SAS/STAT® 14.2 User's Guide
The MIXED Procedure

This document is an individual chapter from *SAS/STAT® 14.2 User's Guide*.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2016. *SAS/STAT® 14.2 User's Guide*. Cary, NC: SAS Institute Inc.

SAS/STAT® 14.2 User's Guide

SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

November 2016

Chapter 78
The MIXED Procedure

Contents

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 6244.

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 6249.

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 6216.

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:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

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

The mixed model generalizes the standard linear model as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

Here, $\boldsymbol{\gamma}$ is an unknown vector of random-effects parameters with known design matrix $\mathbf{Z}$, and $\boldsymbol{\epsilon}$ 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 $\boldsymbol{\gamma}$ and $\boldsymbol{\epsilon}$ are Gaussian random variables that are uncorrelated and have expectations $\mathbf{0}$ and variances $\mathbf{G}$ and $\mathbf{R}$, respectively. The variance of $\mathbf{y}$ is thus

$$\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$$

Note that, when $\mathbf{R} = \sigma^2\mathbf{I}$ and $\mathbf{Z} = \mathbf{0}$, the mixed model reduces to the standard linear model.

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

- variance components
- compound symmetry (common covariance plus diagonal)
- unstructured (general covariance)
- autoregressive

- 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.

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 $\boldsymbol{\gamma}$ 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 VARCOMP procedure. PROC MIXED provides a wide variety of covariance structures, while PROC VARCOMP estimates only simple random effects. PROC MIXED carries out several analyses that are absent in PROC VARCOMP, 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

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:

```
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:

```
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 $\mathbf{X}$ matrix for the linear model. A column of 1s is also included as the first column of $\mathbf{X}$ to model a global intercept. There are no $\mathbf{Z}$ or $\mathbf{G}$ matrices for this model, and $\mathbf{R}$ is assumed to equal $\sigma^2\mathbf{I}$, where $\mathbf{I}$ is an 18×18 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 78.1–Figure 78.7.

The "Model Information" table in Figure 78.1 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.

Figure 78.1 Model Information

The Mixed Procedure

Model Information	
Data Set	WORK.HEIGHTS
Dependent Variable	Height
Covariance Structure	Diagonal
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Residual

The "Class Level Information" table in Figure 78.2 lists the levels of all variables specified in the CLASS statement. You can check this table to make sure that the data are correct.

Figure 78.2 Class Level Information

Class Level Information		
Class	**Levels**	**Values**
Family	4	1 2 3 4
Gender	2	F M

The "Dimensions" table in Figure 78.3 lists the sizes of relevant matrices. This table can be useful in determining CPU time and memory requirements.

Figure 78.3 Dimensions

Dimensions	
Covariance Parameters	1
Columns in X	15
Columns in Z	0
Subjects	1
Max Obs per Subject	18

The "Number of Observations" table in Figure 78.4 displays information about the sample size being processed.

Figure 78.4 Number of Observations

Number of Observations	
Number of Observations Read	18
Number of Observations Used	18
Number of Observations Not Used	0

The "Covariance Parameter Estimates" table in Figure 78.5 displays the estimate of σ^2 for the model.

Figure 78.5 Covariance Parameter Estimates

Covariance Parameter Estimates	
Cov Parm	**Estimate**
Residual	2.1000

The "Fit Statistics" table in Figure 78.6 lists several pieces of information about the fitted mixed model, including values derived from the computed value of the restricted/residual likelihood.

Figure 78.6 Fit Statistics

Fit Statistics	
-2 Res Log Likelihood	41.6
AIC (Smaller is Better)	43.6
AICC (Smaller is Better)	44.1
BIC (Smaller is Better)	43.9

The "Type 3 Tests of Fixed Effects" table in Figure 78.7 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 78.7 Tests of Fixed Effects

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Gender	1	10	17.63	0.0018
Family	3	10	5.90	0.0139
Family*Gender	3	10	2.89	0.0889

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:

```
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 $\mathbf{Z}$ matrix in the mixed model. The $\mathbf{X}$ matrix now consists of a column of 1s and the dummy variables for Gender.

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

The output from this analysis is as follows.

Figure 78.8 Model Information

The Mixed Procedure

Model Information	
Data Set	WORK.HEIGHTS
Dependent Variable	Height
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

The "Model Information" table in Figure 78.8 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; for more information, see the description of the DDFM= option.

Figure 78.9 Class Level Information

Class Level Information		
Class	**Levels**	**Values**
Family	4	1 2 3 4
Gender	2	F M

The "Class Level Information" table in Figure 78.9 is the same as before. The "Dimensions" table in Figure 78.10 displays the new sizes of the $\mathbf{X}$ and $\mathbf{Z}$ matrices.

Figure 78.10 Dimensions and Number of Observations

Dimensions	
Covariance Parameters	3
Columns in X	3
Columns in Z	12
Subjects	1
Max Obs per Subject	18

Number of Observations	
Number of Observations Read	18
Number of Observations Used	18
Number of Observations Not Used	0

The "Iteration History" table in Figure 78.11 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 78.11 REML Estimation Iteration History

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	74.11074833	
1	2	71.51614003	0.01441208
2	1	71.13845990	0.00412226
3	1	71.03613556	0.00058188
4	1	71.02281757	0.00001689
5	1	71.02245904	0.00000002
6	1	71.02245869	0.00000000

Convergence criteria met.

The "Covariance Parameter Estimates" table in Figure 78.12 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 σ^2.

Figure 78.12 Covariance Parameter Estimates (REML)

Covariance Parameter Estimates	
Cov Parm	**Estimate**
Family	2.4010
Family*Gender	1.7657
Residual	2.1668

The "Fit Statistics" table in Figure 78.13 contains basic information about the REML fit.

Figure 78.13 Fit Statistics

Fit Statistics	
-2 Res Log Likelihood	71.0
AIC (Smaller is Better)	77.0
AICC (Smaller is Better)	79.0
BIC (Smaller is Better)	75.2

The "Type 3 Tests of Fixed Effects" table in Figure 78.14 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 78.14 Type 3 Tests of Fixed Effects

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Gender	1	3	7.95	0.0667

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* > **;**
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 78.1 summarizes the basic functions and important *options* of each PROC MIXED statement. The syntax of each statement in Table 78.1 is described in the following sections in alphabetical order after the description of the PROC MIXED statement.

Table 78.1 Summary of PROC MIXED Statements

Statement	Description	Options
PROC MIXED	Invokes the procedure	DATA= specifies input data set, METHOD= specifies estimation method
BY	Performs multiple PROC MIXED analyses in one invocation	None

Table 78.1 *continued*

Statement	Description	Options
CLASS	Declares qualitative variables that create indicator variables in design matrices	None
CODE	Requests that the procedure write SAS DATA step code to a file or catalog entry	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
ID	Lists additional variables to be included in predicted values tables	None
MODEL	Specifies dependent variable and fixed effects, setting up **X**	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
RANDOM	Specifies random effects, setting up **Z** and **G**	SUBJECT= creates block-diagonality, TYPE= specifies covariance structure, S requests solution for random-effects parameters, G displays estimated **G**
REPEATED	Sets up **R**	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**
PARMS	Specifies a grid of initial values for the covariance parameters	HOLD= and NOITER hold the covariance parameters or their ratios constant, PARMSDATA= reads the initial values from a SAS data set
PRIOR	Performs a sampling-based Bayesian analysis for variance component models	NSAMPLE= specifies the sample size, SEED= specifies the starting seed
CONTRAST	Constructs custom hypothesis tests	E displays the **L** matrix coefficients
ESTIMATE	Constructs custom scalar estimates	CL produces confidence limits
LSMEANS	Computes least squares means for classification fixed effects	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

Table 78.1 *continued*

Statement	Description	Options
LSMESTIMATE	Provides custom hypothesis tests among the least squares means	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
SLICE	Performs a partitioned analysis of LS–means for an interaction	ADJUST= determines the method for multiple comparison adjustment of LS-mean differences, DIFF requests differences of LS-means
STORE	Saves the context and results of the analysis	LABEL= adds a custom label
WEIGHT	Specifies a variable by which to weight **R**	None

PROC MIXED Statement

PROC MIXED < *options* > **;**

The PROC MIXED statement invokes the MIXED procedure. Table 78.2 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 78.2 PROC MIXED Statement Options

Option	Description
Basic Options	
DATA=	Specifies input data set
METHOD=	Specifies the estimation method
NOPROFILE	Includes scale parameter in optimization
ORDER=	Determines the sort order of CLASS variables
Displayed Output	
ASYCORR	Displays asymptotic correlation matrix of covariance parameter estimates
ASYCOV	Displays asymptotic covariance matrix of covariance parameter estimates
CL	Requests confidence limits for covariance parameter estimates
COVTEST	Displays asymptotic standard errors and Wald tests for covariance parameters
IC	Displays a table of information criteria
ITDETAILS	Displays estimates and gradients added to “Iteration History”
LOGNOTE	Writes periodic status notes to the log
MMEQ	Displays mixed model equations
MMEQSOL	Displays the solution to the mixed model equations

Table 78.2 *continued*

Option	Description
NOCLPRINT	Suppresses "Class Level Information" completely or in parts
NOITPRINT	Suppresses "Iteration History" table
PLOTS=	Produces ODS statistical graphics
RANKS=	Displays a table of ranks of design matrices **X** and (**XZ**)
RATIO	Produces ratio of covariance parameter estimates with residual variance
Optimization Options	
MAXFUNC=	Specifies the maximum number of likelihood evaluations
MAXITER=	Specifies the maximum number of iterations
Computational Options	
CONVF	Requests and tunes the relative function convergence criterion
CONVG	Requests and tunes the relative gradient convergence criterion
CONVH	Requests and tunes the relative Hessian convergence criterion
DFBW	Selects between-within degree of freedom method
EMPIRICAL	Computes empirical ("sandwich") estimators
NOBOUND	Unbounds covariance parameter estimates
RIDGE=	Specifies starting value for minimum ridge value
SCORING=	Applies Fisher scoring where applicable

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 $1 - number$. The value of *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 6226.

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. For more information about this matrix, see the section "Covariance Parameter Estimates" on page 6242. 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

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

where $\nu = 2Z^2$, Z is the Wald statistic $\widehat{\sigma}^2/\mathrm{se}(\widehat{\sigma}^2)$, and the denominators are quantiles of the χ^2-distribution with ν 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

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

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

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

where f_k is the value of the objective function, and g_{jk} is the jth 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

$$\frac{\mathbf{g}_k{}'\mathbf{H}_k^{-1}\mathbf{g}_k}{|f_k|} \leq number$$

where f_k is the value of the objective function, $\mathbf{g}_k$ is the gradient (first derivative) of the objective function, and $\mathbf{H}_k$ is the Hessian (second derivative) of the objective function, all at iteration k.

If $\mathbf{H}_k$ is singular, then PROC MIXED uses the following relative criterion:

$$\frac{\mathbf{g}_k'\mathbf{g}_k}{|f_k|} \leq number$$

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=*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:

$$(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-}\left(\sum_{i=1}^{S}\mathbf{X}_i'\widehat{\mathbf{V}}_i{}^{-1}\widehat{\boldsymbol{\epsilon}}_i\widehat{\boldsymbol{\epsilon}}_i'\widehat{\mathbf{V}}_i{}^{-1}\mathbf{X}_i\right)(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-}$$

Here, $\widehat{\boldsymbol{\epsilon}}_i = y_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}$, S is the number of subjects, and matrices with an i subscript are those for the ith 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 78.26): Contrast, CorrB, CovB, Diffs, Estimates, InvCovB, LSMeans, 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 78.3.

Table 78.3 Information Criteria

Criterion	Formula	Reference
AIC	$-2\ell + 2d$	Akaike (1974)
AICC	$-2\ell + 2dn^*/(n^* - d - 1)$	Hurvich and Tsai (1989) Burnham and Anderson (1998)
HQIC	$-2\ell + 2d \log\log n$ for $n > 1$	Hannan and Quinn (1979)
BIC	$-2\ell + d \log n$ for $n > 0$	Schwarz (1978)
CAIC	$-2\ell + d(\log n + 1)$ for $n > 0$	Bozdogan (1987)

Here ℓ 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 $\mathbf{X}$. 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 $n \leq 1$, the value of the HQIC criterion is -2ℓ. When n=0, the values of the BIC and CAIC criteria are -2ℓ 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 $q + p$ where p is by default the sum of the Type 3 degrees of freedom associated with each fixed effect or the rank of X if you specify NOTEST option. The value of d is displayed in the "Information Criteria" table as the value of Parms variable; see Table 78.27.

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.

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 47, "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 $\widehat{\mathbf{G}}$ is nonsingular, the coefficient matrix and the right-hand side have the following form:

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

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

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

See the section "Estimating Fixed and Random Effects in the Mixed Model" on page 6223 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 $\widehat{\mathbf{G}}$ is nonsingular, the formula is the same as the preceding description of the MMEQ option. If $\widehat{\mathbf{G}}$ is singular, $\widehat{\boldsymbol{\beta}}$ and $\mathbf{G}\widehat{\boldsymbol{\tau}}$ 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 6223 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:

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

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 607 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 6249.

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 $\mathbf{X}$ 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:

