sense that they do not contain a component corresponding to a new observational error. See Henderson (1984) for information about predicting future records. The Lower and Upper columns usually contain confidence limits for the predicted values; they are missing here because the random-effects parameter degrees of freedom equals 0.

**Example 63.5: Random Coefficients**

This example comes from a pharmaceutical stability data simulation performed by Obenchain (1990). The observed responses are replicate assay results, expressed in percent of label claim, at various shelf ages, expressed in months. The desired mixed model involves three batches of product that differ randomly in intercept (initial potency) and slope (degradation rate). This type of model is also known as a hierarchical or multilevel model (Singer 1998; Sullivan, Dukes, and Losina 1999).

The SAS statements are as follows:
In the DATA step, Monthc is created as a duplicate of Month in order to enable both a continuous and a classification version of the same variable. The variable Monthc is used in a subsequent analysis.

In the PROC MIXED statements, Batch is listed as the only classification variable. The fixed effect Month in the MODEL statement is not declared as a classification variable; thus it models a linear trend in time. An intercept is included as a fixed effect by default, and the S option requests that the fixed-effects parameter estimates be produced.

The two random effects are Int and Month, modeling random intercepts and slopes, respectively. Note that Intercept and Month are used as both fixed and random effects. The TYPE=UN option in the RANDOM statement specifies an unstructured covariance matrix for the random intercept and slope effects. In mixed model notation, G is block diagonal with unstructured 2×2 blocks. Each block corresponds to a different level of Batch, which is the SUBJECT= effect. The unstructured type provides a mechanism for estimating the correlation between the random coefficients. The S option requests the production of the random-effects parameter estimates.

The results from this analysis are shown in Output 63.5.1–Output 63.5.9. The “Unstructured” covariance structure in Output 63.5.1 applies to G here.
Output 63.5.1 Model Information in Random Coefficients Analysis

The Mixed Procedure

Model Information

<table>
<thead>
<tr>
<th>Data Set</th>
<th>WORK.RC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Y</td>
</tr>
<tr>
<td>Covariance Structure</td>
<td>Unstructured</td>
</tr>
<tr>
<td>Subject Effect</td>
<td>Batch</td>
</tr>
<tr>
<td>Estimation Method</td>
<td>REML</td>
</tr>
<tr>
<td>Residual Variance Method</td>
<td>Profile</td>
</tr>
<tr>
<td>Fixed Effects SE Method</td>
<td>Model-Based</td>
</tr>
<tr>
<td>Degrees of Freedom Method</td>
<td>Containment</td>
</tr>
</tbody>
</table>

Batch is the only classification variable in this analysis, and it has three levels (Output 63.5.2).

Output 63.5.2 Random Coefficients Analysis (continued)

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch</td>
<td>3</td>
<td>1 2 3</td>
</tr>
</tbody>
</table>

The “Dimensions” table in Output 63.5.3 indicates that there are three subjects (corresponding to batches). The 24 observations not used correspond to the missing values of Y in the input data set.

Output 63.5.3 Random Coefficients Analysis (continued)

Dimensions

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariance Parameters</td>
<td>4</td>
</tr>
<tr>
<td>Columns in X</td>
<td>2</td>
</tr>
<tr>
<td>Columns in Z per Subject</td>
<td>2</td>
</tr>
<tr>
<td>Subjects</td>
<td>3</td>
</tr>
<tr>
<td>Max Obs per Subject</td>
<td>28</td>
</tr>
</tbody>
</table>

Number of Observations

| Number of Observations Read | 108 |
| Number of Observations Used | 84  |
| Number of Observations Not Used | 24  |
As Output 63.5.4 shows, only one iteration is required for convergence.

**Output 63.5.4** Random Coefficients Analysis (*continued*)

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Evaluations</th>
<th>-2 Res Log Like</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>367.02768461</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>350.32813577</td>
<td>0.00000000</td>
</tr>
</tbody>
</table>

Convergence criteria met.

The Estimate column in Output 63.5.5 lists the estimated elements of the unstructured 2×2 matrix comprising the blocks of G. Note that the random coefficients are negatively correlated.

**Output 63.5.5** Random Coefficients Analysis (*continued*)

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN(1,1)</td>
<td>Batch</td>
<td>0.9768</td>
</tr>
<tr>
<td>UN(2,1)</td>
<td>Batch</td>
<td>-0.1045</td>
</tr>
<tr>
<td>UN(2,2)</td>
<td>Batch</td>
<td>0.03717</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>3.2932</td>
</tr>
</tbody>
</table>

The null model likelihood ratio test indicates a significant improvement over the null model consisting of no random effects and a homogeneous residual error (Output 63.5.6).

**Output 63.5.6** Random Coefficients Analysis (*continued*)

<table>
<thead>
<tr>
<th>-2 Res Log Likelihood</th>
<th>AIC (Smaller is Better)</th>
<th>AICC (Smaller is Better)</th>
<th>BIC (Smaller is Better)</th>
</tr>
</thead>
<tbody>
<tr>
<td>350.3</td>
<td>358.3</td>
<td>358.8</td>
<td>354.7</td>
</tr>
</tbody>
</table>

Null Model Likelihood Ratio Test

<table>
<thead>
<tr>
<th>DF</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>16.70</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

The fixed-effects estimates represent the estimated means for the random intercept and slope, respectively (Output 63.5.7).
Output 63.5.7 Random Coefficients Analysis (continued)

| Effect   | Estimate | Error  | DF  | t Value | Pr > |t| |
|----------|----------|--------|-----|---------|-------|
| Intercept| 102.70   | 0.6456 | 2   | 159.08  | <.0001|
| Month    | -0.5259  | 0.1194 | 2   | -4.41   | 0.0478|

The random-effects estimates represent the estimated deviation from the mean intercept and slope for each batch (Output 63.5.8). Therefore, the intercept for the first batch is close to $102.7 - 1 = 101.7$, while the intercepts for the other two batches are greater than 102.7. The second batch has a slope less than the mean slope of $-0.526$, while the other two batches have slopes greater than $-0.526$.

Output 63.5.8 Random Coefficients Analysis (continued)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Batch</th>
<th>Estimate</th>
<th>Std Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>-1.0010</td>
<td>0.6842</td>
</tr>
<tr>
<td>Month</td>
<td>1</td>
<td>0.1287</td>
<td>0.1245</td>
</tr>
<tr>
<td>Intercept</td>
<td>2</td>
<td>0.3934</td>
<td>0.6842</td>
</tr>
<tr>
<td>Month</td>
<td>2</td>
<td>-0.2060</td>
<td>0.1245</td>
</tr>
<tr>
<td>Intercept</td>
<td>3</td>
<td>0.6076</td>
<td>0.6842</td>
</tr>
<tr>
<td>Month</td>
<td>3</td>
<td>0.07731</td>
<td>0.1245</td>
</tr>
</tbody>
</table>

The $F$ statistic in the “Type 3 Tests of Fixed Effects” table in Output 63.5.9 is the square of the $t$ statistic used in the test of $Month$ in the preceding “Solution for Fixed Effects” table (compare Output 63.5.7 and Output 63.5.9). Both statistics test the null hypothesis that the slope assigned to $Month$ equals 0, and this hypothesis can barely be rejected at the 5% level.

Output 63.5.9 Random Coefficients Analysis (continued)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>1</td>
<td>2</td>
<td>19.41</td>
<td>0.0478</td>
</tr>
</tbody>
</table>

It is also possible to fit a random coefficients model with error terms that follow a nested structure (Fuller and Battese 1973). The following SAS statements represent one way of doing this:
proc mixed data=rc;
  class Batch Monthc;
  model Y = Month / s;
  random Int Month Monthc / sub=Batch s;
run;

The variable Monthc is added to the CLASS and RANDOM statements, and it models the nested errors. Note that Month and Monthc are continuous and classification versions of the same variable. Also, the TYPE=UN option is dropped from the RANDOM statement, resulting in the default variance components model instead of correlated random coefficients. The results from this analysis are shown in Output 63.5.10.

**Output 63.5.10** Random Coefficients with Nested Errors Analysis

<table>
<thead>
<tr>
<th>The Mixed Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Information</td>
</tr>
<tr>
<td>Data Set</td>
</tr>
<tr>
<td>Dependent Variable</td>
</tr>
<tr>
<td>Covariance Structure</td>
</tr>
<tr>
<td>Subject Effect</td>
</tr>
<tr>
<td>Estimation Method</td>
</tr>
<tr>
<td>Residual Variance Method</td>
</tr>
<tr>
<td>Fixed Effects SE Method</td>
</tr>
<tr>
<td>Degrees of Freedom Method</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class Level Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
</tr>
<tr>
<td>Batch</td>
</tr>
<tr>
<td>Monthc</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariance Parameters</td>
</tr>
<tr>
<td>Columns in X</td>
</tr>
<tr>
<td>Columns in Z per Subject</td>
</tr>
<tr>
<td>Subjects</td>
</tr>
<tr>
<td>Max Obs per Subject</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations Read</td>
</tr>
<tr>
<td>Number of Observations Used</td>
</tr>
<tr>
<td>Number of Observations Not Used</td>
</tr>
</tbody>
</table>
Output 63.5.10 continued

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Evaluations</th>
<th>(-2) Res Log Like</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>367.02768461</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>277.51945360</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>276.97551718</td>
<td>0.00104208</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>276.90304909</td>
<td>0.00003174</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>276.90100316</td>
<td>0.00000004</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>276.90100092</td>
<td>0.00000000</td>
</tr>
</tbody>
</table>

Convergence criteria met.

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Batch</td>
<td>0</td>
</tr>
<tr>
<td>Month</td>
<td>Batch</td>
<td>0.01243</td>
</tr>
<tr>
<td>Monthc</td>
<td>Batch</td>
<td>3.7411</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.7969</td>
</tr>
</tbody>
</table>

For this analysis, the Newton-Raphson algorithm requires five iterations and nine likelihood evaluations to achieve convergence. The missing value in the Criterion column in iteration 1 indicates that a boundary constraint has been dropped.

The estimate for the Intercept variance component equals 0. This occurs frequently in practice and indicates that the restricted likelihood is maximized by setting this variance component equal to 0. Whenever a zero variance component estimate occurs, the following note appears in the SAS log:

NOTE: Estimated G matrix is not positive definite.

The remaining variance component estimates are positive, and the estimate corresponding to the nested errors (MONTHC) is much larger than the other two.

A comparison of AIC and BIC for this model with those of the previous model favors the nested error model (compare Output 63.5.11 and Output 63.5.6). Strictly speaking, a likelihood ratio test cannot be carried out between the two models because one is not contained in the other; however, a cautious comparison of likelihoods can be informative.

Output 63.5.11 Random Coefficients with Nested Errors Analysis (continued)
The better-fitting covariance model affects the standard errors of the fixed-effects parameter estimates more than the estimates themselves (Output 63.5.12).

**Output 63.5.12** Random Coefficients with Nested Errors Analysis (continued)

| Effect   | Estimate | Error  | DF | t Value | Pr > |t| |
|----------|----------|--------|----|---------|-------|---|
| Intercept| 102.56   | 0.7287 | 2  | 140.74  | <.0001|
| Month    | -0.5003  | 0.1259 | 2  | -3.97   | 0.0579|

The random-effects solution provides the empirical best linear unbiased predictions (EBLUPs) for the realizations of the random intercept, slope, and nested errors (Output 63.5.13). You can use these values to compare batches and months.

**Output 63.5.13** Random Coefficients with Nested Errors Analysis (continued)

| Effect   | Batch | Monthc | Estimate | Pred | DF | t Value | Pr > |t| |
|----------|-------|--------|----------|------|----|---------|-------|---|
| Intercept| 1     | 0      | .        | .    | .  | .       | .     |   |
| Month    | 1     | 0      | -0.00028 | 0.09268 | 66 | -0.00 | 0.9976|
| Monthc   | 1     | 1      | -2.5690  | 0.7571 | 66 | -3.39 | 0.0012|
| Monthc   | 1     | 3      | -2.3067  | 0.6865 | 66 | -3.36 | 0.0013|
| Monthc   | 1     | 6      | 1.8726   | 0.7328 | 66 | 2.56  | 0.0129|
| Monthc   | 1     | 9      | -1.2350  | 0.9300 | 66 | -1.33 | 0.1888|
| Monthc   | 1     | 12     | 0.7736   | 1.1992 | 66 | 0.65  | 0.5211|
| Intercept| 2     | 0      | .        | .    | .  | .       | .     |   |
| Month    | 2     | 0      | -0.07571 | 0.09268 | 66 | -0.82 | 0.4169|
| Monthc   | 2     | 1      | -2.2126  | 0.7571 | 66 | -2.92 | 0.0048|
| Monthc   | 2     | 3      | -3.1063  | 0.6865 | 66 | 4.53  | <.0001|
| Monthc   | 2     | 6      | 2.0649   | 0.7328 | 66 | 2.82  | 0.0064|
| Monthc   | 2     | 9      | -1.4450  | 0.9300 | 66 | -1.55 | 0.1250|
| Monthc   | 2     | 12     | -2.4405  | 1.1992 | 66 | -2.04 | 0.0459|
| Intercept| 3     | 0      | .        | .    | .  | .       | .     |   |
| Month    | 3     | 0      | 0.07600  | 0.09268 | 66 | 0.82  | 0.4152|
| Monthc   | 3     | 1      | 1.9574   | 0.7896 | 66 | 2.48  | 0.0157|
| Monthc   | 3     | 3      | -0.8850  | 0.7571 | 66 | -1.17 | 0.2466|
| Monthc   | 3     | 6      | 0.3006   | 0.6865 | 66 | 0.44  | 0.6629|
| Monthc   | 3     | 9      | 0.7972   | 0.7328 | 66 | 1.09  | 0.2806|
| Monthc   | 3     | 12     | 2.0059   | 0.9300 | 66 | 2.16  | 0.0347|
| Monthc   | 3     | 12     | 0.002293 | 1.1992 | 66 | 0.00  | 0.9985|
Output 63.5.14 Random Coefficients with Nested Errors Analysis (continued)

<table>
<thead>
<tr>
<th>Type 3 Tests of Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>Month</td>
</tr>
</tbody>
</table>

The test of Month is similar to that from the previous model, although it is no longer significant at the 5% level (Output 63.5.14).

Example 63.6: Line-Source Sprinkler Irrigation

These data appear in Hanks et al. (1980); Johnson, Chaudhuri, and Kanemasu (1983); Stroup (1989b). Three cultivars (Cult) of winter wheat are randomly assigned to rectangular plots within each of three blocks (Block). The nine plots are located side by side, and a line-source sprinkler is placed through the middle. Each plot is subdivided into twelve subplots—six to the north of the line source, six to the south (Dir). The two plots closest to the line source represent the maximum irrigation level (Irrig=6), the two next-closest plots represent the next-highest level (Irrig=5), and so forth.

This example is a case where both G and R can be modeled. One of Stroup’s models specifies a diagonal G containing the variance components for Block, Block*Dir, and Block*Irrig, and a Toeplitz R with four bands. The SAS statements to fit this model and carry out some further analyses follow.

CAUTION: This analysis can require considerable CPU time.

data line;
  length Cult$ 8;
  input Block Cult$ @;
  row = _n_;
  do Sbplt=1 to 12;
    if Sbplt le 6 then do;
      Irrig = Sbplt;
      Dir = 'North';
    end; else do;
      Irrig = 13 - Sbplt;
      Dir = 'South';
    end;
    input Y @; output;
  end;
datalines;
1 Luke  2.4  2.7  5.6  7.5  7.9  7.1  6.1  7.3  7.4  6.7  3.8  1.8
1 Nugaines  2.2  2.2  4.3  6.3  7.9  7.1  6.2  5.3  5.3  5.2  5.4  2.9
1 Bridger  2.9  3.2  5.1  6.9  6.1  7.5  5.6  6.5  6.6  5.3  4.1  3.1
2 Nugaines  2.4  2.2  4.0  5.8  6.1  6.2  7.0  6.4  6.7  6.4  3.7  2.2
2 Bridger  2.6  3.1  5.7  6.4  7.7  6.8  6.3  6.2  6.6  6.5  4.2  2.7
2 Luke  2.2  2.7  4.3  6.9  6.8  8.0  6.5  7.3  5.9  6.6  3.0  2.0
3 Nugaines  1.8  1.9  3.7  4.9  5.4  5.1  5.7  5.0  5.6  5.1  4.2  2.2
3 Luke  2.1  2.3  3.7  5.8  6.3  6.3  6.5  5.7  5.8  4.5  2.7  2.3
The levels of each classification variable are listed as a single string in the Values column, regardless of whether the levels are numeric or character (Output 63.6.2).
Even though there is a \texttt{SUBJECT=} effect in the \texttt{REPEATED} statement, the analysis considers all of the data to be from one subject because there is no corresponding \texttt{SUBJECT=} effect in the \texttt{RANDOM} statement (Output 63.6.3).

The Newton-Raphson algorithm converges successfully in seven iterations (Output 63.6.4).
The first block of the estimated \( \mathbf{R} \) matrix has the TOEP(4) structure, and the observations that are three plots apart exhibit a negative correlation (Output 63.6.5).

**Output 63.6.5**  Estimated \( \mathbf{R} \) Matrix for the First Subject

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
<th>Col5</th>
<th>Col6</th>
<th>Col7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2850</td>
<td>0.007986</td>
<td>0.001452</td>
<td>-0.09253</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.007986</td>
<td>0.2850</td>
<td>0.007986</td>
<td>0.001452</td>
<td>-0.09253</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.001452</td>
<td>0.007986</td>
<td>0.2850</td>
<td>0.007986</td>
<td>0.001452</td>
<td>-0.09253</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.09253</td>
<td>0.001452</td>
<td>0.007986</td>
<td>0.2850</td>
<td>0.007986</td>
<td>0.001452</td>
<td>-0.09253</td>
</tr>
<tr>
<td>5</td>
<td>-0.09253</td>
<td>0.001452</td>
<td>0.007986</td>
<td>0.2850</td>
<td>0.007986</td>
<td>0.001452</td>
<td>-0.09253</td>
</tr>
<tr>
<td>6</td>
<td>-0.09253</td>
<td>0.001452</td>
<td>0.007986</td>
<td>0.2850</td>
<td>0.007986</td>
<td>0.001452</td>
<td>-0.09253</td>
</tr>
<tr>
<td>7</td>
<td>-0.09253</td>
<td>0.001452</td>
<td>0.007986</td>
<td>0.2850</td>
<td>0.007986</td>
<td>0.001452</td>
<td>-0.09253</td>
</tr>
<tr>
<td>8</td>
<td>-0.09253</td>
<td>0.001452</td>
<td>0.007986</td>
<td>0.2850</td>
<td>0.007986</td>
<td>0.001452</td>
<td>-0.09253</td>
</tr>
</tbody>
</table>

Output 63.6.6 lists the estimated covariance parameters from both \( \mathbf{G} \) and \( \mathbf{R} \). The first three are the variance components making up the diagonal \( \mathbf{G} \), and the final four make up the Toeplitz structure in the blocks of \( \mathbf{R} \). The Residual row corresponds to the variance of the Toeplitz structure, and it represents the parameter profiled out during the optimization process.
Output 63.6.6  Estimated Covariance Parameters

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td></td>
<td>0.2194</td>
</tr>
<tr>
<td>Block*Dir</td>
<td></td>
<td>0.01768</td>
</tr>
<tr>
<td>Block*Irrig</td>
<td></td>
<td>0.03539</td>
</tr>
<tr>
<td>TOEP (2)</td>
<td>Block*Cult</td>
<td>0.007986</td>
</tr>
<tr>
<td>TOEP (3)</td>
<td>Block*Cult</td>
<td>0.001452</td>
</tr>
<tr>
<td>TOEP (4)</td>
<td>Block*Cult</td>
<td>-0.09253</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.2850</td>
</tr>
</tbody>
</table>

The “–2 Res Log Likelihood” value in Output 63.6.7 is the same as the final value listed in the “Iteration History” table (Output 63.6.4).

Output 63.6.7  Fit Statistics Based on the Residual Log Likelihood

<table>
<thead>
<tr>
<th>Fit Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>–2 Res Log Likelihood</td>
</tr>
<tr>
<td>AIC (Smaller is Better)</td>
</tr>
<tr>
<td>AICC (Smaller is Better)</td>
</tr>
<tr>
<td>BIC (Smaller is Better)</td>
</tr>
</tbody>
</table>

Every fixed effect except for Dir and Cult*Irrig is significant at the 5% level (Output 63.6.8).

Output 63.6.8  Tests for Fixed Effects

<table>
<thead>
<tr>
<th>Type 3 Tests of Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>Cult</td>
</tr>
<tr>
<td>Dir</td>
</tr>
<tr>
<td>Cult*Dir</td>
</tr>
<tr>
<td>Irrig</td>
</tr>
<tr>
<td>Cult*Irrig</td>
</tr>
<tr>
<td>Dir*Irrig</td>
</tr>
</tbody>
</table>

The “Estimates” table lists the results from the various linear combinations of fixed effects specified in the ESTIMATE statements (Output 63.6.9). Bridger is not significantly different from Luke, and Irrig possesses a strong linear component. This strength appears to be influencing the significance of the interaction.
Example 63.6: Line-Source Sprinkler Irrigation

Output 63.6.9 Estimates

| Label                     | Estimate | Standard Error | DF | t Value | Pr > |t| |
|---------------------------|----------|----------------|----|---------|-------|
| Bridger vs Luke           | -0.03889 | 0.09524        | 68 | -0.41   | 0.6843|
| Linear Irrig              | 30.6444  | 1.4412         | 10 | 21.26   | <.0001|
| B vs L x Linear Irrig     | -9.8667  | 2.7400         | 68 | -3.60   | 0.0006|

The least squares means shown in Output 63.6.10 are useful in comparing the levels of the various fixed effects. For example, it appears that irrigation levels 5 and 6 have virtually the same effect.

Output 63.6.10 Least Squares Means for Cult, Irrig, and Their Interaction

| Effect     | Cult | Irrig | Estimate | Standard Error | DF | t Value | Pr > |t| |
|------------|------|-------|----------|----------------|----|---------|-------|
| Cult       | Bridger | 5.0306 | 0.2874 | 68 | 17.51 | <.0001|
| Cult       | Luke   | 5.0694 | 0.2874 | 68 | 17.64 | <.0001|
| Cult       | Nugaines | 4.7222 | 0.2874 | 68 | 16.43 | <.0001|
| Irrig      | 1      | 2.4222 | 0.3220 | 10 | 7.52  | <.0001|
| Irrig      | 2      | 3.1833 | 0.3220 | 10 | 9.88  | <.0001|
| Irrig      | 3      | 5.0556 | 0.3220 | 10 | 15.70 | <.0001|
| Irrig      | 4      | 6.1889 | 0.3220 | 10 | 19.22 | <.0001|
| Irrig      | 5      | 6.4000 | 0.3140 | 10 | 20.38 | <.0001|
| Irrig      | 6      | 6.3944 | 0.3227 | 10 | 19.81 | <.0001|
| Cult*Irrig | Bridger 1 | 2.8500 | 0.3679 | 68 | 7.75  | <.0001|
| Cult*Irrig | Bridger 2 | 3.4167 | 0.3679 | 68 | 9.29  | <.0001|
| Cult*Irrig | Bridger 3 | 5.1500 | 0.3679 | 68 | 14.00 | <.0001|
| Cult*Irrig | Bridger 4 | 6.2500 | 0.3679 | 68 | 16.99 | <.0001|
| Cult*Irrig | Bridger 5 | 6.3000 | 0.3463 | 68 | 18.19 | <.0001|
| Cult*Irrig | Bridger 6 | 6.2167 | 0.3697 | 68 | 18.81 | <.0001|
| Cult*Irrig | Luke 1   | 2.1333 | 0.3679 | 68 | 5.80  | <.0001|
| Cult*Irrig | Luke 2   | 2.8667 | 0.3679 | 68 | 7.79  | <.0001|
| Cult*Irrig | Luke 3   | 5.2333 | 0.3679 | 68 | 14.22 | <.0001|
| Cult*Irrig | Luke 4   | 6.5500 | 0.3679 | 68 | 17.80 | <.0001|
| Cult*Irrig | Luke 5   | 6.8833 | 0.3463 | 68 | 19.87 | <.0001|
| Cult*Irrig | Luke 6   | 6.7500 | 0.3679 | 68 | 18.26 | <.0001|
| Cult*Irrig | Nugaines 1 | 2.2833 | 0.3679 | 68 | 6.21  | <.0001|
| Cult*Irrig | Nugaines 2 | 3.2667 | 0.3679 | 68 | 8.88  | <.0001|
| Cult*Irrig | Nugaines 3 | 4.7833 | 0.3679 | 68 | 13.00 | <.0001|
| Cult*Irrig | Nugaines 4 | 5.7667 | 0.3679 | 68 | 15.67 | <.0001|
| Cult*Irrig | Nugaines 5 | 6.0167 | 0.3463 | 68 | 17.37 | <.0001|
| Cult*Irrig | Nugaines 6 | 6.2167 | 0.3679 | 68 | 16.81 | <.0001|

An interesting exercise is to fit other variance-covariance models to these data and to compare them to this one by using likelihood ratio tests, Akaike’s information criterion, or Schwarz’s Bayesian information criterion. In particular, some spatial models are worth investigating (Marx and Thompson 1987; Zimmerman...
and Harville 1991). The following is one example of spatial model statements:

```sas
proc mixed;
    class Block Cult Dir Irrig;
    model Y = Cult|Dir|Irrig@2;
    repeated / type=sp(pow)(Row Sbplt) sub=intercept;
run;
```

The `TYPE=SP(POW)(Row Sbplt)` option in the `REPEATED` statement requests the spatial power structure, with the two defining coordinate variables being `Row` and `Sbplt`. The `SUBJECT=INTERCEPT` option indicates that the entire data set is to be considered as one subject, thereby modeling $\mathbf{R}$ as a dense $108 \times 108$ covariance matrix. See Wolfinger (1993) for further discussion of this example and additional analyses.

---

**Example 63.7: Influence in Heterogeneous Variance Model**

In this example from Snedecor and Cochran (1980, p. 216), a one-way classification model with heterogeneous variances is fit. The data, shown in the following DATA step, represent amounts of different types of fat absorbed by batches of doughnuts during cooking, measured in grams.

```sas
data absorb;
    input FatType Absorbed @@;
datalines;
1 164 1 172 1 168 1 177 1 156 1 195
2 178 2 191 2 197 2 182 2 185 2 177
3 175 3 193 3 178 3 171 3 163 3 176
4 155 4 166 4 149 4 164 4 170 4 168;
```

The statistical model for these data can be written as

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

where $Y_{ij}$ is the amount of fat absorbed by the $j$th batch of the $i$th fat type, and $\tau_i$ denotes the fat-type effects. A quick glance at the data suggests that observations 6, 9, 14, and 21 might be influential on the analysis, because these are extreme observations for the respective fat types.

The following SAS statements fit this model and request influence diagnostics for the fixed effects and covariance parameters. ODS Graphics is used to create plots of the influence diagnostics in addition to the tabular output. The `ESTIMATES` suboption requests plots of “leave-one-out” estimates for the fixed effects and group variances.

```sas
ods graphics on;

proc mixed data=absorb asycov;
    class FatType;
    model Absorbed = FatType / s
        influence(iter=10 estimates);
    repeated / group=FatType;
```
ods output Influence=inf;
run;

ods graphics off;
The “Influence” table is output to the SAS data set inf so that parameter estimates can be printed subsequently. Results from this analysis are shown in Output 63.7.1.