```
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:

```
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=*number*
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

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 78.4 summarizes the *options* available in the CODE statement.

Table 78.4 CODE Statement Options

Option	Description
CATALOG=	Names the catalog entry where the generated code is saved
DUMMIES	Retains the dummy variables in the data set
ERROR	Computes the error function
FILE=	Names the file where the generated code is saved
FORMAT=	Specifies the numeric format for the regression coefficients
GROUP=	Specifies the group identifier for array names and statement labels
IMPUTE	Imputes predicted values for observations with missing or invalid covariates
LINESIZE=	Specifies the line size of the generated code
LOOKUP=	Specifies the algorithm for looking up CLASS levels
RESIDUAL	Computes residuals

For details about the syntax of the CODE statement, see the section "CODE Statement" on page 393 in Chapter 19, "Shared Concepts and Topics."

CONTRAST Statement

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 $\mathbf{L}'\boldsymbol{\phi} = \mathbf{0}$, where $\mathbf{L}' = (\mathbf{K}'\,\mathbf{M}')$ and $\boldsymbol{\phi}' = (\boldsymbol{\beta}'\,\boldsymbol{\gamma}')$, in several inference spaces. The inference space corresponds to the choice of $\mathbf{M}$. When $\mathbf{M} = \mathbf{0}$, 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 $\mathbf{M}$ 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 $\mathbf{M}$ 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 $\mathbf{L}'$ are specified in order and are separated by commas. The rows of the $\mathbf{K}'$ component of $\mathbf{L}'$ 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 $\mathbf{M}'$ component of $\mathbf{L}'$ 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 78.1):

```
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 $\mathbf{L}'$ 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):

```
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 78.1).

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

$$F = \frac{\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{bmatrix}' \mathbf{L}(\mathbf{L}'\widehat{\mathbf{C}}\mathbf{L})^{-1}\mathbf{L}' \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{bmatrix}}{r}$$

where $r = \mathrm{rank}(\mathbf{L}'\widehat{\mathbf{C}}\mathbf{L})$, and approximates its distribution with an F distribution. In this expression, $\widehat{\mathbf{C}}$ is an estimate of the generalized inverse of the coefficient matrix in the mixed model equations. For more information about this F statistic, see the section "Inference and Test Statistics" on page 6225.

The numerator degrees of freedom in the F approximation are $r = \mathrm{rank}(\mathbf{L}'\widehat{\mathbf{C}}\mathbf{L})$, 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 DF= option.

You can specify the following *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=*number*
: specifies the denominator degrees of freedom for the F test. For the degrees-of-freedom methods DDFM=BETWITHIN, DDFM=CONTAIN, and DDFM=RESIDUAL, 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. For DDFM=SATTERTHWAITE, DDFM=KENWARDROGER, and DDFM=KENWARDROGER2, the denominator degrees of freedom are computed separately for each contrast.

E
: requests that the **L** matrix coefficients for the contrast be displayed. The ODS name of the "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, CONTRAST statement coefficients on random effects are distributed equally across groups.

SINGULAR=*number*

tunes the estimability checking. If $\mathbf{v}$ is a vector, define ABS($\mathbf{v}$) to be the absolute value of the element of $\mathbf{v}$ with the largest absolute value. If ABS($\mathbf{K}' - \mathbf{K}'\mathbf{T}$) is greater than c**number* for any row of $\mathbf{K}'$ in the contrast, then $\mathbf{K}$ is declared nonestimable. Here $\mathbf{T}$ is the Hermite form matrix $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$, and c is ABS($\mathbf{K}'$) except when it equals 0, and then c is 1. The value for *number* must be between 0 and 1; the default is 1E–4.

SUBJECT *coeffs*

SUB *coeffs*

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.

ESTIMATE Statement

ESTIMATE *'label'* < *fixed-effect values . . .* >
< | *random-effect values . . .* > < / *options* > ;

The ESTIMATE statement is exactly like a CONTRAST statement, except only one-row $\mathbf{L}$ matrices are permitted. The actual estimate, $\mathbf{L}'\widehat{\mathbf{p}}$, is displayed along with its approximate standard error. An approximate t test that $\mathbf{L}'\widehat{\mathbf{p}} = 0$ 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 78.1) for various inference spaces:

```
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 78.5 summarizes the *options* available in the ESTIMATE statement.

Table 78.5 ESTIMATE Statement Options

Option	Description
ALPHA=	Specifies the confidence level
CL	Constructs *t*-type confidence limits
DF=	Specifies the degrees of freedom
DIVISOR=	Specifies a value by which to divide all coefficients
E	Displays the **L** matrix coefficients
GROUP	Sets up random-effect contrasts between different groups
LOWER	Performs lower-tailed tests
SINGULAR=	Tunes the estimability checking
SUBJECT	Sets up random-effect contrasts between different subjects
UPPER	Performs upper-tailed tests

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 – *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. 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 78.5 constructs the difference between the random slopes of the first two batches.

```
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 $\mathbf{L}\widehat{\boldsymbol{\beta}}$, where $\mathbf{L}$ is the coefficient matrix associated with the least squares mean and $\widehat{\boldsymbol{\beta}}$ is the estimate of the fixed-effects parameter vector (see the section "Estimating Fixed and Random Effects in the Mixed Model" on page 6223). The approximate standard errors for the LS-mean is computed as the square root of $\mathbf{L}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{L}'$.

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 $\mathbf{L}$ 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 6241).

Table 78.6 summarizes the *options* available in the LSMEANS statement. All LSMEANS *options* are subsequently discussed in alphabetical order.

Table 78.6 Summary of LSMEANS Statement Options

Option	Description
Construction and Computation of LS-Means	
AT	Modifies covariate value in computing LS-means
BYLEVEL	Computes separate margins
DIFF	Requests differences of LS-means
OM	Specifies weighting scheme for LS-mean computation
SINGULAR=	Tunes estimability checking
SLICE=	Partitions F tests (simple effects)
Degrees of Freedom and p-Values	
ADJDFE=	Determines whether to compute rowwise denominator degrees of freedom by using DDFM=SATTERTHWAITE, DDFM=KENWARDROGER, or DDFM=KENWARDROGER2
ADJUST=	Determines the method for multiple comparison adjustment of LS-mean differences
ALPHA=α	Determines the confidence level $(1-\alpha)$
DF=	Assigns specific value to degrees of freedom for tests and confidence limits
Statistical Output	
CL	Constructs confidence limits for means and or mean differences
CORR	Displays correlation matrix of LS-means
COV	Displays covariance matrix of LS-means
E	Prints the $\mathbf{L}$ matrix

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=SOURCE, 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, DDFM=KENWARDROGER, or DDFM=KENWARDROGER2 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), whereas 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 47, "The GLM Procedure," and Chapter 80, "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 α is 0.05, and you can change this value with the ALPHA= option in the LSMEANS statement.

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

$$P(|F(\widehat{q}) - (1-\alpha)| \leq \gamma) = 1 - \epsilon$$

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 ACC= and EPS= *sim-options* reset γ and ϵ, respectively; the NSAMP= *sim-option* sets the sample size directly; and the SEED= *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 "LSMEANS Statement" on page 3637 in Chapter 47, "The GLM Procedure."

ALPHA=*number*

requests that a t-type confidence interval be constructed for each of the LS-means with confidence level 1 – *number*. The value of *number* must be between 0 and 1; the default is 0.05.

AT *variable* **=** *value*

AT (*variable-list***)= (***value-list***)**

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:

```
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 $\overline{x_1}$ (the mean of X1) and for X2 is $\overline{x_2}$ (the mean of X2). However, for the first LSMEANS statement, the coefficient for X1*X2 is $\overline{x_1 x_2}$, but for the second LSMEANS statement, the coefficient is $\overline{x_1} \cdot \overline{x_2}$. The third LSMEANS statement sets the coefficient for X1 equal to 1.2 and leaves it at $\overline{x_2}$ 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 you specify the DDFM=SATTERTHWAITE, DDFM=KENWARDROGER, or DDFM=KENWARDROGER2 option in the MODEL statement. For these DDFM= methods, degrees of freedom are determined separately for each test; for more information, see the DDFM= option.

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.

PDIFF
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. For more information about this F test, see the section "Inference and Test Statistics" on page 6225.

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

LSMESTIMATE Statement

LSMESTIMATE *model-effect* < '*label*' > *values* < *divisor*=*n* >
< , . . . < '*label*' > *values* < *divisor*=*n* > >
< / *options* > ;

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

Table 78.7 summarizes the *options* available in the LSMESTIMATE statement.

Table 78.7 LSMESTIMATE Statement Options

Option	Description
Construction and Computation of LS-Means	
AT	Modifies covariate values in computing LS-means

Table 78.7 *continued*

Option	Description
BYLEVEL	Computes separate margins
DIVISOR=	Specifies a list of values to divide the coefficients
OM=	Specifies the weighting scheme for LS-means computation as determined by a data set
SINGULAR=	Tunes estimability checking
Degrees of Freedom and *p*-values	
ADJUST=	Determines the method for multiple-comparison adjustment of LS-means differences
ALPHA=α	Determines the confidence level ($1-\alpha$)
LOWER	Performs one-sided, lower-tailed inference
STEPDOWN	Adjusts multiple-comparison *p*-values further in a step-down fashion
TESTVALUE=	Specifies values under the null hypothesis for tests
UPPER	Performs one-sided, upper-tailed inference
Statistical Output	
CL	Constructs confidence limits for means and mean differences
CORR	Displays the correlation matrix of LS-means
COV	Displays the covariance matrix of LS-means
E	Prints the **L** matrix
ELSM	Prints the **K** matrix
JOINT	Produces a joint *F* or chi-square test for the LS-means and LS-means differences
PLOTS=	Requests graphs of means and mean comparisons
SEED=	Specifies the seed for computations that depend on random numbers
Generalized Linear Modeling	
CATEGORY=	Specifies how to construct estimable functions with multinomial data
EXP	Exponentiates and displays LS-means estimates
ILINK	Computes and displays estimates and standard errors of LS-means (but not differences) on the inverse linked scale

For details about the syntax of the LSMESTIMATE statement, see the section "LSMESTIMATE Statement" on page 477 in Chapter 19, "Shared Concepts and Topics."

MODEL Statement

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 6228 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 78.8 summarizes the *options* available in the MODEL statement. These are subsequently discussed in detail in alphabetical order.

Table 78.8 Summary of MODEL Statement Options

Option	Description
Model Building	
NOINT	Excludes fixed-effect intercept from model
Statistical Computations	
ALPHA=α	Determines the confidence level $(1-\alpha)$ for fixed effects
ALPHAP=α	Determines the confidence level $(1-\alpha)$ for predicted values
CHISQ	Requests chi-square tests
DDF=	Specifies denominator degrees of freedom (list)
DDFM=	Specifies the method for computing denominator degrees of freedom
HTYPE=	Selects the type of hypothesis test
INFLUENCE	Requests influence and case-deletion diagnostics
NOTEST	Suppresses hypothesis tests for the fixed effects
OUTP=	Specifies output data set for predicted values and related quantities
OUTPM=	Specifies output data set for predicted means and related quantities
RESIDUAL	Adds Pearson-type and studentized residuals to output data sets
VCIRY	Adds scaled marginal residual to output data sets
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=*number*
requests that a *t*-type confidence interval be constructed for each of the fixed-effects parameters with confidence level 1 – *number*. The value of *number* must be between 0 and 1; the default is 0.05.

ALPHAP=*number*
requests that a *t*-type confidence interval be constructed for the predicted values with confidence level 1 – *number*. The value of *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.

CL
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 ALPHA= option.

CONTAIN
has the same effect as the DDFM=CONTAIN option.

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

COVB
produces the approximate variance-covariance matrix of the fixed-effects parameter estimates $\widehat{\boldsymbol{\beta}}$. By default, this matrix equals $(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-}$ and results from sweeping $(\mathbf{X}\ \mathbf{y})'\widehat{\mathbf{V}}^{-1}(\mathbf{X}\ \mathbf{y})$ on all but its last pivot and removing the y border. The EMPIRICAL option in the 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 (DDFM=KENWARDROGER or DDFM=KENWARDROGER2), 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).

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

DDF=*value-list*
enables you to specify your own denominator degrees of freedom for the fixed effects. The *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:

```
model Y = A B A*B / ddf=3,.,4.7;
```

If you specify DDFM=SATTERTHWAITE, DDFM=KENWARDROGER, or DDFM=KENWARDROGER2, the DDF= option has no effect.

DDFM=
DDFM=CONTAIN
DDFM=BETWITHIN
DDFM=RESIDUAL
DDFM=SATTERTHWAITE
DDFM=KENWARDROGER<(FIRSTORDER)>
DDFM=KENWARDROGER<(LINEAR)>
DDFM=KENWARDROGER2

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

Table 78.9 lists syntax aliases for the degrees-of-freedom methods.

Table 78.9 Aliases for DDFM= Option

DDFM= Option	Alias
BETWITHIN	**BW**
CONTAIN	**CON**
KENWARDROGER	**KENROG**, **KR**
KENWARDROGER2	**KENROG2**, **KR2**
RESIDUAL	**RES**
SATTERTHWAITE	**SATTERTH**, **SAT**

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 $(\mathbf{X}\ \mathbf{Z})$ 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, $N - \mathrm{rank}(\mathbf{X}\ \mathbf{Z})$. 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 $\mathbf{Z}$ 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 or DDFM=KENWARDROGER2 might be a better option to try for this case.

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

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

Consider the one-dimensional case, and consider ℓ to be a vector defining an estimable linear combination of $\boldsymbol{\beta}$. The Satterthwaite degrees of freedom for the t statistic

$$t = \frac{\ell\widehat{\boldsymbol{\beta}}}{\sqrt{\ell\hat{\mathbf{C}}\ell'}}$$

is computed as

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

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

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

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

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

$$\nu_m = \frac{2(D_m)^2}{\mathbf{g}_m'\mathbf{A}\mathbf{g}_m}$$

where D_m is the mth diagonal element of $\mathbf{D}$ and $\mathbf{g}_m$ is the gradient of $\ell_m\mathbf{C}\ell_m'$ with respect to $\boldsymbol{\theta}$, evaluated at $\widehat{\boldsymbol{\theta}}$. Then let

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

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

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

provided $E > q$; otherwise ν 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 $\widehat{\boldsymbol{\gamma}}$ to $\widehat{\boldsymbol{\beta}}$ and change $\widehat{\mathbf{C}}$ 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 78.29, 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 $\widetilde{m}^{@}$ 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.

The DDFM=KENWARDROGER2 option specifies an improved approximation of the DDFM=KENWARDROGER method that uses a less biased precision estimator, as proposed by Kenward and Roger (2009). For an intrinsically linear covariance parameterization, this option produces the same precision estimator as that obtained using DDFM=KR(FIRSTORDER).

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, DDFM=KENWARDROGER, or DDFM=KENWARDROGER2, and a message is written to the log.

This method changes output in the following tables (listed in Table 78.26): 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 78.10 briefly describes important *suboptions* and their effect on the influence analysis.

Table 78.10 Summary of INFLUENCE Default and Suboptions

Description	Suboption
Compute influence diagnostics for individual observations	Default
Measure influence of sets of observations chosen according to a classification variable or effect	EFFECT=
Remove pairs of observations and report the results sorted by degree of influence	SIZE=2
Remove triples, quadruples of observations, etc.	SIZE=

Table 78.10 *continued*

Description	Suboption
Allow selection of individual observations, observations sharing specific levels of effects, and construction of tuples from specified subsets of observations	SELECT=
Update fixed effects and covariance parameters by refitting the mixed model, adding up to n iterations	ITER=$n > 0$
Compute influence diagnostics for the covariance parameters	ITER=$n > 0$
Update only fixed effects and the residual variance, if it is profiled	ITER=0
Add the reduced-data estimates to the data set created with ODS OUTPUT	ESTIMATES

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 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 6249.

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 $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}$ 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 $\widehat{\boldsymbol{\beta}}$ 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–Parm$p$, 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 6235).