**Output 63.7.1 Heterogeneous Variance Analysis**

<table>
<thead>
<tr>
<th>Data Set</th>
<th>WORK.ABSORB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Absorbed</td>
</tr>
<tr>
<td>Covariance Structure</td>
<td>Variance Components</td>
</tr>
<tr>
<td>Group Effect</td>
<td>FatType</td>
</tr>
<tr>
<td>Estimation Method</td>
<td>REML</td>
</tr>
<tr>
<td>Residual Variance Method</td>
<td>None</td>
</tr>
<tr>
<td>Fixed Effects SE Method</td>
<td>Model-Based</td>
</tr>
<tr>
<td>Degrees of Freedom Method</td>
<td>Between-Within</td>
</tr>
</tbody>
</table>

**Covariance Parameter Estimates**

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Group</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual</td>
<td>FatType 1</td>
<td>178.00</td>
</tr>
<tr>
<td>Residual</td>
<td>FatType 2</td>
<td>60.4000</td>
</tr>
<tr>
<td>Residual</td>
<td>FatType 3</td>
<td>97.6000</td>
</tr>
<tr>
<td>Residual</td>
<td>FatType 4</td>
<td>67.6000</td>
</tr>
</tbody>
</table>

**Solution for Fixed Effects**

| Effect | Fat Type | Estimate | Standard Error | DF  | t Value | Pr > |t| |
|--------|----------|----------|----------------|-----|---------|------|---|
| Intercept |         | 162.00   | 3.3566         | 20  | 48.26   | <.0001 |
| FatType 1 |        | 10.0000  | 6.3979         | 20  | 1.56    | 0.1337 |
| FatType 2 |        | 23.0000  | 4.6188         | 20  | 4.98    | <.0001 |
| FatType 3 |        | 14.0000  | 5.2472         | 20  | 2.67    | 0.0148 |
| FatType 4 |        | 0        | .              | .   | .       | .     |

The fixed-effects solutions correspond to estimates of the following parameters:

- Intercept : $\mu + \tau_4$
- FatType1 : $\tau_1 - \tau_4$
- FatType2 : $\tau_2 - \tau_4$
- FatType3 : $\tau_3 - \tau_4$
- FatType4 : 0
You can easily verify that these estimates are simple functions of the arithmetic means \( \bar{y}_i \) in the groups. For example, \( \mu + \tau_4 = \bar{y}_4 = 162.0 \), \( \tau_1 - \tau_4 = \bar{y}_1 - \bar{y}_4 = 10.0 \), and so forth. The covariance parameter estimates are the sample variances in the groups and are uncorrelated.

The variances in the four groups are shown in the “Covariance Parameter Estimates” table (Output 63.7.1). The estimated variance in the first group is two to three times larger than the variance in the other groups.

**Output 63.7.2** Asymptotic Variances of Group Variance Estimates

<table>
<thead>
<tr>
<th>Row</th>
<th>Cov Parm</th>
<th>CovP1</th>
<th>CovP2</th>
<th>CovP3</th>
<th>CovP4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Residual</td>
<td>12674</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Residual</td>
<td></td>
<td>1459.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Residual</td>
<td></td>
<td></td>
<td>3810.30</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Residual</td>
<td></td>
<td></td>
<td></td>
<td>1827.90</td>
</tr>
</tbody>
</table>

In groups where the residual variance estimate is large, the precision of the estimate is also small (Output 63.7.2).

The following statements print the “leave-one-out” estimates for fixed effects and covariance parameters that were written to the inf data set with the ESTIMATES suboption (Output 63.7.3):

```plaintext
proc print data=inf label;
  var parm1-parm5 covp1-covp4;
run;
```
### Output 63.7.3 Leave-One-Out Estimates

<table>
<thead>
<tr>
<th>Obs</th>
<th>Intercept</th>
<th>Fat Type 1</th>
<th>Fat Type 2</th>
<th>Fat Type 3</th>
<th>Fat Type 4</th>
<th>Residual FatType 1</th>
<th>Residual FatType 2</th>
<th>Residual FatType 3</th>
<th>Residual FatType 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>162.00</td>
<td>11.600</td>
<td>23.000</td>
<td>14.000</td>
<td>0</td>
<td>203.30</td>
<td>60.400</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>2</td>
<td>162.00</td>
<td>10.000</td>
<td>23.000</td>
<td>14.000</td>
<td>0</td>
<td>222.47</td>
<td>60.400</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>3</td>
<td>162.00</td>
<td>10.800</td>
<td>23.000</td>
<td>14.000</td>
<td>0</td>
<td>217.68</td>
<td>60.400</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>4</td>
<td>162.00</td>
<td>9.000</td>
<td>23.000</td>
<td>14.000</td>
<td>0</td>
<td>214.99</td>
<td>60.400</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>5</td>
<td>162.00</td>
<td>13.200</td>
<td>23.000</td>
<td>14.000</td>
<td>0</td>
<td>145.70</td>
<td>60.400</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>6</td>
<td>162.00</td>
<td>5.400</td>
<td>23.000</td>
<td>14.000</td>
<td>0</td>
<td>63.80</td>
<td>60.400</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>7</td>
<td>162.00</td>
<td>10.000</td>
<td>24.400</td>
<td>14.000</td>
<td>0</td>
<td>178.00</td>
<td>60.795</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>8</td>
<td>162.00</td>
<td>10.000</td>
<td>21.800</td>
<td>14.000</td>
<td>0</td>
<td>178.00</td>
<td>64.691</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>9</td>
<td>162.00</td>
<td>10.000</td>
<td>20.600</td>
<td>14.000</td>
<td>0</td>
<td>178.00</td>
<td>32.296</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>10</td>
<td>162.00</td>
<td>10.000</td>
<td>23.600</td>
<td>14.000</td>
<td>0</td>
<td>178.00</td>
<td>72.797</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>11</td>
<td>162.00</td>
<td>10.000</td>
<td>23.000</td>
<td>14.000</td>
<td>0</td>
<td>178.00</td>
<td>75.490</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>12</td>
<td>162.00</td>
<td>10.000</td>
<td>24.600</td>
<td>14.000</td>
<td>0</td>
<td>178.00</td>
<td>56.285</td>
<td>97.60</td>
<td>67.60</td>
</tr>
<tr>
<td>13</td>
<td>162.00</td>
<td>10.000</td>
<td>23.000</td>
<td>14.200</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>121.68</td>
<td>67.60</td>
</tr>
<tr>
<td>14</td>
<td>162.00</td>
<td>10.000</td>
<td>23.000</td>
<td>10.600</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>35.30</td>
<td>67.60</td>
</tr>
<tr>
<td>15</td>
<td>162.00</td>
<td>10.000</td>
<td>23.000</td>
<td>13.600</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>120.79</td>
<td>67.60</td>
</tr>
<tr>
<td>16</td>
<td>162.00</td>
<td>10.000</td>
<td>23.000</td>
<td>15.000</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>114.50</td>
<td>67.60</td>
</tr>
<tr>
<td>17</td>
<td>162.00</td>
<td>10.000</td>
<td>23.000</td>
<td>16.600</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>71.30</td>
<td>67.60</td>
</tr>
<tr>
<td>18</td>
<td>162.00</td>
<td>10.000</td>
<td>23.000</td>
<td>14.000</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>121.98</td>
<td>67.60</td>
</tr>
<tr>
<td>19</td>
<td>163.40</td>
<td>8.600</td>
<td>21.600</td>
<td>12.600</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>97.60</td>
<td>69.799</td>
</tr>
<tr>
<td>20</td>
<td>161.20</td>
<td>10.800</td>
<td>23.800</td>
<td>14.800</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>97.60</td>
<td>79.698</td>
</tr>
<tr>
<td>21</td>
<td>164.60</td>
<td>7.400</td>
<td>20.400</td>
<td>11.400</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>97.60</td>
<td>33.800</td>
</tr>
<tr>
<td>22</td>
<td>161.60</td>
<td>10.400</td>
<td>23.400</td>
<td>14.400</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>97.60</td>
<td>83.292</td>
</tr>
<tr>
<td>23</td>
<td>160.40</td>
<td>11.600</td>
<td>24.600</td>
<td>15.600</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>97.60</td>
<td>65.299</td>
</tr>
<tr>
<td>24</td>
<td>160.80</td>
<td>11.200</td>
<td>24.200</td>
<td>15.200</td>
<td>0</td>
<td>178.00</td>
<td>60.400</td>
<td>97.60</td>
<td>73.677</td>
</tr>
</tbody>
</table>

The graphical displays in Output 63.7.4 and Output 63.7.5 are created when ODS Graphics is enabled. For general information about ODS Graphics, see Chapter 21, “Statistical Graphics Using ODS.” For specific information about the graphics available in the MIXED procedure, see the section “ODS Graphics” on page 5346.
Output 63.7.4 Fixed-Effects Deletion Estimates

Fixed Effects Deletion Estimates for Absorbed

- Intercept
- FatType 1
- FatType 2
- FatType 3

Deleted Observation
The estimate of the intercept is affected only when observations from the last group are removed. The estimate of the “FatType 1” effect reacts to removal of observations in the first and last group (Output 63.7.4).

While observations can affect one or more fixed-effects solutions in this model, they can affect only one covariance parameter, the variance in their group (Output 63.7.5). Observations 6, 9, 14, and 21, which are extreme in their group, reduce the group variance considerably.

Diagnostics related to residuals and predicted values are printed with the following statements:

```
proc print data=inf label;
   var observed predicted residual pressres
       student Rstudent;
run;
```
Observations 6, 9, 14, and 21 have large studentized residuals (Output 63.7.6). That the externally studentized residuals are much larger than the internally studentized residuals for these observations indicates that the variance estimate in the group shrinks when the observation is removed. Also important to note is that comparisons based on raw residuals in models with heterogeneous variance can be misleading. Observation 5, for example, has a larger residual but a smaller studentized residual than observation 21. The variance for the first fat type is much larger than the variance in the fourth group. A “large” residual is more “surprising” in the groups with small variance.

A measure of the overall influence on the analysis is the (restricted) likelihood distance, shown in Output 63.7.7. Observations 6, 9, 14, and 21 clearly displace the REML solution more than any other observations.
The following statements list the restricted likelihood distance and various diagnostics related to the fixed-effects estimates (Output 63.7.8):

```plaintext
proc print data=inf label;
   var leverage observed CookD DFFITS CovRatio RLD;
run;
```
Output 63.7.8  Restricted Likelihood Distance and Fixed-Effects Diagnostics

<table>
<thead>
<tr>
<th>Obs</th>
<th>Leverage</th>
<th>Observed Value</th>
<th>Cook's D</th>
<th>DFFITS</th>
<th>COVRATIO</th>
<th>Likelihood Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.167</td>
<td>164</td>
<td>0.02157</td>
<td>-0.27487</td>
<td>1.3706</td>
<td>0.1178</td>
</tr>
<tr>
<td>2</td>
<td>0.167</td>
<td>172</td>
<td>0.00000</td>
<td>-0.00000</td>
<td>1.4998</td>
<td>0.1156</td>
</tr>
<tr>
<td>3</td>
<td>0.167</td>
<td>168</td>
<td>0.00539</td>
<td>-0.13282</td>
<td>1.4675</td>
<td>0.1124</td>
</tr>
<tr>
<td>4</td>
<td>0.167</td>
<td>177</td>
<td>0.00843</td>
<td>0.16706</td>
<td>1.4494</td>
<td>0.1117</td>
</tr>
<tr>
<td>5</td>
<td>0.167</td>
<td>156</td>
<td>0.08629</td>
<td>-0.64938</td>
<td>0.9822</td>
<td>0.5290</td>
</tr>
<tr>
<td>6</td>
<td>0.167</td>
<td>195</td>
<td>0.17831</td>
<td>1.41069</td>
<td>0.4301</td>
<td>5.8101</td>
</tr>
<tr>
<td>7</td>
<td>0.167</td>
<td>178</td>
<td>0.04868</td>
<td>-0.43982</td>
<td>1.2078</td>
<td>0.1935</td>
</tr>
<tr>
<td>8</td>
<td>0.167</td>
<td>191</td>
<td>0.03576</td>
<td>0.36546</td>
<td>1.2853</td>
<td>0.1451</td>
</tr>
<tr>
<td>9</td>
<td>0.167</td>
<td>197</td>
<td>0.14305</td>
<td>1.03446</td>
<td>0.6416</td>
<td>2.2909</td>
</tr>
<tr>
<td>10</td>
<td>0.167</td>
<td>182</td>
<td>0.00894</td>
<td>-0.17225</td>
<td>1.4663</td>
<td>0.1116</td>
</tr>
<tr>
<td>11</td>
<td>0.167</td>
<td>185</td>
<td>0.00000</td>
<td>-0.00000</td>
<td>1.4998</td>
<td>0.1156</td>
</tr>
<tr>
<td>12</td>
<td>0.167</td>
<td>177</td>
<td>0.06358</td>
<td>-0.52239</td>
<td>1.1183</td>
<td>0.2856</td>
</tr>
<tr>
<td>13</td>
<td>0.167</td>
<td>175</td>
<td>0.00061</td>
<td>-0.04441</td>
<td>1.4961</td>
<td>0.1151</td>
</tr>
<tr>
<td>14</td>
<td>0.167</td>
<td>193</td>
<td>0.17766</td>
<td>1.40175</td>
<td>0.4340</td>
<td>5.7044</td>
</tr>
<tr>
<td>15</td>
<td>0.167</td>
<td>178</td>
<td>0.00246</td>
<td>0.08915</td>
<td>1.4851</td>
<td>0.1139</td>
</tr>
<tr>
<td>16</td>
<td>0.167</td>
<td>171</td>
<td>0.01537</td>
<td>-0.22892</td>
<td>1.4078</td>
<td>0.1129</td>
</tr>
<tr>
<td>17</td>
<td>0.167</td>
<td>163</td>
<td>0.10389</td>
<td>-0.75423</td>
<td>0.8766</td>
<td>0.8433</td>
</tr>
<tr>
<td>18</td>
<td>0.167</td>
<td>176</td>
<td>0.00000</td>
<td>0.00000</td>
<td>1.4998</td>
<td>0.1156</td>
</tr>
<tr>
<td>19</td>
<td>0.167</td>
<td>155</td>
<td>0.04349</td>
<td>-0.41047</td>
<td>1.2390</td>
<td>0.1710</td>
</tr>
<tr>
<td>20</td>
<td>0.167</td>
<td>166</td>
<td>0.01420</td>
<td>0.21950</td>
<td>1.4148</td>
<td>0.1124</td>
</tr>
<tr>
<td>21</td>
<td>0.167</td>
<td>149</td>
<td>0.15000</td>
<td>-1.09545</td>
<td>0.6000</td>
<td>2.7343</td>
</tr>
<tr>
<td>22</td>
<td>0.167</td>
<td>164</td>
<td>0.00355</td>
<td>0.10736</td>
<td>1.4786</td>
<td>0.1133</td>
</tr>
<tr>
<td>23</td>
<td>0.167</td>
<td>170</td>
<td>0.05680</td>
<td>0.48500</td>
<td>1.1592</td>
<td>0.2383</td>
</tr>
<tr>
<td>24</td>
<td>0.167</td>
<td>168</td>
<td>0.03195</td>
<td>0.34245</td>
<td>1.3079</td>
<td>0.1353</td>
</tr>
</tbody>
</table>

In this example, observations with large likelihood distances also have large values for Cook’s $D$ and values of CovRatio far less than one (Output 63.7.8). The latter indicates that the fixed effects are estimated more precisely when these observations are removed from the analysis.