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 6228 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:

```
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 $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}$ matrix changes rank or the $\mathbf{G}$ 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 78.11 and Table 78.12 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 6235 for details about the individual statistics.

Table 78.11 Statistics Computed with INFLUENCE Option, Noniterative Analysis (ITER=0)

Suboption	σ^2 Profiled	Statistic	Variable Name
Default	Yes	Observed value	Observed
		Predicted value	Predicted
		Marginal residual	Residual
		Leverage	Leverage
		PRESS residual	PRESSRes
		Internally studentized marginal residual	Student
		Externally studentized marginal residual	RStudent
		RMSE without deleted observations	RMSE
		Cook's *D*	CookD
		DFFITS	DFFITS
		CovRatio	COVRATIO
		(Restricted) likelihood distance	RLD, LD
Default	No	Observed value	Observed
		Predicted value	Predicted
		Marginal residual	Residual
		Leverage	Leverage
		PRESS residual	PRESSRes
		Internally studentized marginal residual	Student
		Cook's *D*	CookD
EFFECT=, **SIZE=**, or **KEEP=**	Yes	Observations in level (tuple)	Nobs
		PRESS statistic	PRESS
		Cook's *D*	CookD
		MDFFITS	MDFFITS
		CovRatio	COVRATIO
		COVTRACE	COVTRACE
		RMSE without deleted level (tuple)	RMSE
		(Restricted) likelihood distance	RLD, LD
EFFECT=, **SIZE=**, or **KEEP=**	No	Observations in level (tuple)	Nobs
		PRESS statistic	PRESS
		Cook's *D*	CookD

Table 78.12 Statistics Computed with INFLUENCE Option, Iterative Analysis (ITER=$n > 0$)

Suboption	Statistic	Variable Name
Default	Number of iterations	Iter
	Observed value	Observed

Table 78.12 *continued*

Suboption	Statistic	Variable Name
	Predicted value	Predicted
	Marginal residual	Residual
	Leverage	Leverage
	PRESS residual	PRESSres
	Internally studentized marginal residual	Student
	Externally studentized marginal residual	RStudent
	RMSE without deleted obs (if possible)	RMSE
	Cook's *D*	CookD
	DFFITS	DFFITS
	CovRatio	COVRATIO
	Cook's *D* CovParms	CookDCP
	CovRatio CovParms	COVRATIOCP
	MDFFITS CovParms	MDFFITSCP
	(Restricted) likelihood distance	RLD, LD
EFFECT=, **SIZE=**, or **KEEP=**	Observations in level (tuple)	Nobs
	Number of iterations	Iter
	PRESS statistic	PRESS
	RMSE without deleted level (tuple)	RMSE
	Cook's *D*	CookD
	MDFFITS	MDFFITS
	CovRatio	COVRATIO
	COVTRACE	COVTRACE
	Cook's *D* CovParms	CookDCP
	CovRatio CovParms	COVRATIOCP
	MDFFITS CovParms	MDFFITSCP
	(Restricted) likelihood distance	RLD, LD

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 $\mathbf{L}$ 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 $\mathbf{L}\boldsymbol{\beta} = \mathbf{0}$. The matrix $\mathbf{L}$ 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 $\mathbf{L}$ 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 $\mathbf{L}$ 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 $\mathbf{L}$ associated with the Type 1 and Type 3 estimable functions for factor A, $(b-1)$ tests for the rows of $\mathbf{L}$ associated with factor B, and a single test for the Type 1 and Type 3 coefficients associated with regressor X:

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

The denominator degrees of freedom associated with a row of $\mathbf{L}$ are the same as those in the corresponding "Tests of Fixed Effects" table, except for DDFM=KENWARDROGER, DDFM=KENWARDROGER2, and DDFM=SATTERTHWAITE. For these degrees-of-freedom methods, the denominator degrees of freedom are computed separately for each row of $\mathbf{L}$.

The ODS name of the table containing all requested component tests is LComponents. See Example 78.9 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 $(\mathbf{X}\ \mathbf{Z})$ as $\mathbf{L}$ matrices. Thus, predicted values from the original data are $\mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{Z}\widehat{\boldsymbol{\gamma}}$. Their approximate standard errors of prediction are formed from the quadratic form of $\mathbf{L}$ with $\widehat{\mathbf{C}}$ defined in the section "Statistical Properties" on page 6224. 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

$$\hat{\mathbf{m}} = \mathbf{X}_{\mathbf{m}}\hat{\boldsymbol{\beta}} + \hat{\mathbf{C}}_{\mathbf{m}}\hat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

where $\mathbf{m}$ represents a hypothetical realization of a missing data vector with associated design matrix $\mathbf{X}_{\mathbf{m}}$. The matrix $\mathbf{C}_{\mathbf{m}}$ is the model-based covariance matrix between $\mathbf{m}$ and the observed data $\mathbf{y}$, and other notation is as presented in the section "Mixed Models Theory" on page 6216.

The estimated prediction variance is as follows:

$$\widehat{\mathrm{Var}}(\hat{\mathbf{m}} - \mathbf{m}) = \hat{\mathbf{V}}_{\mathbf{m}} - \hat{\mathbf{C}}_{\mathbf{m}}\hat{\mathbf{V}}^{-1}\hat{\mathbf{C}}'_{\mathbf{m}} + [\mathbf{X}_{\mathbf{m}} - \hat{\mathbf{C}}_{\mathbf{m}}\hat{\mathbf{V}}^{-1}\mathbf{X}](\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-}[\mathbf{X}_{\mathbf{m}} - \hat{\mathbf{C}}_{\mathbf{m}}\hat{\mathbf{V}}^{-1}\mathbf{X}]'$$

where $\mathbf{V_m}$ is the model-based variance matrix of $\mathbf{m}$. 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 $\mathbf{Z}\widehat{\boldsymbol{\gamma}}$. 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 $\mathbf{X}\widehat{\boldsymbol{\beta}}$ in the OUTPM= data set, and standard errors are quadratic forms in the approximate variance-covariance matrix of $\widehat{\boldsymbol{\beta}}$ 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.

RESIDUAL

RESIDUALS

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 6249. For computational details about studentized and Pearson residuals in MIXED, see the section "Residual Diagnostics" on page 6233.

SINGCHOL=*number*

tunes the sensitivity in computing Cholesky roots. If a diagonal pivot element is less than *D***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***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

S

requests that a solution for the fixed-effects parameters be produced. Using notation from the section "Mixed Models Theory" on page 6216, the fixed-effects parameter estimates are $\widehat{\boldsymbol{\beta}}$ and their approximate standard errors are the square roots of the diagonal elements of $(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-}$. 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 or DDFM=KENWARDROGER2 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 6249.

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 78.1), 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:

```
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 78.3.)

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 78.13 summarizes the *options* available in the PARMS statement.

Table 78.13 PARMS Statement Options

Option	Description
HOLD=	Holds parameter values equal to the specified values
LOGDETH	Evaluates the log determinant of the Hessian matrix
LOWERB=	Specifies lower boundary constraints

Table 78.13 *continued*

Option	Description
NOBOUND	Removes boundary constraints on covariance parameters
NOITER	Performs inferences using the best value from the grid search
NOPRINT	Suppresses the "Parameter Search" table
NOPROFILE	Specifies a different computational method for the residual variance
OLS	Requests starting values corresponding to the usual general linear model
PARMSDATA=	Reads in covariance parameter values from a SAS data set
RATIOS	Indicates that ratios with the residual variance are specified
UPPERB=	Specifies upper boundary constraints

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.

LOWERB=*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 2×2 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–Covp*n* 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.

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–COVP*n*), 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 BETA*n*); 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 GAM*n*). 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 $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ and variance $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}$.

Table 78.14 summarizes the *options* available in the PRIOR statement.

Table 78.14 PRIOR Statement Options

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

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–Parm*n* 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–Parm*n* 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 $2 * number$ 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–CovP*n* 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 $\boldsymbol{\gamma}$ 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 6216, the purpose of the RANDOM statement is to define the $\mathbf{Z}$ matrix of the mixed model, the random effects in the $\boldsymbol{\gamma}$ 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 78.15 summarizes the *options* available in the RANDOM statement. All *options* are subsequently discussed in alphabetical order.

Table 78.15 Summary of RANDOM Statement Options

Option	Description
Construction of Covariance Structure	
GDATA=	Requests that the $\mathbf{G}$ matrix be read from a SAS data set
GROUP=	Varies covariance parameters by groups
LDATA=	Specifies data set with coefficient matrices for TYPE=LIN
NOFULLZ	Eliminates columns in $\mathbf{Z}$ corresponding to missing values
RATIOS	Indicates that ratios are specified in the GDATA= data set
SUBJECT=	Identifies the subjects in the model
TYPE=	Specifies the covariance structure
Statistical Output	
ALPHA=α	Determines the confidence level $(1-\alpha)$
CL	Requests confidence limits for predictors of random effects
G	Displays the estimated $\mathbf{G}$ matrix
GC	Displays the Cholesky root (lower) of estimated $\mathbf{G}$ matrix
GCI	Displays the inverse Cholesky root (lower) of estimated $\mathbf{G}$ matrix
GCORR	Displays the correlation matrix corresponding to estimated $\mathbf{G}$ matrix
GI	Displays the inverse of the estimated $\mathbf{G}$ matrix
SOLUTION	Displays solutions $\widehat{\boldsymbol{\gamma}}$ of the G-side random effects
V	Displays blocks of the estimated $\mathbf{V}$ matrix
VC	Displays the lower-triangular Cholesky root of blocks of the estimated $\mathbf{V}$ matrix
VCI	Displays the inverse Cholesky root of blocks of the estimated $\mathbf{V}$ matrix
VCORR	Displays the correlation matrix corresponding to blocks of the estimated $\mathbf{V}$ matrix

Table 78.15 *continued*

Option	Description
VI	Displays the inverse of the blocks of the estimated **V** matrix

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 – *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–Col*n* 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 78.4.)

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 $\mathbf{G}$ can also be input by using the PARMS statement with the HOLD= option.

GI

displays the inverse of the estimated $\mathbf{G}$ 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 $\mathbf{G}$. 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–Col*n* or Parm, Row, Col, Value. The Parm variable denotes which of the *number* coefficient matrices is currently being constructed, and the Row, Col1–Col*n*, 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 $\mathbf{Z}$ corresponding to missing levels of random effects involving CLASS variables. By default, these columns are included in $\mathbf{Z}$.