The following statements print the values of the $D$ statistic and the CovRatio for the covariance parameters:

```latex
proc print data=inf label;
  var iter CookDCP CovRatioCP;
run;
```

The same conclusions as for the fixed-effects estimates hold for the covariance parameter estimates. Observations 6, 9, 14, and 21 change the estimates and their precision considerably (Output 63.7.9, Output 63.7.10). All iterative updates converged within at most four iterations.
Output 63.7.9 Covariance Parameter Diagnostics

<table>
<thead>
<tr>
<th>Obs</th>
<th>Iterations</th>
<th>Cook's D</th>
<th>COVRATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>0.05050</td>
<td>1.6306</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.15603</td>
<td>1.9520</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.12426</td>
<td>1.8692</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.10796</td>
<td>1.8233</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.08232</td>
<td>0.8375</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>1.02909</td>
<td>0.1606</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.00011</td>
<td>1.2662</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>0.01262</td>
<td>1.4335</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0.54126</td>
<td>0.3573</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>0.10531</td>
<td>1.8156</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>0.15603</td>
<td>1.9520</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>0.01160</td>
<td>1.0849</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>0.15223</td>
<td>1.9425</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>1.01865</td>
<td>0.1635</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>0.14111</td>
<td>1.9141</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>0.07494</td>
<td>1.7203</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>0.18154</td>
<td>0.6671</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>0.15603</td>
<td>1.9520</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>0.00265</td>
<td>1.3326</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>0.08008</td>
<td>1.7374</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>0.62500</td>
<td>0.3125</td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>0.13472</td>
<td>1.8974</td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>0.00290</td>
<td>1.1663</td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>0.02020</td>
<td>1.4839</td>
</tr>
</tbody>
</table>

Output 63.7.10 displays the standard panel of influence diagnostics that is obtained when influence analysis is iterative. The Cook’s D and CovRatio statistics are displayed for each deletion set for both fixed-effects and covariance parameter estimates. This provides a convenient summary of the impact on the analysis for each deletion set, since Cook’s D statistic measures impact on the estimates and the CovRatio statistic measures impact on the precision of the estimates.
Observations 6, 9, 14, and 21 have considerable impact on estimates and precision of fixed effects and covariance parameters. This is not necessarily the case. Observations can be influential on only some aspects of the analysis, as shown in the next example.

**Example 63.8: Influence Analysis for Repeated Measures Data**

This example revisits the repeated measures data of Pothoff and Roy (1964) that were analyzed in Example 63.2. Recall that the data consist of growth measurements at ages 8, 10, 12, and 14 for 11 girls and 16 boys. The model being fit contains fixed effects for Gender and Age and their interaction.

The earlier analysis of these data indicated some unusual observations in this data set. Because of the clustered data structure, it is of interest to study the influence of clusters (children) on the analysis rather than the influence of individual observations. A cluster comprises the repeated measurements for each child.

The repeated measures are first modeled with an unstructured within-child variance-covariance matrix. A residual variance is not profiled in this model. A noniterative influence analysis will update the fixed effects only. The following statements request this noniterative maximum likelihood analysis and produce Output 63.8.1:
Example 63.8: Influence Analysis for Repeated Measures Data

proc mixed data=pr method=ml;
  class person gender;
  model y = gender age gender*age /
    influence(effect=person);
  repeated / type=un subject=person;
  ods select influence;
run;

Output 63.8.1 Default Influence Statistics in Non iterative Analysis

The Mixed Procedure
Influence Diagnostics for Levels of Person

<table>
<thead>
<tr>
<th>Person</th>
<th>Observations in Level</th>
<th>PRESS Statistic</th>
<th>Cook's D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>10.1716</td>
<td>0.01539</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>3.8187</td>
<td>0.03988</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>10.8448</td>
<td>0.02891</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>24.0339</td>
<td>0.04515</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1.6900</td>
<td>0.01613</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>11.8592</td>
<td>0.01634</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>1.1887</td>
<td>0.00521</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>4.6717</td>
<td>0.02742</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>13.4244</td>
<td>0.03949</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>85.1195</td>
<td>0.13848</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>67.9397</td>
<td>0.09728</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>40.6467</td>
<td>0.04438</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>13.0304</td>
<td>0.00924</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>6.1712</td>
<td>0.00411</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>24.5702</td>
<td>0.12727</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>20.5266</td>
<td>0.01026</td>
</tr>
<tr>
<td>17</td>
<td>4</td>
<td>9.9917</td>
<td>0.01526</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>7.9355</td>
<td>0.01070</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>15.5955</td>
<td>0.01982</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>42.6845</td>
<td>0.01973</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>95.3282</td>
<td>0.10075</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>13.9649</td>
<td>0.03778</td>
</tr>
<tr>
<td>23</td>
<td>4</td>
<td>4.9656</td>
<td>0.01245</td>
</tr>
<tr>
<td>24</td>
<td>4</td>
<td>37.2494</td>
<td>0.15094</td>
</tr>
<tr>
<td>25</td>
<td>4</td>
<td>4.3756</td>
<td>0.03375</td>
</tr>
<tr>
<td>26</td>
<td>4</td>
<td>8.1448</td>
<td>0.03470</td>
</tr>
<tr>
<td>27</td>
<td>4</td>
<td>20.2913</td>
<td>0.02523</td>
</tr>
</tbody>
</table>

Each observation in the “Influence Diagnostics for Levels of Person” table in Output 63.8.1 represents the removal of four observations. The subjects 10, 15, and 24 have the greatest impact on the fixed effects (Cook’s D), and subject 10 and 21 have large PRESS statistics. The 21st child has a large PRESS statistic, and its D statistic is not that extreme. This is an indication that the model fits rather poorly for this child, whether it is part of the data or not.

The previous analysis does not take into account the effect on the covariance parameters when a subject is removed from the analysis. If you also update the covariance parameters, the impact of observations on these
can amplify or allay their effect on the fixed effects. To assess the overall influence of subjects on the analysis and to compute separate statistics for the fixed effects and covariance parameters, an iterative analysis is obtained by adding the INFLUENCE suboption ITER=, as follows:

```plaintext
ods graphics on;
proc mixed data=pr method=ml;
   class person gender;
   model y = gender age gender*age /
            influence(effect=person iter=5);
   repeated / type=un subject=person;
run;
```

The number of additional iterations following removal of the observations for a particular subject is limited to five. Graphical displays of influence diagnostics are created when ODS Graphics is enabled. For general information about ODS Graphics, see Chapter 21, “Statistical Graphics Using ODS.” For specific information about the graphics available in the MIXED procedure, see the section “ODS Graphics” on page 5346.

The MIXED procedure produces a plot of the restricted likelihood distance (Output 63.8.2) and a panel of diagnostics for fixed effects and covariance parameters (Output 63.8.3).

**Output 63.8.2** Restricted Likelihood Distance

![Restricted Likelihood Distance](image-url)
Output 63.8.3 Influence Diagnostics Panel

As judged by the restricted likelihood distance, subjects 20 and 24 clearly have the most influence on the overall analysis (Output 63.8.2).

Output 63.8.3 displays Cook’s $D$ and CovRatio statistics for the fixed effects and covariance parameters. Clearly, subject 20 has a dramatic effect on the estimates of variances and covariances. This subject also affects the precision of the covariance parameter estimates more than any other subject in Output 63.8.3 (CovRatio near 0).

The child who exerts the greatest influence on the fixed effects is subject 24. Maybe surprisingly, this subject affects the variance-covariance matrix of the fixed effects more than subject 20 (small CovRatio in Output 63.8.3).

The final model investigated for these data is a random coefficient model as in Stram and Lee (1994) with random effects for the intercept and age effect. The following statements examine the estimates for fixed effects and the entries of the unstructured $2 \times 2$ variance matrix of the random coefficients graphically:

```plaintext
proc mixed data=pr method=ml
   plots(only)=InfluenceEstPlot;
   class person gender;
   model y = gender age gender*age /
```
influence(iter=5 effect=person est);
   random intercept age / type=un subject=person;
run;

The PLOTS(ONLY)=INFLUENCEESTPLOT option restricts the graphical output from this PROC MIXED run to only the panels of deletion estimates (Output 63.8.4 and Output 63.8.5).

Output 63.8.4 Fixed-Effects Deletion Estimates

In Output 63.8.4 the graphs on the left side of the panel represent the intercept and slope estimate for boys; the graphs on the right side represent the difference in intercept and slope between boys and girls. Removing any one of the first eleven children, who are girls, does not alter the intercept or slope in the group of boys. The difference in these parameters between boys and girls is altered by the removal of any child. Subject 24 changes the fixed effects considerably, subject 20 much less so.
The covariance parameter deletion estimates in Output 63.8.5 show several important features.

- The panels do not contain information about subject 24. Estimation of the $G$ matrix following removal of that child did not yield a positive definite matrix. As a consequence, covariance parameter diagnostics are not produced for this subject.

- Subject 20 has great impact on the four covariance parameters. Removing this child from the analysis increases the variance of the random intercept and random slope and reduces the residual variance by almost 80%. The repeated measurements of this child exhibit an up-and-down behavior.

- The variance of the random intercept and slope are reduced when child 15 is removed from the analysis. This child’s growth measurements oscillate about 27.0 from age 10 on.

Examining observed and residual values by levels of classification variables is also a useful tool to diagnose the adequacy of the model and unusual observations. Box plots for effects in the model that consist of only classification variables can be requested with the BOXPLOT option of the PLOTS= option in the PROC MIXED statement. For example, the following statements produce box plots for the SUBJECT= effects in the model:
ods graphics on;
proc mixed data=pr method=ml
   plot=boxplot(observed marginal conditional subject);
   class person gender;
   model y = gender age gender*age;
   random intercept age / type=un subject=person;
run;

The specific boxplot options request a plot of the observed data (Output 63.8.6), the marginal residuals (Output 63.8.7), and the conditional residuals (Output 63.8.8). Box plots of the observed values show the variation within and between children clearly. The group of girls (subjects 1–11) is distinguishable from the group of boys by somewhat lesser average growth and lesser within-child variation (Output 63.8.6). After adjusting for overall (population-averaged) gender and age effects, the residual within-child variation is reduced but substantial differences in the means remain (Output 63.8.7). If child-specific inferences are desired, a model accounting for only Gender, Age, and Gender*Age effects is not adequate for these data.

Output 63.8.6  Distribution of Observed Values
Example 63.8: Influence Analysis for Repeated Measures Data

Output 63.8.7 Distribution of Marginal Residuals

The conditional residuals incorporate the EBLUPs for each child and enable you to examine whether the subject-specific model is adequate (Output 63.8.8). By using each child “as its own control,” the residuals are now centered near zero. Subjects 20 and 24 stand out as unusual in all three sets of box plots.
Example 63.9: Examining Individual Test Components

The LCOMPONENTS option in the MODEL statement enables you to perform single-degree-of-freedom tests for individual rows of the L matrix. Such tests are useful to identify interaction patterns. In a balanced layout, Type 3 components of L associated with A*B interactions correspond to simple contrasts of cell mean differences.

The first example revisits the data from the split-plot design by Stroup (1989a) that was analyzed in Example 63.1. Recall that variables A and B in the following statements represent the whole-plot and subplot factors, respectively:

```r
proc mixed data=sp;
   class a b block;
   model y = a b a*b / LComponents e3;
   random block a*block;
run;
```

The MIXED procedure constructs a separate L matrix for each of the three fixed-effects components. The matrices are displayed in Output 63.9.1. The tests for fixed effects are shown in Output 63.9.2.
Output 63.9.1  Coefficients of Type 3 Estimable Functions

The Mixed Procedure

Type 3 Coefficients for A

<table>
<thead>
<tr>
<th>Effect</th>
<th>A</th>
<th>B</th>
<th>Row1</th>
<th>Row2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>1</td>
<td>2</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>1</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>3</td>
<td>1</td>
<td>-0.5</td>
<td>-0.5</td>
</tr>
<tr>
<td>A*B</td>
<td>3</td>
<td>2</td>
<td>-0.5</td>
<td>-0.5</td>
</tr>
</tbody>
</table>

Type 3 Coefficients for B

<table>
<thead>
<tr>
<th>Effect</th>
<th>A</th>
<th>B</th>
<th>Row1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>1</td>
<td>1</td>
<td>0.3333</td>
</tr>
<tr>
<td>A*B</td>
<td>1</td>
<td>2</td>
<td>-0.333</td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>1</td>
<td>0.3333</td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>2</td>
<td>-0.333</td>
</tr>
<tr>
<td>A*B</td>
<td>3</td>
<td>1</td>
<td>0.3333</td>
</tr>
<tr>
<td>A*B</td>
<td>3</td>
<td>2</td>
<td>-0.333</td>
</tr>
</tbody>
</table>

Type 3 Coefficients for A*B

<table>
<thead>
<tr>
<th>Effect</th>
<th>A</th>
<th>B</th>
<th>Row1</th>
<th>Row2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>1</td>
<td>2</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>2</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>A*B</td>
<td>3</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>A*B</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Output 63.9.2 Type 3 Tests in Split-Plot Example

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>6</td>
<td>4.07</td>
<td>0.0764</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>9</td>
<td>19.39</td>
<td>0.0017</td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>9</td>
<td>4.02</td>
<td>0.0566</td>
</tr>
</tbody>
</table>

If \( \mu_i \) denotes a whole-plot main effect mean, \( \mu_j \) denotes a subplot main effect mean, and \( \mu_{ij} \) denotes a cell mean, the five components shown in Output 63.9.3 correspond to tests of the following:

- \( H_0^{W1} : \mu_1 = \mu_3 \)
- \( H_0^{W2} : \mu_2 = \mu_3 \)
- \( H_0^{W1} : \mu_1 = \mu_2 \)
- \( H_0 : \mu_{11} - \mu_{12} = \mu_{31} - \mu_{32} \)
- \( H_0 : \mu_{21} - \mu_{22} = \mu_{31} - \mu_{32} \)

Output 63.9.3 Type 3 L Components Table

<table>
<thead>
<tr>
<th>L Components of Type 3 Tests of Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>A*B</td>
</tr>
<tr>
<td>A*B</td>
</tr>
</tbody>
</table>

The first three components are comparisons of marginal means. The fourth component compares the effect of factor B at the first whole-plot level against the effect of B at the third whole-plot level. Finally, the last component tests whether the factor B effect changes between the second and third whole-plot level.

The Type 3 component tests can also be produced with these corresponding ESTIMATE statements:

```plaintext
proc mixed data=sp;
   class a b block;
   model y = a b a*b;
   random block a*block;
   estimate 'a 1' a 1 0 -1;
   estimate 'a 2' a 0 1 -1;
   estimate 'b 1' b 1 -1;
```

```plaintext
```
Example 63.9: Examining Individual Test Components

```
estimate 'a*b 1' a*b 1 -1 0 0 -1 1;
estimate 'a*b 2' a*b 0 0 1 -1 -1 1;
ods select Estimates;
run;
```

The results are shown in Output 63.9.4.

**Output 63.9.4** Results from ESTIMATE Statements

| Estimation | Label | Estimate | Standard Error | DF | t Value | Pr > |t| |
|------------|-------|----------|----------------|----|---------|------|---|
|            | a 1   | 7.1250   | 3.1672         | 6  | 2.25    | 0.0655 |
|            | a 2   | 8.3750   | 3.1672         | 6  | 2.64    | 0.0383 |
|            | b 1   | 5.5000   | 1.2491         | 9  | 4.40    | 0.0017 |
|            | a*b 1 | 7.7500   | 3.0596         | 9  | 2.53    | 0.0321 |
|            | a*b 2 | 7.2500   | 3.0596         | 9  | 2.37    | 0.0419 |

A second useful application of the LCOMPONENTS option is in polynomial models, where Type 1 tests are often used to test the entry of model terms sequentially. The SOLUTION option in the MODEL statement displays the regression coefficients that correspond to a Type 3 analysis. That is, the coefficients represent the partial coefficients you would get by adding the regressor variable last in a model containing all other effects, and the tests are identical to those in the “Type 3 Tests of Fixed Effects” table.

Consider the following DATA step and the fit of a third-order polynomial regression model.

```
data polynomial;
  do x=1 to 20; input y@@; output; end;
datalines;
  1.092   1.758  1.997  3.154  3.880
  3.810   4.921  4.573  6.029  6.032
  6.291   7.151  7.154  6.469  7.137
  6.374   5.860  4.866  4.155  2.711
;

proc mixed data=polynomial;
  model y = x x*x x*x*x / s lcomponents htype=1,3;
run;
```

The $t$ tests displayed in the “Solution for Fixed Effects” table are Type 3 tests, sometimes referred to as partial tests. They measure the contribution of a regressor in the presence of all other regressor variables in the model.
Output 63.9.5 Parameter Estimates in Polynomial Model

The Mixed Procedure

Solution for Fixed Effects

| Effect  | Estimate | Error  | DF  | t Value | Pr > |t|  |
|---------|----------|--------|-----|---------|-------|-----|
| Intercept | 0.7837  | 0.3545 | 16  | 2.21    | 0.0420|
| x         | 0.3726  | 0.1426 | 16  | 2.61    | 0.0189|
| x*x       | 0.04756 | 0.01558| 16  | 3.05    | 0.0076|
| x*x*x     | -0.00306| 0.000489| 16  | -6.27   | <.0001|

The Type 3 L components are identical to the tests in the “Solutions for Fixed Effects” table shown in Output 63.9.5. The Type 1 table yields the following:

- sequential (Type 1) tests of regression variables that test the significance of a regressor given all other variables preceding it in the model list
- the regression coefficients for sequential submodels

Output 63.9.6 Type 1 and Type 3 L Components

L Components of Type 1 Tests of Fixed Effects

| L | Effect | Index | Estimate | Error   | DF | t Value | Pr > |t|  |
|---|--------|-------|----------|---------|----|---------|-------|-----|
| 1 | x      | 0.1763| 0.01259  | 16      | 14.01| <.0001  |
| 1 | x*x    | -0.04886| 0.002449 | 16      | -19.95| <.0001  |
| 1 | x*x*x  | -0.00306| 0.000489 | 16      | -6.27 | <.0001  |

L Components of Type 3 Tests of Fixed Effects

| L | Effect | Index | Estimate | Error   | DF | t Value | Pr > |t|  |
|---|--------|-------|----------|---------|----|---------|-------|-----|
| 1 | x      | 0.3726| 0.1426   | 16      | 2.61| 0.0189  |
| 1 | x*x    | 0.04756| 0.01558  | 16      | 3.05| 0.0076  |
| 1 | x*x*x  | -0.00306| 0.000489 | 16      | -6.27| <.0001  |

The estimate of 0.1763 is the regression coefficient in a simple linear regression of $Y$ on $X$. The estimate of $-0.04886$ is the partial coefficient for the quadratic term when it is added to a model containing only a linear component. Similarly, the value $-0.00306$ is the partial coefficient for the cubic term when it is added to a model containing a linear and quadratic component. The last Type 1 component is always identical to the corresponding Type 3 component.
Example 63.10: Isotonic Contrasts for Ordered Mean Values

It is often of interest to test whether the mean values of the dependent variable increases or decreases monotonically with certain factors. Hirotsu and Srivastava (2000) demonstrate one approach by using data (Moriguchi 1976). The data consist of ferrite cores subjected to four increasing temperatures. The response variable is the magnetic force of each core.

data FerriteCores;
  do Temp = 1 to 4;
    do rep = 1 to 5; drop rep;
      input MagneticForce @@;
      output;
    end;
  end;
datalines;
10.8 9.9 10.7 10.4 9.7
10.7 10.6 11.0 10.8 10.9
11.9 11.2 11.0 11.1 11.3
11.4 10.7 10.9 11.3 11.7
;

The method presented by Hirotsu and Srivastava (2000) to test whether the magnetic force of the cores rises monotonically with temperature depends on the lower confidence limits of the isotonic contrasts of the force means at each temperature, adjusted for multiplicity. The corresponding isotonic contrast compares the average of a particular group and the preceding groups with the average of the succeeding groups. You can compute adjusted confidence intervals for isotonic contrasts by using the LSMESTIMATE statement.

The following statements analyze the FerriteCores data as a one-way design and multiplicity-adjusted lower confidence limits for the isotonic contrasts. For the multiplicity adjustment, the LSMESTIMATE statement employs simulation, which provides adjusted p-values and lower confidence limits that are exact up to Monte Carlo error.

proc mixed data=FerriteCores;
  class Temp;
  model MagneticForce = Temp;
  lsmestimate Temp
    'avg(1:1)<avg(2:4)' -3 1 1 1 divisor=3,
    'avg(1:2)<avg(3:4)' -1 -1 1 1 divisor=2,
    'avg(1:3)<avg(4:4)' -1 -1 -1 3 divisor=3
    / adjust=simulate(seed=1) cl upper;
  ods select LSMestimates;
run;

The results are shown in Output 63.10.1.
Output 63.10.1 Analysis of LS-Means with Isotonic Contrasts

The Mixed Procedure

Least Squares Means Estimates
Adjustment for Multiplicity: Simulated

<table>
<thead>
<tr>
<th>Effect Label</th>
<th>Estimate</th>
<th>Error</th>
<th>DF</th>
<th>t Value</th>
<th>Tails</th>
<th>Pr &gt; t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp avg(1:1)&lt;avg(2:4)</td>
<td>0.8000</td>
<td>0.1906</td>
<td>16</td>
<td>4.20</td>
<td>Upper</td>
<td>0.0003</td>
</tr>
<tr>
<td>Temp avg(1:2)&lt;avg(3:4)</td>
<td>0.7000</td>
<td>0.1651</td>
<td>16</td>
<td>4.24</td>
<td>Upper</td>
<td>0.0003</td>
</tr>
<tr>
<td>Temp avg(1:3)&lt;avg(4:4)</td>
<td>0.4000</td>
<td>0.1906</td>
<td>16</td>
<td>2.10</td>
<td>Upper</td>
<td>0.0260</td>
</tr>
</tbody>
</table>

Least Squares Means Estimates
Adjustment for Multiplicity: Simulated

<table>
<thead>
<tr>
<th>Effect Label</th>
<th>Adj P</th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp avg(1:1)&lt;avg(2:4)</td>
<td>0.0010</td>
<td>0.05</td>
<td>0.4672 Infty</td>
<td>Infty</td>
</tr>
<tr>
<td>Temp avg(1:2)&lt;avg(3:4)</td>
<td>0.0009</td>
<td>0.05</td>
<td>0.4118 Infty</td>
<td>Infty</td>
</tr>
<tr>
<td>Temp avg(1:3)&lt;avg(4:4)</td>
<td>0.0625</td>
<td>0.05</td>
<td>0.06721 Infty</td>
<td>Infty</td>
</tr>
</tbody>
</table>

Least Squares Means Estimates
Adjustment for Multiplicity: Simulated

<table>
<thead>
<tr>
<th>Effect Label</th>
<th>Adj Lower</th>
<th>Adj Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp avg(1:1)&lt;avg(2:4)</td>
<td>0.3771 Infty</td>
<td>Infty</td>
</tr>
<tr>
<td>Temp avg(1:2)&lt;avg(3:4)</td>
<td>0.3337 Infty</td>
<td>Infty</td>
</tr>
<tr>
<td>Temp avg(1:3)&lt;avg(4:4)</td>
<td>-0.02291 Infty</td>
<td>Infty</td>
</tr>
</tbody>
</table>

With an adjusted $p$-value of 0.001, the magnetic force at the first temperature is significantly less than the average of the other temperatures. Likewise, the average of the first two temperatures is significantly less than the average of the last two ($p = 0.0009$). However, the magnetic force at the last temperature is not significantly greater than the average magnetic force of the others ($p = 0.0625$). These results indicate a significant monotone increase over the first three temperatures, but not across all four temperatures.

References


Marx, D. and Thompson, K. (1987), *Practical Aspects of Agricultural Kriging*, Bulletin 903, Arkansas Agricultural Experiment Station, Fayetteville, NC.