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

S

requests that the solution for the random-effects parameters be produced. Using notation from the section "Mixed Models Theory" on page 6216, these estimates are the empirical best linear unbiased predictors (EBLUPs) $\widehat{\boldsymbol{\gamma}} = \widehat{\mathbf{G}}\mathbf{Z}'\widehat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}})$. 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 $\widehat{\boldsymbol{\gamma}}$ displayed in the Estimate column; rather, they are the standard errors of predictions $\widehat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i$, where $\widehat{\boldsymbol{\gamma}}_i$ is the ith EBLUP and $\boldsymbol{\gamma}_i$ is the ith random-effect parameter.

SUBJECT=*effect*

SUB=*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 78.17 and Table 78.18. 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<=*value-list***>**

requests that blocks of the estimated **V** matrix be displayed. The first block determined by the SUBJECT= effect is the default displayed block. PROC MIXED displays entries that are 0 as blanks in the table.

You can optionally use the *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:

```
random int time / type=un subject=person v=1,3,7;
```

The ODS table name is V.

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

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

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

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

REPEATED Statement

REPEATED < *repeated-effect* > < / *options* > **;**

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 $\sigma^2\mathbf{I}$.

For many repeated measures models, no repeated effect is required in the REPEATED statement. Simply use the SUBJECT= option to define the blocks of **R** and the 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 78.16 summarizes the *options* available in the REPEATED statement. All *options* are subsequently discussed in alphabetical order.

Table 78.16 Summary of REPEATED Statement Options

Option	Description
Construction of Covariance Structure	
GROUP=	Defines an effect specifying heterogeneity in the R-side covariance structure

Table 78.16 *continued*

Option	Description
LDATA=	Specifies data set with coefficient matrices for TYPE=LIN
LOCAL	Requests that a diagonal matrix be added to **R**
LOCALW	Specifies that only the local effects are weighted
NONLOCALW	Specifies that only the nonlocal effects are weighted
SUBJECT=	Identifies the subjects in the R-side model
TYPE=	Specifies the R-side covariance structure
Statistical Output	
HLM	Produces a table of Hotelling-Lawley-McKeon statistics (McKeon 1974)
HLPS	Produces a table of Hotelling-Lawley-Pillai-Samson statistics (Pillai and Samson 1959)
R	Displays blocks of the estimated **R** matrix
RC	Display the Cholesky root (lower) of blocks of the estimated **R** matrix
RCI	Displays the inverse Cholesky root (lower) of blocks of the estimated **R** matrix
RCORR	Displays the correlation matrix corresponding to blocks of the estimated **R** matrix
RI	Displays the inverse of blocks of the estimated **R** matrix

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

GROUP=*effect*

GRP=*effect*

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.

HLM

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–Col*n* or Parm, Row, Col, Value. The Parm variable denotes which of the *number* coefficient matrices is currently being constructed, and the Row, Col1–Col*n*, 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 $\mathbf{R}$. With just the LOCAL option, this diagonal matrix equals $\sigma^2\mathbf{I}$, and σ^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

$$\sigma^2\mathrm{diag}[\exp(\mathbf{U}\boldsymbol{\delta})]$$

where $\mathbf{U}$ is the full-rank design matrix corresponding to the effects that you specify and $\boldsymbol{\delta}$ are the parameters that PROC MIXED estimates. An intercept is not included in $\mathbf{U}$ because it is accounted for by σ^2. PROC MIXED constructs the full-rank $\mathbf{U}$ in terms of 1s and –1s for classification effects. Be sure to scale continuous effects in $\mathbf{U}$ sensibly.

The LOCAL=POM(*POM-data-set*) option specifies the power-of-the-mean structure. This structure possesses a variance of the form $\sigma^2|\mathbf{x}_i'\boldsymbol{\beta}^*|^\theta$ for the ith observation, where $\mathbf{x}_i$ is the ith row of $\mathbf{X}$ (the design matrix of the fixed effects) and $\boldsymbol{\beta}^*$ 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 $\boldsymbol{\beta}^*$, where p is the number of columns of $\mathbf{X}$. 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 $\boldsymbol{\beta}^*$ corresponding to ordinary least squares is illustrated by the following statements:

```
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 $\boldsymbol{\beta}^*$. 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:

```
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:

```
repeated / type=cs subject=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 $\mathbf{R}$ 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 $\mathbf{R}$ 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 $\mathbf{R}$ matrix. The SUBJECT= option defines the blocks of $\mathbf{R}$, and the TYPE= option specifies the structure of these blocks. Valid values for *covariance-structure* and their descriptions are provided in Table 78.17 and Table 78.18. The default structure is VC.

Table 78.17 Covariance Structures

Structure	Description	Parms	(i, j) element
ANTE(1)	Antedependence	$2t-1$	$\sigma_i\sigma_j\prod_{k=i}^{j-1}\rho_k$
AR(1)	Autoregressive(1)	2	$\sigma^2\rho^{\lvert i-j\rvert}$
ARH(1)	Heterogeneous AR(1)	$t+1$	$\sigma_i\sigma_j\rho^{\lvert i-j\rvert}$
ARMA(1,1)	ARMA(1,1)	3	$\sigma^2[\gamma\rho^{\lvert i-j\rvert-1}1(i\neq j)+1(i=j)]$
CS	Compound symmetry	2	$\sigma_1+\sigma^2 1(i=j)$
CSH	Heterogeneous CS	$t+1$	$\sigma_i\sigma_j[\rho 1(i\neq j)+1(i=j)]$
FA(q)	Factor analytic	$\frac{q}{2}(2t-q+1)+t$	$\Sigma_{k=1}^{\min(i,j,q)}\lambda_{ik}\lambda_{jk}+\sigma_i^2 1(i=j)$
FA0(q)	No diagonal FA	$\frac{q}{2}(2t-q+1)$	$\Sigma_{k=1}^{\min(i,j,q)}\lambda_{ik}\lambda_{jk}$
FA1(q)	Equal diagonal FA	$\frac{q}{2}(2t-q+1)+1$	$\Sigma_{k=1}^{\min(i,j,q)}\lambda_{ik}\lambda_{jk}+\sigma^2 1(i=j)$
HF	Huynh-Feldt	$t+1$	$(\sigma_i^2+\sigma_j^2)/2+\lambda 1(i\neq j)$
LIN(q)	General linear	q	$\Sigma_{k=1}^{q}\theta_k\mathbf{A}_{ij}$
TOEP	Toeplitz	t	$\sigma_{\lvert i-j\rvert+1}$
TOEP(q)	Banded Toeplitz	q	$\sigma_{\lvert i-j\rvert+1}1(\lvert i-j\rvert<q)$
TOEPH	Heterogeneous TOEP	$2t-1$	$\sigma_i\sigma_j\rho_{\lvert i-j\rvert}$
TOEPH(q)	Banded hetero TOEP	$t+q-1$	$\sigma_i\sigma_j\rho_{\lvert i-j\rvert}1(\lvert i-j\rvert<q)$
UN	Unstructured	$t(t+1)/2$	σ_{ij}
UN(q)	Banded	$\frac{q}{2}(2t-q+1)$	$\sigma_{ij}1(\lvert i-j\rvert<q)$
UNR	Unstructured corrs	$t(t+1)/2$	$\sigma_i\sigma_j\rho_{\max(i,j)\min(i,j)}$
UNR(q)	Banded correlations	$\frac{q}{2}(2t-q+1)$	$\sigma_i\sigma_j\rho_{\max(i,j)\min(i,j)}$
UN@AR(1)	Direct product AR(1)	$t_1(t_1+1)/2+1$	$\sigma_{i_1j_1}\rho^{\lvert i_2-j_2\rvert}$
UN@CS	Direct product CS	$t_1(t_1+1)/2+1$	$\begin{cases}\sigma_{i_1j_1} & i_2=j_2\\ \sigma^2\sigma_{i_1j_1} & i_2\neq j_2\\ 0\le\sigma^2\le 1\end{cases}$
UN@UN	Direct product UN	$t_1(t_1+1)/2+$ $t_2(t_2+1)/2-1$	$\sigma_{1,i_1j_1}\sigma_{2,i_2j_2}$
VC	Variance components	q	$\sigma_k^2 1(i=j)$ and i corresponds to kth effect

In Table 78.17, "Parms" is the number of covariance parameters in the structure, t is the overall dimension of the covariance matrix, and $1(A)$ equals 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, $\rho_{ii} = 1$ 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 78.18 Spatial Covariance Structures

Structure	Description	Parms	(i, j) element
SP(EXP)(*c-list*)	Exponential	2	$\sigma^2 \exp\{-d_{ij}/\theta\}$
SP(EXPA)(*c-list*)	Anisotropic exponential	$2c + 1$	$\sigma^2 \prod_{k=1}^{c} \exp\{-\theta_k d(i, j, k)^{p_k}\}$
SP(EXPGA)(c_1 c_2)	2D exponential, geometrically anisotropic	4	$\sigma^2 \exp\{-d_{ij}(\theta, \lambda)/\rho\}$
SP(GAU)(*c-list*)	Gaussian	2	$\sigma^2 \exp\{-d_{ij}^2/\rho^2\}$
SP(GAUGA)(c_1 c_2)	2D Gaussian, geometrically anisotropic	4	$\sigma^2 \exp\{-d_{ij}(\theta, \lambda)^2/\rho^2\}$
SP(LIN)(*c-list*)	Linear	2	$\sigma^2(1 - \rho d_{ij})\ 1(\rho d_{ij} \le 1)$
SP(LINL)(*c-list*)	Linear log	2	$\sigma^2(1 - \rho \log(d_{ij}))$ $\times 1(\rho \log(d_{ij}) \le 1, d_{ij} > 0)$
SP(MATERN)(*c-list*)	Matérn	3	$\sigma^2 \frac{1}{\Gamma(\nu)} \left(\frac{d_{ij}}{2\rho}\right)^{\nu} 2K_\nu(d_{ij}/\rho)$
SP(MATHSW)(*c-list*)	Matérn (Handcock-Stein-Wallis)	3	$\sigma^2 \frac{1}{\Gamma(\nu)} \left(\frac{d_{ij}\sqrt{\nu}}{\rho}\right)^{\nu} 2K_\nu \left(\frac{2d_{ij}\sqrt{\nu}}{\rho}\right)$
SP(POW)(*c-list*)	Power	2	$\sigma^2 \rho^{d_{ij}}$
SP(POWA)(*c-list*)	Anisotropic power	$c + 1$	$\sigma^2 \rho_1^{d(i,j,1)} \rho_2^{d(i,j,2)} \ldots \rho_c^{d(i,j,c)}$
SP(SPH)(*c-list*)	Spherical	2	$\sigma^2[1 - (\frac{3d_{ij}}{2\rho}) + (\frac{d_{ij}^3}{2\rho^3})]\ 1(d_{ij} \le \rho)$
SP(SPHGA)(c_1 c_2)	2D Spherical, geometrically anisotropic	4	$\sigma^2[1 - (\frac{3d_{ij}(\theta,\lambda)}{2\rho}) + (\frac{d_{ij}(\theta,\lambda)^3}{2\rho^3})]$ $\times 1(d_{ij}(\theta, \lambda) \le \rho)$

In Table 78.18, *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 ith and jth vectors of these coordinates, which correspond to the ith and jth 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 kth coordinate, $k = 1, \ldots, c$, of the ith and jth 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 θ and scaling λ 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_ν is the modified Bessel function of the second kind of (real) order $\nu > 0$; the parameter ν governs the smoothness of the process (see below for more details).

Table 78.19 lists some examples of the structures in Table 78.17 and Table 78.18.

Table 78.19 Covariance Structure Examples

Description	Structure	Example
Variance components	VC (default)	$\begin{bmatrix} \sigma_B^2 & 0 & 0 & 0 \\ 0 & \sigma_B^2 & 0 & 0 \\ 0 & 0 & \sigma_{AB}^2 & 0 \\ 0 & 0 & 0 & \sigma_{AB}^2 \end{bmatrix}$
Compound symmetry	CS	$\begin{bmatrix} \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 \end{bmatrix}$
Unstructured	UN	$\begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} & \sigma_{42} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix}$
Banded main diagonal	UN(1)	$\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$
First-order autoregressive	AR(1)	$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$
Toeplitz	TOEP	$\begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{bmatrix}$
Toeplitz with two bands	TOEP(2)	$\begin{bmatrix} \sigma^2 & \sigma_1 & 0 & 0 \\ \sigma_1 & \sigma^2 & \sigma_1 & 0 \\ 0 & \sigma_1 & \sigma^2 & \sigma_1 \\ 0 & 0 & \sigma_1 & \sigma^2 \end{bmatrix}$
Spatial power	SP(POW)(*c*)	$\sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} \\ \rho^{d_{21}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} \\ \rho^{d_{31}} & \rho^{d_{32}} & 1 & \rho^{d_{34}} \\ \rho^{d_{41}} & \rho^{d_{42}} & \rho^{d_{43}} & 1 \end{bmatrix}$
Heterogeneous AR(1)	ARH(1)	$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$

Table 78.19 *continued*

Description	Structure	Example
First-order autoregressive moving average	ARMA(1,1)	$\sigma^2 \begin{bmatrix} 1 & \gamma & \gamma\rho & \gamma\rho^2 \\ \gamma & 1 & \gamma & \gamma\rho \\ \gamma\rho & \gamma & 1 & \gamma \\ \gamma\rho^2 & \gamma\rho & \gamma & 1 \end{bmatrix}$
Heterogeneous CS	CSH	$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$
First-order factor analytic	FA(1)	$\begin{bmatrix} \lambda_1^2+d_1 & \lambda_1\lambda_2 & \lambda_1\lambda_3 & \lambda_1\lambda_4 \\ \lambda_2\lambda_1 & \lambda_2^2+d_2 & \lambda_2\lambda_3 & \lambda_2\lambda_4 \\ \lambda_3\lambda_1 & \lambda_3\lambda_2 & \lambda_3^2+d_3 & \lambda_3\lambda_4 \\ \lambda_4\lambda_1 & \lambda_4\lambda_2 & \lambda_4\lambda_3 & \lambda_4^2+d_4 \end{bmatrix}$
Huynh-Feldt	HF	$\begin{bmatrix} \sigma_1^2 & \frac{\sigma_1^2+\sigma_2^2}{2}-\lambda & \frac{\sigma_1^2+\sigma_3^2}{2}-\lambda \\ \frac{\sigma_2^2+\sigma_1^2}{2}-\lambda & \sigma_2^2 & \frac{\sigma_2^2+\sigma_3^2}{2}-\lambda \\ \frac{\sigma_3^2+\sigma_1^2}{2}-\lambda & \frac{\sigma_3^2+\sigma_2^2}{2}-\lambda & \sigma_3^2 \end{bmatrix}$
First-order antedependence	ANTE(1)	$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_2 \\ \sigma_3\sigma_1\rho_2\rho_1 & \sigma_3\sigma_2\rho_2 & \sigma_3^2 \end{bmatrix}$
Heterogeneous Toeplitz	TOEPH	$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_2 & \sigma_1\sigma_4\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_1 & \sigma_2\sigma_4\rho_2 \\ \sigma_3\sigma_1\rho_2 & \sigma_3\sigma_2\rho_1 & \sigma_3^2 & \sigma_3\sigma_4\rho_1 \\ \sigma_4\sigma_1\rho_3 & \sigma_4\sigma_2\rho_2 & \sigma_4\sigma_3\rho_1 & \sigma_4^2 \end{bmatrix}$
Unstructured correlations	UNR	$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & \sigma_1\sigma_3\rho_{31} & \sigma_1\sigma_4\rho_{41} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & \sigma_2\sigma_4\rho_{42} \\ \sigma_3\sigma_1\rho_{31} & \sigma_3\sigma_2\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ \sigma_4\sigma_1\rho_{41} & \sigma_4\sigma_2\rho_{42} & \sigma_4\sigma_3\rho_{43} & \sigma_4^2 \end{bmatrix}$
Direct product AR(1)	UN@AR(1)	$\begin{bmatrix} \sigma_1^2 & \sigma_{21} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \otimes \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix} =$
		$\begin{bmatrix} \sigma_1^2 & \sigma_1^2\rho & \sigma_1^2\rho^2 & \sigma_{21} & \sigma_{21}\rho & \sigma_{21}\rho^2 \\ \sigma_1^2\rho & \sigma_1^2 & \sigma_1^2\rho & \sigma_{21}\rho & \sigma_{21} & \sigma_{21}\rho \\ \sigma_1^2\rho^2 & \sigma_1^2\rho & \sigma_1^2 & \sigma_{21}\rho^2 & \sigma_{21}\rho & \sigma_{21} \\ \sigma_{21} & \sigma_{21}\rho & \sigma_{21}\rho^2 & \sigma_2^2 & \sigma_2^2\rho & \sigma_2^2\rho^2 \\ \sigma_{21}\rho & \sigma_{21} & \sigma_{21}\rho & \sigma_2^2\rho & \sigma_2^2 & \sigma_2^2\rho \\ \sigma_{21}\rho^2 & \sigma_{21}\rho & \sigma_{21} & \sigma_2^2\rho^2 & \sigma_2^2\rho & \sigma_2^2 \end{bmatrix}$

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 78.17, σ_i^2 is the ith variance parameter, and ρ_k is the kth 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 78.17, ρ is the autoregressive parameter, γ models a moving-average component, and σ^2 is the residual variance. In the notation of Fuller (1976, p. 68), $\rho = \theta_1$ and

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

The example in Table 78.19 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 ρ and γ in this region the resulting covariance matrix is not positive definite. When the estimated value of ρ 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.

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 78.17, σ_i^2 is the ith variance parameter, and ρ 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 $\mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{D}$, where $\mathbf{\Lambda}$ 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 $\mathbf{\Lambda}$ in its upper-right corner (that is, the elements in the ith row and jth 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 kth coordinate ($k = 1, \ldots, c$) of the ith and jth observations in the input data set. There are $2c + 1$ parameters to be estimated: θ_k, p_k ($k = 1, \ldots, c$), and σ^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 θ_k and σ^2 estimates are close to 3, 4, 5, and 1, respectively. Then specify the following statements:

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

SP(EXPGA)(c_1 c_2) specify modification of the isotropic SP(EXP) covariance structure.

SP(GAUGA)(c_1 c_2) specify modification of the isotropic SP(GAU) covariance structure.

SP(SPHGA)(c_1 c_2) 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 $\mathbf{c} = [c_{i1}, c_{i2}]$, then it is isotropic in the coordinate system

$$\mathbf{Ac} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \mathbf{c} = \mathbf{c}^*$$

for a properly chosen angle θ and scaling factor λ. Elliptical isocorrelation contours are thereby transformed to spherical contours, adding two parameters to the respective isotropic covariance structures. Euclidean distances (see Table 78.18) are expressed in terms of $\mathbf{c}^*$.

The angle θ of the clockwise rotation is reported in radians, $0 \leq \theta \leq 2\pi$. The scaling parameter λ 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 θ, isotropy is achieved by compressing or magnifying distances in one coordinate by the factor λ.

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:

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

If λ 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-list*) | **SP(MATHSW)**(*c-list*) specifies covariance structures in the Matérn class of covariance functions (Matérn 1986). Two observations for the same subject (block of $\mathbf{R}$) that are Euclidean distance d_{ij} apart have covariance

$$\sigma^2 \frac{1}{\Gamma(\nu)} \left(\frac{d_{ij}}{2\rho} \right)^{\nu} 2K_\nu(d_{ij}/\rho) \qquad \nu > 0,\ \rho > 0$$

where K_ν is the modified Bessel function of the second kind of (real) order $\nu > 0$. The smoothness (continuity) of a stochastic process with covariance function in this class increases with ν. The Matérn class thus enables data-driven estimation of the smoothness properties. The covariance is identical to the exponential model for $\nu = 0.5$ (TYPE=SP(EXP)(*c-list*)), while for $\nu = 1$ the model advocated by Whittle (1954) results. As $\nu \to \infty$ 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),

$$\sigma^2 \frac{1}{\Gamma(\nu)} \left(\frac{d_{ij}\sqrt{\nu}}{\rho} \right)^{\nu} 2K_\nu \left(\frac{2d_{ij}\sqrt{\nu}}{\rho} \right)$$

Since computation of the function K_ν 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-list*) | **SP(POWA)**(*c-list*) specifies the spatial power structures. When the estimated value of ρ becomes negative, the computed covariance is multiplied by $\cos(\pi d_{ij})$ to account for the negativity.

TOEP<(*q*)> 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*)> specifies a heterogeneous banded Toeplitz structure. In Table 78.17, σ_i^2 is the *i*th variance parameter and ρ_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*)> 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*)> 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 σ_i^2. The parameter ρ_{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
- Var, indicating whether the measurement is a height or a weight
- Year, indicating the year of measurement
- 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:

```
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 Var in the MODEL statement. The preceding mean model consists of cell means for all combinations of VAR and 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.