### Subject Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D geometric anisotropic structure</td>
<td>5303</td>
</tr>
<tr>
<td>Akaike’s information criterion</td>
<td>5358, 5372, 5401</td>
</tr>
<tr>
<td>MIXED procedure</td>
<td>5249, 5321, 5340</td>
</tr>
<tr>
<td>Akaike’s information criterion (finite sample corrected version)</td>
<td>5249, 5340</td>
</tr>
<tr>
<td>alpha level</td>
<td>5247, 5264, 5268, 5273, 5295</td>
</tr>
<tr>
<td>anisotropic power covariance structure</td>
<td>5304</td>
</tr>
<tr>
<td>MIXED procedure</td>
<td>5304</td>
</tr>
<tr>
<td>anisotropic spatial power structure</td>
<td>5304</td>
</tr>
<tr>
<td>ANTE(1) structure</td>
<td>5303</td>
</tr>
<tr>
<td>antedependence structure</td>
<td>5303</td>
</tr>
<tr>
<td>AR(1) structure</td>
<td>5303</td>
</tr>
<tr>
<td>asymptotic covariance</td>
<td>5248</td>
</tr>
<tr>
<td>at sign (@) operator</td>
<td>5327, 5397</td>
</tr>
<tr>
<td>autoregressive moving-average structure</td>
<td>5303</td>
</tr>
<tr>
<td>autoregressive structure</td>
<td>5303</td>
</tr>
<tr>
<td>example (MIXED), 5366</td>
<td>5303</td>
</tr>
<tr>
<td>banded Toeplitz structure</td>
<td>5303</td>
</tr>
<tr>
<td>bar (l) operator</td>
<td>5326, 5327, 5397</td>
</tr>
<tr>
<td>Bayesian analysis</td>
<td>5290</td>
</tr>
<tr>
<td>BLUE</td>
<td>5321</td>
</tr>
<tr>
<td>BLUP</td>
<td>5321</td>
</tr>
<tr>
<td>Bonferroni adjustment</td>
<td>5267</td>
</tr>
<tr>
<td>boundary constraints</td>
<td>5288, 5289, 5352</td>
</tr>
<tr>
<td>CALIS procedure</td>
<td>5237</td>
</tr>
<tr>
<td>compared to MIXED procedure</td>
<td>5237</td>
</tr>
<tr>
<td>chi-square test</td>
<td>5237</td>
</tr>
<tr>
<td>class level</td>
<td>5251, 5338</td>
</tr>
<tr>
<td>compound symmetry structure</td>
<td>5314, 5367, 5372</td>
</tr>
<tr>
<td>MIXED procedure</td>
<td>5303</td>
</tr>
<tr>
<td>computational details</td>
<td>5351</td>
</tr>
<tr>
<td>computational problems</td>
<td>5351</td>
</tr>
<tr>
<td>confidence limits</td>
<td>5330</td>
</tr>
<tr>
<td>adjusted (MIXED), 5425</td>
<td>5330</td>
</tr>
<tr>
<td>and isotronic contrasts (MIXED), 5425</td>
<td>5330</td>
</tr>
<tr>
<td>MIXED procedure</td>
<td>5248</td>
</tr>
<tr>
<td>constraints</td>
<td>5288, 5289</td>
</tr>
<tr>
<td>containment method</td>
<td>5274, 5275</td>
</tr>
<tr>
<td>continuous-by-class effects</td>
<td>5328</td>
</tr>
<tr>
<td>continuous-nesting-class effects</td>
<td>5327</td>
</tr>
<tr>
<td>contrasts</td>
<td>5260, 5263</td>
</tr>
<tr>
<td>convergence criterion</td>
<td>5247, 5249, 5338, 5353</td>
</tr>
<tr>
<td>convergence problems</td>
<td>5353</td>
</tr>
<tr>
<td>convergence status</td>
<td>5339</td>
</tr>
<tr>
<td>Cook’s D</td>
<td>5334</td>
</tr>
<tr>
<td>for covariance parameters</td>
<td>5334</td>
</tr>
<tr>
<td>correlation estimates (MIXED), 5295, 5298, 5302, 5369</td>
<td>5334</td>
</tr>
<tr>
<td>covariance</td>
<td>5248, 5249</td>
</tr>
<tr>
<td>parameter estimates (MIXED), 5257</td>
<td>5248, 5249</td>
</tr>
<tr>
<td>parameters (MIXED)</td>
<td>5234</td>
</tr>
<tr>
<td>covariance parameter estimates</td>
<td>5339</td>
</tr>
<tr>
<td>anisotropic power (MIXED), 5310</td>
<td>5339</td>
</tr>
<tr>
<td>autoregressive (MIXED), 5307</td>
<td>5339</td>
</tr>
<tr>
<td>chi-square test</td>
<td>5237</td>
</tr>
</tbody>
</table>
autoregressive moving-average (MIXED), 5307
banded (MIXED), 5310
compound symmetry (MIXED), 5307
equi-correlation (MIXED), 5307
examples (MIXED), 5305, 5361
exponential anisotropic (MIXED), 5308
factor-analytic (MIXED), 5308
general linear (MIXED), 5308
heterogeneous autoregressive (MIXED), 5307
heterogeneous compound symmetry (MIXED), 5308
heterogeneous Toeplitz (MIXED), 5310
Huynh-Feldt (MIXED), 5308
Kronecker (MIXED), 5311
Matérn (MIXED), 5310
MIXED procedure, 5236, 5303
power (MIXED), 5310
simple (MIXED), 5308
spatial geometric anisotropic (MIXED), 5309
Toeplitz (MIXED), 5310
unstructured (MIXED), 5310
unstructured, correlation (MIXED), 5310
variance components (MIXED), 5311
covariates
MIXED procedure, 5326
covRatio
MIXED procedure, 5335
covRatio for covariance parameters
MIXED procedure, 5335
covTrace
MIXED procedure, 5335
covTrace for covariance parameters
MIXED procedure, 5335
crossed effects
MIXED procedure, 5326
default output
MIXED procedure, 5337
degrees of freedom
between-within method (MIXED), 5249, 5275
covariance parameter method (MIXED), 5274, 5275
Kenward-Roger method (GLIMMIX), 5276
method (MIXED), 5274
MIXED procedure, 5262, 5264, 5269, 5274
residual method (MIXED), 5275
Satterthwaite method (MIXED), 5275
dFfiting
MIXED procedure, 5334
dimension information
MIXED procedure, 5338
dimensions
MIXED procedure, 5252
direct product structure
MIXED procedure, 5303
Dunnett’s adjustment
MIXED procedure, 5267
EBLUP
MIXED procedure, 5284
effect
name length (MIXED), 5251
empirical best linear unbiased prediction
MIXED procedure, 5284
empirical estimator
MIXED procedure, 5249
estimability
MIXED procedure, 5261
estimable functions
MIXED procedure, 5283
estimation
mixed model (MIXED), 5318
estimation methods
MIXED procedure, 5251
examples
MIXED
ASYCOV matrix, 5364
asymptotic covariance of covariance parameters, 5364
autoregressive structure, R-side, 5366
box plots, 5418
box plots, paneling, 5256
broad inference space, 5261, 5263
compound symmetry, G-side setup, 5315, 5370
compound symmetry, R-side setup, 5315, 5367
constrained anisotropic model, 5309
covariates in LS-mean construction, 5268
COVTEST option, 5361, 5374
deletion estimates, 5402
doubly repeated measure, 5311
estimate, with subject, 5265
fat absorption data, 5402
ferrite cores data, 5425
fixed-effect solutions, 5389
full-rank parameterization, 5366
GDATA= option in RANDOM statement, 5382
geometrically anisotropic model, 5309
getting started, 5238
GLM procedure, split-plot design, 5359
graphics, box plots, 5418
graphics, influence diagnostics, 5402, 5414
graphics, residual panel, 5347
graphics, studentized residual panel, 5347
GROUP= effect in RANDOM statement, 5370
height data, 5238
holding covariance parameters fixed, 5287, 5309
IML procedure, reading ASYCOV, 5365
inference space, broad, 5261, 5263
inference space, intermediate, 5263
inference space, narrow, 5261, 5263
inference spaces, 5359
influence analysis, iterative, 5402, 5414
influence analysis, non-iterative, 5412
influence analysis, set deletion, 5412, 5414
influence analysis, tuples, 5280
intermediate inference space, 5263
isotonic contrast, 5425
known covariance parameters, 5286
Kronecker covariance structure, 5311
L-components, 5283, 5420, 5423
least squares means estimate, 5425
least squares means, AT option, 5268
least squares means, covariate, 5268
least squares means, differences against control, 5269
least squares means, slice, 5271
line-source sprinkler data, 5396
local power-of-mean model, 5301
maximum likelihood estimation, 5361
mixed model equations, 5374, 5382
mixed model equations, solution, 5374, 5382
multiple plot requests, 5257
multiple traits data, 5381
multiplicity adjustment, 5425
multivariate analysis, 5311
nested error structure, 5392
nested random effects, 5241
NOITER option, 5286, 5382
oven data (Hemmerle and Hartley, 1973), 5373
parameter grid search, 5374
pharmaceutical stability data, 5388
polynomial model, 5423
POM data set, 5301
POM fitting, iterated, 5301
Pothoff and Roy growth measurements, 5361, 5412
random coefficient model, 5367, 5389, 5416
random-effect solutions, 5389
residual panel, 5347
row-wise multiplicity adjustment, 5267
Satterthwaite method, 5267
set deletion, 5414
SGRENDER procedure, 5380
slice F test, 5271
spatial power structure, 5402
specifying lower bounds, 5288
specifying values for degrees of freedom, 5274
split-plot design, 5316, 5336
split-plot design, data, 5355, 5420
split-plot design, equivalent model, 5360
starting values, 5374
studentized maximum modulus, 5267
studentized residual panel, 5347
subject and no-subject formulation, 5315
subject contrasts, 5265
subject v. no-subject formulation, 5315
subject-specific R matrices, 5302
subject-specific V matrices, 5298
Toeplitz structure, 5397
tuples, influence analysis, 5280
two-way analysis of variance, 5238
unstructured covariance, G-side, 5297
unstructured covariance, R-side, 5361, 5412
varying covariance parameters, 5370
exponential covariance structure
MIXED procedure, 5304
external studentization
MIXED procedure, 5330
factor analytic structures
MIXED procedure, 5303
Fisher information matrix
element (MIXED), 5374
MIXED procedure, 5339
Fisher’s scoring method
MIXED procedure, 5248, 5257, 5353
fixed effects
MIXED procedure, 5236
fixed-effects parameters
MIXED procedure, 5234, 5314
G matrix
MIXED procedure, 5236, 5294, 5295, 5314, 5394
Gaussian covariance structure
MIXED procedure, 5304
general linear covariance structure
MIXED procedure, 5303
generalized inverse, 5321
MIXED procedure, 5262
GLM procedure
compared to other procedures, 5237
gradient
MIXED procedure, 5248, 5249, 5338
grid search
element (MIXED), 5374
growth curve analysis
element (MIXED), 5314
Hannan-Quinn information criterion
MIXED procedure, 5249
Hessian matrix
MIXED procedure, 5248, 5249, 5257, 5287, 5338, 5339, 5353, 5354, 5364, 5374
heterogeneity
element (MIXED), 5370
MIXED procedure, 5296, 5299
heterogeneous
AR(1) structure (MIXED), 5303
compound-symmetry structure (MIXED), 5303
covariance structure (MIXED), 5312
Toeplitz structure (MIXED), 5303
hierarchical model
element (MIXED), 5388
Hotelling-Lawley-McKeon statistic
MIXED procedure, 5300
Hotelling-Lawley-Pillai-Samson statistic
MIXED procedure, 5300
Hsu’s adjustment
MIXED procedure, 5267
Huynh-Feldt
structure (MIXED), 5303
hypothesis tests
mixed model (MIXED), 5322, 5340
inference
mixed model (MIXED), 5322
space, mixed model (MIXED), 5260, 5261, 5263, 5359
infinite likelihood
MIXED procedure, 5298, 5352, 5353
influence diagnostics
MIXED procedure, 5331
influence plots
MIXED procedure, 5349
information criteria
MIXED procedure, 5249
initial values
MIXED procedure, 5286
interaction effects
MIXED procedure, 5326
intercept
MIXED procedure, 5325
internal studentization
MIXED procedure, 5330
intraclass correlation coefficient
MIXED procedure, 5369
iteration history
MIXED procedure, 5338
iterations
history (MIXED), 5338
Kenward-Roger method
MIXED procedure, 5276
Kronecker product structure
MIXED procedure, 5303
L matrices
mixed model (MIXED), 5260, 5265, 5322
MIXED procedure, 5260, 5265, 5322
LATTICE procedure
compared to MIXED procedure, 5237
least squares means
Bonferroni adjustment (MIXED), 5267
BYLEVEL processing (MIXED), 5269
comparison types (MIXED), 5269
covariate values (MIXED), 5268
Dunnett’s adjustment (MIXED), 5267
elements (MIXED), 5374, 5397
Hsu’s adjustment (MIXED), 5267
mixed model (MIXED), 5265
multiple comparison adjustment (MIXED), 5266, 5267
nonstandard weights (MIXED), 5270
observed margins (MIXED), 5270
Sidak’s adjustment (MIXED), 5267
simple effects (MIXED), 5271
simulation-based adjustment (MIXED), 5267
Tukey’s adjustment (MIXED), 5267
leverage
MIXED procedure, 5333
likelihood distance
MIXED procedure, 5335
likelihood ratio test, 5358
elements (MIXED), 5372
mixed model (MIXED), 5322, 5323
MIXED procedure, 5340
linear covariance structure
MIXED procedure, 5303
log-linear variance model
MIXED procedure, 5300
main effects
MIXED procedure, 5326
marginal residuals
MIXED procedure, 5330
Matérn covariance structure
MIXED procedure, 5303
matrix
notation, theory (MIXED), 5313
maximum likelihood estimation
mixed model (MIXED), 5319
MDFFITS
MIXED procedure, 5334
MDFFITS for covariance parameters
MIXED procedure, 5335
memory requirements
MIXED procedure, 5354
missing level combinations
MIXED procedure, 5329
mixed model (MIXED), see also MIXED procedure
estimation, 5318
formulation, 5314
hypothesis tests, 5322, 5340
inference, 5322
inference space, 5260, 5261, 5263, 5359
least squares means, 5265
likelihood ratio test, 5322, 5323
linear model, 5234
maximum likelihood estimation, 5319
notation, 5236
objective function, 5338
parameterization, 5325
predicted values, 5265
restricted maximum likelihood, 5357
theory, 5313
Wald test, 5322, 5364

mixed model equations
example (MIXED), 5374
MIXED procedure, 5320

MIXED procedure, see also mixed model
2D geometric anisotropic structure, 5303
Akaike’s information criterion, 5249, 5321, 5340
Akaike’s information criterion (finite sample corrected version), 5249, 5340
alpha level, 5247, 5264, 5268, 5273, 5295
anisotropic power covariance structure, 5304
anisotropic spatial power structure, 5304
ANTE(1) structure, 5303
antedependence structure, 5303
AR(1) structure, 5303
ARIMA procedure, compared, 5237
ARMA structure, 5303
assumptions, 5234
asymptotic covariance, 5248
AUTOREG procedure, compared, 5237
autoregressive moving-average structure, 5303
autoregressive structure, 5303, 5366
banded Toeplitz structure, 5303
basic features, 5235
Bayesian analysis, 5290
between-within method, 5249, 5275
BLUE, 5231
BLUP, 5231, 5388
Bonferroni adjustment, 5267
boundary constraints, 5288, 5289, 5352
BYLEVEL processing of LSMEANS, 5269
CALIS procedure, compared, 5237
chi-square test, 5262, 5273
Cholesky root, 5285, 5331, 5351
class level, 5251, 5338
compound symmetry structure, 5303, 5314, 5367, 5372
computational details, 5351
computational order, 5352
conditional residuals, 5330
confidence interval, 5264, 5295
confidence limits, 5248, 5264, 5269, 5274, 5295
containment method, 5274, 5275
continuous effects, 5296, 5297, 5299, 5303
continuous-by-class effects, 5328
covariance parameter estimates, 5248, 5249, 5339
covariance parameter estimates, ratio, 5257
covariance parameters, 5234
covariance structure, 5236, 5303, 5305, 5361
covariances of least squares means, 5269
covariate values for LSMEANS, 5268
covariates, 5326
covariate values for covariance parameters, 5334
covariance parameter estimates, 5248, 5249, 5339
covariance structures, 5303, 5314, 5367, 5372
continuous effects, 5296, 5297, 5299, 5303
continuous-nesting-class effects, 5327
desktop models, 5234
default output, 5337
degrees of freedom, 5261–5264, 5266, 5269, 5274, 5285, 5233, 5329, 5340, 5352, 5388
DFFITS, 5334
dimension information, 5338
dimensions, 5250, 5252
direct product structure, 5303
Dunnett’s adjustment, 5267
EBLUPs, 5297, 5321, 5380, 5395
effect name length, 5251
empirical best linear unbiased prediction, 5284
empirical estimator, 5249
estimability, 5261, 5262, 5264, 5265, 5270, 5285, 5286, 5322, 5329
estimable functions, 5283
estimation methods, 5251
exponential covariance structure, 5304
factor analytic structures, 5303
Fisher information matrix, 5339, 5374
Fisher’s scoring method, 5248, 5257, 5353
fitting information, 5340
fixed effects, 5236
fixed-effects parameters, 5234, 5285, 5314
fixed-effects variance matrix, 5286
function evaluations, 5250
G matrix, 5236, 5294, 5295, 5314, 5394
Gaussian covariance structure, 5304
general linear covariance structure, 5303
generalized inverse, 5262, 5321
GLIMMIX procedure, compared, 5237
gradient, 5248, 5249, 5327
grid search, 5286, 5374
growth curve analysis, 5314
Hannan-Quinn information criterion, 5249
Hessian matrix, 5248, 5249, 5257, 5287, 5338, 5339, 5353, 5354, 5364, 5374
heterogeneity, 5296, 5299, 5370
heterogeneous AR(1) structure, 5303
heterogeneous compound-symmetry structure, 5303
heterogeneous covariance structures, 5312
heterogeneous Toeplitz structure, 5303
hierarchical model, 5388
Hotelling-Lawley-McKeon statistic, 5300
Hotelling-Lawley-Pillai-Sampson statistic, 5300
Hsu's adjustment, 5267
Huynh-Feldt structure, 5303
infinite likelihood, 5298, 5352, 5353
influence diagnostics, 5281, 5331
influence plots, 5349
information criteria, 5249
initial values, 5286
input data sets, 5249
interaction effects, 5326
intercept, 5325
intercept effect, 5283, 5294
intraclass correlation coefficient, 5369
iterative step, 5338
iterations, 5251, 5338
Kenward-Roger method, 5276
Kronecker product structure, 5303
LATTICE procedure, compared, 5237
least squares means, 5269, 5374, 5397
leave-one-out-estimates, 5349
leverage, 5333
likelihood distance, 5335
likelihood ratio test, 5340
linear covariance structure, 5303
log-linear covariance model, 5300
main effects, 5326
marginal residuals, 5330
Matérn covariance structure, 5303
matrix notation, 5313
MDFFFITS, 5334
MDFFFITS for covariance parameters, 5335
memory requirements, 5354
missing level combinations, 5329
mixed linear model, 5234
mixed model, 5314
mixed model equations, 5320, 5374
mixed model theory, 5313
model information, 5252, 5337
model selection, 5320
multi-threading, 5289
multilevel model, 5388
multiple comparisons of least squares means, 5266, 5267, 5269
multiple tables, 5343
multiplicity adjustment, 5266
multivariate tests, 5300
nested effects, 5327
nested error structure, 5392
NESTED procedure, compared, 5237
Newton-Raphson algorithm, 5319
non-full-rank parameterization, 5237, 5301, 5329
nonstandard weights for LSMEANS, 5270
nugget effect, 5300
number of observations, 5338
oblique projector, 5333
observed margins for LSMEANS, 5270
ODS graph names, 5346
ODS Graphics, 5252, 5346
ODS table names, 5341
ordering of effects, 5328
over-parameterization, 5326
parameter constraints, 5287, 5352
parameterization, 5325
Pearson residual, 5285
pharmaceutical stability, example, 5388
plotting the likelihood, 5380
polynomial effects, 5326
power-of-the-mean model, 5300
predicted means, 5284
predicted value confidence intervals, 5273
predicted values, 5284, 5374
PRESS residual, 5332
PRESS statistic, 5332
prior density, 5291
profiling residual variance, 5252, 5288, 5300, 5319, 5351
R matrix, 5236, 5298, 5302, 5314
random coefficients, 5367, 5388
random effects, 5236, 5294
random-effects parameters, 5235, 5297, 5314
regression effects, 5326
rejection sampling, 5292
repeated measures, 5235, 5298, 5361
residual diagnostics, details, 5330
residual method, 5275
residual plots, 5347
residual variance tolerance, 5285
restricted maximum likelihood (REML), 5235
ridging, 5257, 5319
sandwich estimator, 5249
Satterthwaite method, 5275
scaled residual, 5286, 5331
Schwarz's Bayesian information criterion, 5249, 5321, 5340
scoring, 5248, 5257, 5353
Sidak’s adjustment, 5267
simple effects, 5271
simulation-based adjustment, 5267
singularities, 5354
spatial anisotropic exponential structure, 5303
spatial covariance structure, 5304, 5305, 5312, 5353
split-plot design, 5316, 5355
standard linear model, 5236
statement positions, 5244
studentized residual, 5285, 5333
subject effect, 5261, 5297, 5303, 5355, 5360
summary of commands, 5244
sweep operator, 5334, 5351
table names, 5341
test components, 5283
Toeplitz structure, 5303, 5397
TSCSREG procedure, compared, 5237
Tukey’s adjustment, 5267
type 1 estimation, 5251
type 1 testing, 5277
type 2 estimation, 5251
type 2 testing, 5277
type 3 estimation, 5251
type 3 testing, 5277, 5340
unstructured correlations, 5303
unstructured covariance matrix, 5303
unstructured R matrix, 5302
V matrix, 5298
VARCOMP procedure, example, 5373
variance components, 5235, 5303
variance ratios, 5287, 5296
Wald test, 5339, 5340
weighted LSMEANS, 5270
weighting, 5312
zero design columns, 5277
zero variance component estimates, 5352
MIXED procedure
ordering of effects, 5252
model
information (MIXED), 5252
model information
MIXED procedure, 5337
model selection
MIXED procedure, 5320
multi-threading
MIXED procedure, 5289
multilevel model
example (MIXED), 5388
multiple comparison adjustment (MIXED)
least squares means, 5266, 5267
multiple comparisons of least squares means
MIXED procedure, 5266, 5267, 5269
multiple tables
MIXED procedure, 5343
multiplicity adjustment
MIXED procedure, 5266
row-wise (MIXED), 5266
multivariate tests
MIXED procedure, 5300
nested effects
MIXED procedure, 5327
nested error structure
MIXED procedure, 5392
NESTED procedure
compared to other procedures, 5237
Newton-Raphson algorithm
MIXED procedure, 5319
non-full-rank parameterization
MIXED procedure, 5273, 5301, 5329
nugget effect
MIXED procedure, 5300
number of observations
MIXED procedure, 5338
objective function
mixed model (MIXED), 5338
oblique projector
MIXED procedure, 5333
ODS graph names
MIXED procedure, 5346
ODS Graphics
MIXED procedure, 5252, 5346
options summary
LSMEANS statement, (MIXED), 5266
MODEL statement (MIXED), 5272
PROC MIXED statement, 5246
RANDOM statement (MIXED), 5294
REPEATED statement (MIXED), 5299
over-parameterization
MIXED procedure, 5326
parameter constraints
MIXED procedure, 5287, 5352
parameterization
mixed model (MIXED), 5325
MIXED procedure, 5325
Pearson residual
MIXED procedure, 5285
pharmaceutical stability
example (MIXED), 5388
plots
likelihood (MIXED), 5380
polynomial effects
MIXED procedure, 5326
power-of-the-mean model
MIXED procedure, 5300
predicted means
   MIXED procedure, 5284
predicted value confidence intervals
   MIXED procedure, 5273
predicted values
   example (MIXED), 5374
   mixed model (MIXED), 5265
   MIXED procedure, 5284
PRESS residual
   MIXED procedure, 5332
PRESS statistic
   MIXED procedure, 5332
prior density
   MIXED procedure, 5291
profiling residual variance
   MIXED procedure, 5351

R matrix
   MIXED procedure, 5236, 5298, 5302, 5314
random coefficients
   example (MIXED), 5367, 5388
random effects
   MIXED procedure, 5236, 5294
random-effects parameters
   MIXED procedure, 5235, 5314
regression effects
   MIXED procedure, 5326
rejection sampling
   MIXED procedure, 5292

REML, see restricted maximum likelihood

repeated measures
   MIXED procedure, 5235, 5298, 5361
residual maximum likelihood (REML), see also
   restricted maximum likelihood (REML)
   MIXED procedure, 5319, 5357
residual plots
   MIXED procedure, 5347
residuals, details
   MIXED procedure, 5330
restricted maximum likelihood
   MIXED procedure, 5235
restricted maximum likelihood (REML)
   MIXED procedure, 5319, 5357
ridging
   MIXED procedure, 5257, 5319

sandwich estimator
   MIXED procedure, 5249
Satterthwaite method
   MIXED procedure, 5275
scaled residual
   MIXED procedure, 5286, 5331
Schwarz’s Bayesian information criterion
   example (MIXED), 5358, 5372, 5401

MIXED procedure, 5249, 5321, 5340
scoring
   MIXED procedure, 5248, 5257, 5353
Sidak’s adjustment
   MIXED procedure, 5267
simple effects
   MIXED procedure, 5271
simulation-based adjustment
   MIXED procedure, 5267
singularities
   MIXED procedure, 5354
spatial anisotropic exponential structure
   MIXED procedure, 5303
spatial covariance structure
   examples (MIXED), 5305
   MIXED procedure, 5304, 5312, 5353
split-plot design
   MIXED procedure, 5316, 5355
standard linear model
   MIXED procedure, 5236
studentized residual
   external, 5333
   internal, 5333
   MIXED procedure, 5285, 5333
subject effect
   MIXED procedure, 5261, 5297, 5303, 5355, 5360
summary of commands
   MIXED procedure, 5244

table names
   MIXED procedure, 5341
test components
   MIXED procedure, 5283
Toeplitz structure
   example (MIXED), 5397
   MIXED procedure, 5303
Tukey’s adjustment
   MIXED procedure, 5267
Type 1 estimation
   MIXED procedure, 5251
Type 1 testing
   MIXED procedure, 5277
Type 2 estimation
   MIXED procedure, 5251
Type 2 testing
   MIXED procedure, 5277
Type 3 estimation
   MIXED procedure, 5251
Type 3 testing
   MIXED procedure, 5277, 5340
unstructured correlations
   MIXED procedure, 5303
unstructured covariance matrix
MIXED procedure, 5303