Jennrich and Schluchter (1986) provide general information about the use of covariance structures, and Wolfinger (1996) presents details about many of the heterogeneous structures. Modeling with spatial covariance structures is discussed in many sources (Marx and Thompson 1987; Zimmerman and Harville 1991; Cressie 1993; Brownie, Bowman, and Burton 1993; Stroup, Baenziger, and Mulitze 1994; Brownie and Gumpertz 1997; Gotway and Stroup 1997; Chilès and Delfiner 1999; Schabenberger and Gotway 2005; Littell et al. 2006).

SLICE Statement

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 506 in Chapter 19, "Shared Concepts and Topics."

NOTE: Use the section "LSMEANS Statement" on page 458 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 6169.

STORE Statement

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 509 in Chapter 19, "Shared Concepts and Topics."

WEIGHT Statement

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 $\mathbf{X}'\mathbf{X}$ and $\mathbf{Z}'\mathbf{Z}$ with $\mathbf{X}'\mathbf{W}\mathbf{X}$ and $\mathbf{Z}'\mathbf{W}\mathbf{Z}$, where $\mathbf{W}$ is the diagonal weight matrix. If you specify a REPEATED statement, then the WEIGHT statement replaces $\mathbf{R}$ with $\mathbf{LRL}$, where $\mathbf{L}$ is a diagonal matrix with elements $\mathbf{W}^{-1/2}$. Observations with nonpositive or missing weights are not included in the PROC MIXED analysis.

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

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

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

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 σ^2.

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

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

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

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{y}$ denotes the vector of observed y_i's, $\mathbf{X}$ is the known matrix of x_{ij}'s, $\boldsymbol{\beta}$ is the unknown fixed-effects parameter vector, and $\boldsymbol{\epsilon}$ 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 ($^{-}$), determinant ($|\cdot|$), 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 ϵ is too restrictive. The mixed model extends the general linear model by allowing a more flexible specification of the covariance matrix of ϵ. In other words, it allows for both correlation and heterogeneous variances, although you still assume normality.

The mixed model is written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

where everything is the same as in the general linear model except for the addition of the known design matrix, $\mathbf{Z}$, and the vector of unknown *random-effects parameters,* $\boldsymbol{\gamma}$. The matrix $\mathbf{Z}$ can contain either continuous or dummy variables, just like $\mathbf{X}$. The name *mixed model* comes from the fact that the model contains both fixed-effects parameters, $\boldsymbol{\beta}$, and random-effects parameters, $\boldsymbol{\gamma}$. 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 $\boldsymbol{\gamma}$ and $\boldsymbol{\epsilon}$ are normally distributed with

$$\mathrm{E}\begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\epsilon} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

$$\mathrm{Var}\begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\epsilon} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}$$

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

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 $\mathbf{Z}$

containing dummy variables, $\mathbf{G}$ containing variance components in a diagonal structure, and $\mathbf{R} = \sigma^2\mathbf{I}_n$, where $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. The general linear model is a further special case with $\mathbf{Z} = \mathbf{0}$ and $\mathbf{R} = \sigma^2\mathbf{I}_n$.

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 $\mathbf{X}$ matrix is as follows:

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

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 $\mathbf{Z}$ and $\mathbf{G}$ matrices and let the $\mathbf{R}$ 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 $\mathbf{R}$ would then be as follows:

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

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:

```
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 $\mathbf{R}$ matrix: TYPE=CS specifies the compound symmetry structure, and SUBJECT=INDIV specifies the blocks of $\mathbf{R}$.

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

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

$$\mathbf{G} = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_1^2 & & \\ & & \ddots & \\ & & & \sigma_1^2 \end{bmatrix}$$

and $\mathbf{R} = \sigma^2 \mathbf{I}_n$. The $\mathbf{Z}$ matrix has $3s$ rows and s columns, and $\mathbf{G}$ is $s \times s$.

You can set up this model in PROC MIXED in two different but equivalent ways:

```
proc mixed;
   class indiv;
   model y = time;
   random indiv;
run;

proc mixed;
   class indiv;
   model y = time;
   random intercept / subject=indiv;
run;
```

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:

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

Here

$$\mathbf{R} = \sigma^2 \mathbf{I}_{24}$$

and $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{G}$ have the following form:

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

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

$$\mathbf{G} = \begin{bmatrix} \sigma^2_B & & & & & & & \\ & \sigma^2_B & & & & & & \\ & & \sigma^2_B & & & & & \\ & & & \sigma^2_B & & & & \\ & & & & \sigma^2_{AB} & & & \\ & & & & & \sigma^2_{AB} & & \\ & & & & & & \ddots & \\ & & & & & & & \sigma^2_{AB} \end{bmatrix}$$

where σ^2_B is the variance component for Block and σ^2_{AB} 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;
```

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

$$\mathbf{G} = \begin{bmatrix} \sigma^2_B & & & & & & & & & & & & & & & & & & & \\ & \sigma^2_{AB} & & & & & & & & & & & & & & & & & & & \\ & & \sigma^2_{AB} & & & & & & & & & & & & & & & & & & \\ & & & \sigma^2_{AB} & & & & & & & & & & & & & & & & & \\ & & & & \ddots & & & & & & & & & & & & & & & & \\ & & & & & \sigma^2_B & & & & & & & & & & & & & & & \\ & & & & & & \sigma^2_{AB} & & & & & & & & & & & & & & & \\ & & & & & & & \sigma^2_{AB} & & & & & & & & & & & & & \\ & & & & & & & & \sigma^2_{AB} & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & & & \end{bmatrix}$$

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 $\boldsymbol{\beta}$ as in the general linear model, but you have unknown parameters in $\boldsymbol{\gamma}$, $\mathbf{G}$, and $\mathbf{R}$ as well. Least squares is no longer the best method. *Generalized least squares* (GLS) is more appropriate, minimizing

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

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

In many situations, the best approach is to use *likelihood-based* methods, exploiting the assumption that $\boldsymbol{\gamma}$ and $\boldsymbol{\epsilon}$ 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 $\mathbf{G}$ and $\mathbf{R}$. The corresponding log-likelihood functions are as follows:

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

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

where $\mathbf{r} = \mathbf{y} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ and p is the rank of $\mathbf{X}$. 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 $\mathbf{H}$, the asymptotic theory of maximum likelihood (see Serfling 1980) shows that $2\mathbf{H}^{-1}$ is an asymptotic variance-covariance matrix of the estimated parameters of $\mathbf{G}$ and $\mathbf{R}$. 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 σ^2 is a part of your mixed model, it can usually be *profiled* out of the likelihood. This means solving analytically for the optimal σ^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 $\mathbf{R}$ equals $\sigma^2\mathbf{I}$ and when it has blocks with a compound symmetry, time series, or spatial structure. PROC MIXED does not profile the log likelihood when $\mathbf{R}$ 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 $\mathbf{G}$ and $\mathbf{R}$ (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 $\mathbf{G}$ and $\mathbf{R}$, which are denoted $\widehat{\mathbf{G}}$ and $\widehat{\mathbf{R}}$, respectively. To obtain estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, the standard method is to solve the *mixed model equations* (Henderson 1984):

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

The solutions can also be written as

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

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

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

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

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

$$\mathbf{P} = \begin{bmatrix} \mathbf{I} & \\ & \widehat{\mathbf{G}} \end{bmatrix}$$

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 $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{G}$, and $\mathbf{R}$. Even though $\mathbf{X}$ and $\mathbf{Z}$ 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 $\mathbf{G}$ and $\mathbf{R}$ 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 $\mathbf{G}$ and $\mathbf{R}$ are known, $\widehat{\boldsymbol{\beta}}$ is the *best linear unbiased estimator* (BLUE) of $\boldsymbol{\beta}$, and $\widehat{\boldsymbol{\gamma}}$ is the *best linear unbiased predictor* (BLUP) of $\boldsymbol{\gamma}$ (Searle 1971; Harville 1988, 1990; Robinson 1991; McLean, Sanders, and Stroup 1991). Here, "best" means minimum mean squared error. The covariance matrix of $(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}, \widehat{\boldsymbol{\gamma}}-\boldsymbol{\gamma})$ is

$$\mathbf{C} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1} \end{bmatrix}^{-}$$

where $^-$ denotes a generalized inverse (see Searle 1971).