V matrix
   MIXED procedure, 5298

VARCOMP procedure
   compared to MIXED procedure, 5237
   example (MIXED), 5373

variance components
   MIXED procedure, 5235, 5303

variance ratios
   MIXED procedure, 5287, 5296

Wald test
   mixed model (MIXED), 5322, 5364
   MIXED procedure, 5339, 5340

weighting
   MIXED procedure, 5312

zero variance component estimates
   MIXED procedure, 5352
### Syntax Index

- **ABSOLUTE** option  
  - PROC MIXED statement, 5247, 5338
- **ADJDFE=** option  
  - LSMEANS statement (MIXED), 5266
- **ADJUST=** option  
  - LSMEANS statement (MIXED), 5267
- **ALG=** option  
  - PRIOR statement (MIXED), 5292
- **ALPHA=** option  
  - ESTIMATE statement (MIXED), 5264  
  - LSMEANS statement (MIXED), 5268  
  - MODEL statement (MIXED), 5273  
  - PROC MIXED statement, 5247  
  - RANDOM statement (MIXED), 5295
- **ALPHAP=** option  
  - MODEL statement (MIXED), 5273
- **ANOVA** option  
  - PROC MIXED statement, 5247
- **ASYCORR** option  
  - PROC MIXED statement, 5247
- **ASYCOV** option  
  - PROC MIXED statement, 5248, 5374
- **AT MEANS** option  
  - LSMEANS statement (MIXED), 5268
- **AT** option  
  - LSMEANS statement (MIXED), 5268, 5269
- **BDATA=** option  
  - PRIOR statement (MIXED), 5292
- **BY** statement  
  - MIXED procedure, 5258
- **BYLEVEL** option  
  - LSMEANS statement (MIXED), 5269, 5270
- **CHISQ** option  
  - CONTRAST statement (MIXED), 5262  
  - MODEL statement (MIXED), 5273
- **CL** option  
  - ESTIMATE statement (MIXED), 5264  
  - LSMEANS statement (MIXED), 5269  
  - MODEL statement (MIXED), 5274  
  - RANDOM statement (MIXED), 5295
- **CL=** option  
  - PROC MIXED statement, 5248
- **CLASS** statement  
  - MIXED procedure, 5338  
  - MIXED procedure, 5258
- **CODE** statement  
  - MIXED procedure, 5259
- **CONTAIN** option  
  - MODEL statement (MIXED), 5274, 5275
- **CONTRAST** statement  
  - MIXED procedure, 5260
- **CONVF** option  
  - PROC MIXED statement, 5248, 5338
- **CONVG** option  
  - PROC MIXED statement, 5248, 5338
- **CONVH** option  
  - PROC MIXED statement, 5249, 5338
- **CORR** option  
  - LSMEANS statement (MIXED), 5269
- **CORRB** option  
  - MODEL statement (MIXED), 5274
- **COV** option  
  - LSMEANS statement (MIXED), 5269
- **COVB** option  
  - MODEL statement (MIXED), 5274
- **COVBI** option  
  - MODEL statement (MIXED), 5274
- **COVTEST** option  
  - PROC MIXED statement, 5249, 5339
- **CPUCOUNT=** option  
  - PERFORMANCE statement (MIXED procedure), 5289
- **DATA=** option  
  - PRIOR statement (MIXED), 5291  
  - PROC MIXED statement, 5249
- **DDF=** option  
  - MODEL statement (MIXED), 5274
- **DDFM=** option  
  - MODEL statement (MIXED), 5274
- **DF=** option  
  - CONTRAST statement (MIXED), 5262  
  - ESTIMATE statement (MIXED), 5264  
  - LSMEANS statement (MIXED), 5269
- **DFBW** option  
  - PROC MIXED statement, 5249
- **DIFF** option  
  - LSMEANS statement (MIXED), 5269
- **DIVISOR=** option  
  - ESTIMATE statement (MIXED), 5264
- **E** option  
  - CONTRAST statement (MIXED), 5262  
  - ESTIMATE statement (MIXED), 5264  
  - LSMEANS statement (MIXED), 5270  
  - MODEL statement (MIXED), 5277
E1 option
MODEL statement (MIXED), 5277

E2 option
MODEL statement (MIXED), 5277

E3 option
MODEL statement (MIXED), 5277
EFFECT= modifier
INFLUENCE option, MODEL statement (MIXED), 5279

EMPIRICAL option
MIXED, 5249

EQCONS= option
PARMS statement (MIXED), 5287

ESTIMATE statement
MIXED procedure, 5263

ESTIMATES modifier
INFLUENCE option, MODEL statement (MIXED), 5279

FLAT option
PRIOR statement (MIXED), 5292

FULLX option
MODEL statement (MIXED), 5269, 5277

G option
RANDOM statement (MIXED), 5295
GC option
RANDOM statement (MIXED), 5295
GCI option
RANDOM statement (MIXED), 5295
GCORR option
RANDOM statement (MIXED), 5295
GDATA= option
RANDOM statement (MIXED), 5295
GI option
RANDOM statement (MIXED), 5296
GRID= option
PRIOR statement (MIXED), 5292
GRIDT= option
PRIOR statement (MIXED), 5292
GROUP option
CONTRAST statement (MIXED), 5262
ESTIMATE statement (MIXED), 5264
GROUP= option
RANDOM statement (MIXED), 5296
REPEATED statement (MIXED), 5299
HLM option
REPEATED statement (MIXED), 5300
HLPS option
REPEATED statement (MIXED), 5300
HOLD= option
PARMS statement (MIXED), 5287
HTYPE= option
MODEL statement (MIXED), 5277

IC option
PROC MIXED statement, 5249
ID statement
MIXED procedure, 5265
IFACTOR= option
PRIOR statement (MIXED), 5292
INFLUENCE option
MODEL statement (MIXED), 5278
INFO option
PROC MIXED statement, 5250
INTERCEPT option
MODEL statement (MIXED), 5250
ITDETAILS option
PROC MIXED statement, 5250
ITER= modifier
INFLUENCE option, MODEL statement (MIXED), 5279

JEFFREYS option
PRIOR statement (MIXED), 5292
KEEP= modifier
INFLUENCE option, MODEL statement (MIXED), 5280

LCOMPONENTS option
MODEL statement (MIXED), 5283
LDATA= option
RANDOM statement (MIXED), 5296
REPEATED statement (MIXED), 5300
LOCAL= option
REPEATED statement (MIXED), 5300
LOCALW option
REPEATED statement (MIXED), 5302
LOGDETH option
PARMS statement (MIXED), 5287
LOGNOTE option
PROC MIXED statement, 5250
LOGNOTE= option
PRIOR statement (MIXED), 5292
LOGRBOUND= option
PRIOR statement (MIXED), 5292
LOWERB= option
PARMS statement (MIXED), 5288
LOWERTAILED option
ESTIMATE statement (MIXED), 5264
LSMEANS statement
MIXED procedure, 5265
LSMESTIMATE statement
MIXED procedure, 5271
MAXFUNC= option
PROC MIXED statement, 5250
MAXITER= option
PROC MIXED statement, 5251
METHOD= option
PROC MIXED statement, 5251, 5362
MIXED procedure, 5244
   INFLUENCE option, 5278
   PERFORMANCE statement, 5289
   syntax, 5244
MIXED procedure, BY statement, 5258
MIXED procedure, CLASS statement, 5338
MIXED procedure, CONTRAST statement, 5260
   CHISQ option, 5262
   DF= option, 5262
   E option, 5262
   GROUP option, 5262
   SINGULAR= option, 5262
   SUBJECT option, 5263
MIXED procedure, ESTIMATE statement, 5263
   ALPHA= option, 5264
   CL option, 5264
   DF= option, 5264
   DIVISOR= option, 5264
   E option, 5264
   GROUP option, 5264
   LOWERTAILED option, 5264
   SINGULAR= option, 5264
   SUBJECT option, 5265
   UPPERTAILED option, 5265
MIXED procedure, ID statement, 5265
MIXED procedure, LSMEANS statement, 5265, 5374
   ADJUST= option, 5267
   ALPHA= option, 5268
   AT MEANS option, 5268
   AT option, 5268, 5269
   BYLEVEL option, 5269, 5270
   CL option, 5269
   CORR option, 5269
   COV option, 5269
   DF= option, 5269
   DIFF option, 5269
   E option, 5270
   OBSMARGINS option, 5270
   PDIFF option, 5269, 5271
   SINGULAR= option, 5271
   SLICE= option, 5271
MIXED procedure, MODEL statement, 5272
   ALPHA= option, 5273
   ALPHAP= option, 5273
   CHISQ option, 5273
   CL option, 5274
   CONTAIN option, 5274, 5275
   CORRBB option, 5274
   COVB option, 5274
   COVBI option, 5274
   DDF= option, 5274
   DDFM= option, 5274
MIXED procedure, PRIOR statement, 5290
   ALG= option, 5292
BDATA= option, 5292
DATA= option, 5291
FLAT option, 5292
GRID= option, 5292
GRIDT= option, 5292
IFACTOR= option, 5292
JEFFREYS option, 5292
LOGNOTE= option, 5292
LOGRBOUND= option, 5292
NSAMPLE= option, 5293
NSEARCH= option, 5293
OUT= option, 5293
OUTG= option, 5293
OUTGT= option, 5293
PSEARCH option, 5293
PTRANS option, 5293
SEED= option, 5293
SFACTOR= option, 5293
TDATA= option, 5293
TRANS= option, 5294
UPDATE= option, 5294

MIXED procedure, PROC MIXED statement, 5246
ABSOLUTE option, 5247, 5338
ALPHA= option, 5247
ANOVA F option, 5247
ASYCORR option, 5247
ASYCOV option, 5248, 5374
CL= option, 5248
CONVF option, 5248, 5338
CONVG option, 5248, 5338
CONVH option, 5249, 5338
COVTEST option, 5249, 5339
DATA= option, 5249
DFBW option, 5249
IC option, 5249
INFO option, 5250
ITDETAILS option, 5250
LOGNOTE option, 5250
MAXFUNC= option, 5250
MAXITER= option, 5251
METHOD= option, 5251, 5362
MMEQ option, 5251, 5374
MMEQSOL option, 5251, 5374
NAMELEN= option, 5251
NOBOUND option, 5251
NOCLPRINT option, 5251
NOINFO option, 5252
NOITPRINT option, 5252
NOPROFILE option, 5252, 5319
ORD option, 5252
ORDER= option, 5252
PLOTS= option, 5252
RANKS option, 5257
RATIO option, 5257, 5339
RIDGE= option, 5257
SCORING= option, 5257
SIGHTER option, 5258
UPDATE option, 5258

MIXED procedure, RANDOM statement, 5237, 5294, 5355
ALPHA= option, 5295
CL option, 5295
G option, 5295
GC option, 5295
GCI option, 5295
GCORR option, 5295
GDATA= option, 5295
GI option, 5296
GROUP= option, 5296
LDATA= option, 5296
NOFULLZ option, 5296
RATIOS option, 5296
SOLUTION option, 5297
SUBJECT= option, 5261, 5297
TYPE= option, 5297
V option, 5298
VC option, 5298
VCI option, 5298
VCORR option, 5298
VI option, 5298

MIXED procedure, REPEATED statement, 5237, 5298, 5361
GROUP= option, 5299
HLM option, 5300
HLPS option, 5300
LDATA= option, 5300
LOCAL= option, 5300
LOCALW option, 5302
NONLOCALW option, 5302
R option, 5302
RC option, 5302
RCI option, 5302
RCORR option, 5302
RI option, 5302
SSCP option, 5302
SUBJECT= option, 5303
TYPE= option, 5303

MIXED procedure, WEIGHT statement, 5312

MIXED procedure, CLASS statement, 5258
REF= option, 5259
REF= variable option, 5259
TRUNCATE option, 5259

MIXED procedure, CODE statement, 5259

MIXED procedure, LSMESTIMATE statement, 5271

MIXED procedure, PROC MIXED statement
ORDER= option, 5252

MIXED procedure, SLICE statement, 5312

MIXED procedure, STORE statement, 5312
MMEQ option  
PROC MIXED statement, 5251, 5374

MMEQSOL option  
PROC MIXED statement, 5251, 5374

MODEL statement  
MIXED procedure, 5272

Modifiers of INFLUENCE option  
MODEL statement (MIXED), 5278

NAMELEN= option  
PROC MIXED statement, 5251

NOBOUND option  
PARMS statement (MIXED), 5288
PROC MIXED statement, 5251

NOCLPRINT option  
PROC MIXED statement, 5251

NOCONTAIN option  
MODEL statement (MIXED), 5283

NOFULLZ option  
RANDOM statement (MIXED), 5296

NOINFO option  
PROC MIXED statement, 5252

NOINT option  
MODEL statement (MIXED), 5283, 5325

NOIT option  
PARMS statement (MIXED), 5288

NOITPRINT option  
PROC MIXED statement, 5252

NONLOCALW option  
REPEATED statement (MIXED), 5302

NOSUMMARY option  
PROC MIXED statement, 5251

PDIFF option  
LSMEANS statement (MIXED), 5269, 5271

PERFORMANCE statement  
MIXED procedure, 5289

PLOTS= option  
PROC MIXED statement, 5252

PRIOR statement  
MIXED procedure, 5290
PROC MIXED statement, 5293

PROC MIXED statement, see MIXED procedure

PSEARCH option  
PRIOR statement (MIXED), 5293

PTRANS option  
PRIOR statement (MIXED), 5293

R option  
REPEATED statement (MIXED), 5302

RANDOM statement  
MIXED procedure, 5294

RANKS option  
PROC MIXED statement, 5257

RATIO option  
PROC MIXED statement, 5257, 5339

RATIOS option  
PARMS statement (MIXED), 5289
RANDOM statement (MIXED), 5296

RC option  
REPEATED statement (MIXED), 5302

RCI option  
REPEATED statement (MIXED), 5302

RCORR option  
REPEATED statement (MIXED), 5302

ORDER= option  
MODEL statement (MIXED), 5329
PROC MIXED statement, 5326
PROC MIXED statement, 5252

OUT= option  
PRIOR statement (MIXED), 5293
OUTG= option  
PRIOR statement (MIXED), 5293
OUTGT= option  
PRIOR statement (MIXED), 5293
OUTP= option  
MODEL statement (MIXED), 5374
OUTPRED= option  
MODEL statement (MIXED), 5284
OUTPREDM= option  
MODEL statement (MIXED), 5284

PARMS statement  
MIXED procedure, 5286, 5374

PARMSDATA= option  
PARMS statement (MIXED), 5289

PDATA= option  
PARMS statement (MIXED), 5289

PDIFF option  
LSMEANS statement (MIXED), 5269, 5271

PERFORMANCE statement  
MIXED procedure, 5289

PLOTS= option  
PROC MIXED statement, 5252

PRIOR statement  
MIXED procedure, 5290
PROC MIXED statement, 5293

PROC MIXED statement, see MIXED procedure

PSEARCH option  
PRIOR statement (MIXED), 5293

PTRANS option  
PRIOR statement (MIXED), 5293

R option  
REPEATED statement (MIXED), 5302

RANDOM statement  
MIXED procedure, 5294

RANKS option  
PROC MIXED statement, 5257

RATIO option  
PROC MIXED statement, 5257, 5339

RATIOS option  
PARMS statement (MIXED), 5289
RANDOM statement (MIXED), 5296

RC option  
REPEATED statement (MIXED), 5302

RCI option  
REPEATED statement (MIXED), 5302

RCORR option  
REPEATED statement (MIXED), 5302

OBSMARGINS option  
LSMEANS statement (MIXED), 5270

OLS option  
PARMS statement (MIXED), 5288

ORD option  
PROC MIXED statement, 5252