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

$$\widehat{\mathbf{C}} = \begin{bmatrix} \mathbf{X}'\widehat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\widehat{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\widehat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\widehat{\mathbf{R}}^{-1}\mathbf{Z}+\widehat{\mathbf{G}}^{-1} \end{bmatrix}^{-}$$

as the approximate variance-covariance matrix of $(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}, \widehat{\boldsymbol{\gamma}}-\boldsymbol{\gamma})$. 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 $\widehat{\mathbf{C}}$ can also be written as

$$\widehat{\mathbf{C}} = \begin{bmatrix} \widehat{\mathbf{C}}_{11} & \widehat{\mathbf{C}}_{21}' \\ \widehat{\mathbf{C}}_{21} & \widehat{\mathbf{C}}_{22} \end{bmatrix}$$

where

$$\widehat{\mathbf{C}}_{11} = (\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-}$$
$$\widehat{\mathbf{C}}_{21} = -\widehat{\mathbf{G}}\mathbf{Z}'\widehat{\mathbf{V}}^{-1}\mathbf{X}\widehat{\mathbf{C}}_{11}$$
$$\widehat{\mathbf{C}}_{22} = (\mathbf{Z}'\widehat{\mathbf{R}}^{-1}\mathbf{Z}+\widehat{\mathbf{G}}^{-1})^{-1} - \widehat{\mathbf{C}}_{21}\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{Z}\widehat{\mathbf{G}}$$

Note that $\widehat{\mathbf{C}}_{11}$ is the familiar estimated generalized least squares formula for the variance-covariance matrix of $\widehat{\boldsymbol{\beta}}$.

As a cautionary note, $\widehat{\mathbf{C}}$ tends to underestimate the true sampling variability of $(\widehat{\boldsymbol{\beta}}\ \widehat{\boldsymbol{\gamma}})$ because no account is made for the uncertainty in estimating $\mathbf{G}$ and $\mathbf{R}$. 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 or DDFM=KENWARDROGER2 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 χ^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 χ^2 statistic computed in this fashion has a large-sample χ^2 distribution that is χ^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 χ^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:

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

The estimability requirement (Searle 1971) applies only to the $\boldsymbol{\beta}$ portion of $\mathbf{L}$, because any linear combination of $\boldsymbol{\gamma}$ is estimable. Such a formulation in terms of a general $\mathbf{L}$ 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 $\mathbf{L}$ matrices. Typically, inference on fixed effects is the focus, and, in this case, the $\boldsymbol{\gamma}$ portion of $\mathbf{L}$ is assumed to contain all 0s.

Statistical inferences are obtained by testing the hypothesis

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

or by constructing point and interval estimates.

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

$$t = \frac{\mathbf{L}\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{bmatrix}}{\sqrt{\mathbf{L}\widehat{\mathbf{C}}\mathbf{L}'}}$$

Under the assumed normality of $\boldsymbol{\gamma}$ and $\boldsymbol{\epsilon}$, 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 $\widehat{\nu}$ being the approximate degrees of freedom, the associated confidence interval is

$$\mathbf{L}\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{bmatrix} \pm t_{\widehat{\nu},\alpha/2}\sqrt{\mathbf{L}\widehat{\mathbf{C}}\mathbf{L}'}$$

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

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

$$F = \frac{\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{bmatrix}' \mathbf{L}'(\mathbf{L}\widehat{\mathbf{C}}\mathbf{L}')^{-1}\mathbf{L}\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{bmatrix}}{r}$$

where $r = \mathrm{rank}(\mathbf{L}\widehat{\mathbf{C}}\mathbf{L}')$. Analogous to t, F in general has an approximate F distribution with r numerator degrees of freedom and $\widehat{\nu}$ 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 χ^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 $\mathbf{L}$ denote the matrix of estimable functions for the hypothesis $H\colon \mathbf{L}\boldsymbol{\beta} = \mathbf{0}$, where $\boldsymbol{\beta}$ are the fixed-effects parameters. Let $\mathbf{M} = \mathbf{L}'(\mathbf{L}\mathbf{L}')^{-}\mathbf{L}$, and suppose that $\widehat{\mathbf{C}}$ denotes the estimated variance-covariance matrix of $\widehat{\boldsymbol{\beta}}$ (see the section "Statistical Properties" for the construction of $\widehat{\mathbf{C}}$).

The ANOVAF F statistics are computed as

$$F_A = \widehat{\boldsymbol{\beta}}'\mathbf{L}'\left(\mathbf{L}\mathbf{L}'\right)^{-1}\mathbf{L}\widehat{\boldsymbol{\beta}}\Big/t_1 = \widehat{\boldsymbol{\beta}}'\mathbf{M}\widehat{\boldsymbol{\beta}}\Big/t_1$$

Notice that this is a modification of the usual F statistic where $(\mathbf{L}\widehat{\mathbf{C}}\mathbf{L}')^{-1}$ is replaced with $(\mathbf{L}\mathbf{L}')^{-1}$ and $\mathrm{rank}(\mathbf{L})$ is replaced with $t_1 = \mathrm{trace}(\mathbf{M}\widehat{\mathbf{C}})$; see, for example, Brunner, Domhof, and Langer (2002, Sec. 5.4).

The p-values for this statistic are computed from either an F_{ν_1,ν_2} or an $F_{\nu_1,\infty}$ distribution. The respective degrees of freedom are determined by the MIXED procedure as follows:

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

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

$$\nu_2 = \begin{cases} \max\{\min\{\nu_2^*, df_e\}, 1\} & \mathbf{g}'\mathbf{A}\mathbf{g} > 1\text{E}3 \times \text{MACEPS} \\ 1 & \text{otherwise} \end{cases}$$

The term $\mathbf{g}'\mathbf{A}\mathbf{g}$ in the term ν_2^* for the denominator degrees of freedom is based on approximating $\text{Var}[\text{trace}(\mathbf{M}\widehat{\mathbf{C}})]$ 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 $\mathbf{g} = [g_1, \ldots, g_q]$ contains the partial derivatives

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

and $\mathbf{A}$ is the asymptotic variance-covariance matrix of the covariance parameter estimates (ASYCOV option).

PROC MIXED reports ν_1 and ν_2 as "NumDF" and "DenDF" under the "ANOVA F" heading in the output. The corresponding p-values are denoted as "Pr > F(DDF)" for F_{ν_1,ν_2} and "Pr > F(infty)" for $F_{\nu_1,\infty}$, 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:

```
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);
      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 Pr > F(DDF) in the MIXED output. The subsequent DATA step computes the *p*-values that correspond to Pr > F(infty) in the PROC MIXED output.

Parameterization of Mixed Models

Recall that a mixed model is of the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

where $\mathbf{y}$ represents univariate data, $\boldsymbol{\beta}$ is an unknown vector of fixed effects with known model matrix $\mathbf{X}$, $\boldsymbol{\gamma}$ is an unknown vector of random effects with known model matrix $\mathbf{Z}$, and $\boldsymbol{\epsilon}$ 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 $\mathbf{X}$, and each effect in the RANDOM statement generates one or more columns in the model matrix $\mathbf{Z}$. 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 $\mathbf{X}$ and $\mathbf{Z}$.

Intercept

By default, all models automatically include a column of 1s in **X** to estimate a fixed-effect intercept parameter μ. 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 X*X X*X*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 78.20 is an example.

Table 78.20 Example of Main Effects

Data		**I**	A		B		
A	B	μ	A1	A2	B1	B2	B3
1	1	1	1	0	1	0	0
1	2	1	1	0	0	1	0
1	3	1	1	0	0	0	1
2	1	1	0	1	1	0	0
2	2	1	0	1	0	1	0
2	3	1	0	1	0	0	1

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 78.21). Empty columns (that would contain all 0s) are not generated for **X**, but they are for **Z**.

Table 78.21 Example of Interaction Effects

Data		I	A		B			A*B					
A	B	μ	A1	A2	B1	B2	B3	A1B1	A1B2	A1B3	A2B1	A2B2	A2B3
1	1	1	1	0	1	0	0	1	0	0	0	0	0
1	2	1	1	0	0	1	0	0	1	0	0	0	0
1	3	1	1	0	0	0	1	0	0	1	0	0	0
2	1	1	0	1	1	0	0	0	0	0	1	0	0
2	2	1	0	1	0	1	0	0	0	0	0	1	0
2	3	1	0	1	0	0	1	0	0	0	0	0	1

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):

```
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 78.22).

Table 78.22 Example of Nested Effects

Data		I	A		B(A)					
A	B	μ	A1	A2	B1A1	B2A1	B3A1	B1A2	B2A2	B3A2
1	1	1	1	0	1	0	0	0	0	0
1	2	1	1	0	0	1	0	0	0	0
1	3	1	1	0	0	0	1	0	0	0
2	1	1	0	1	0	0	0	1	0	0
2	2	1	0	1	0	0	0	0	1	0
2	3	1	0	1	0	0	0	0	0	1

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 78.23).

Table 78.23 Example of Continuous-Nesting-Class Effects

Data		**I**	A		X(A)	
X	A	μ	A1	A2	X(A1)	X(A2)
21	1	1	1	0	21	0
24	1	1	1	0	24	0
22	1	1	1	0	22	0
28	2	1	0	1	0	28
19	2	1	0	1	0	19
23	2	1	0	1	0	23

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 78.24 shows an example.

Table 78.24 Example of Continuous-by-Class Effects

Data		**I**	**X**	A		X*A	
X	A	μ	X	A1	A2	X*A1	X*A2
21	1	1	21	1	0	21	0
24	1	1	24	1	0	24	0
22	1	1	22	1	0	22	0
28	2	1	28	0	1	0	28
19	2	1	19	0	1	0	19
23	2	1	23	0	1	0	23

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

$$\begin{array}{llll} A_1B_1C_1D_1 \rightarrow & A_1B_2C_1D_1 \rightarrow & A_2B_1C_1D_1 \rightarrow & A_2B_2C_1D_1 \rightarrow \\ A_1B_1C_1D_2 \rightarrow & A_1B_2C_1D_2 \rightarrow & A_2B_1C_1D_2 \rightarrow & A_2B_2C_1D_2 \rightarrow \\ A_1B_1C_2D_1 \rightarrow & A_1B_2C_2D_1 \rightarrow & A_2B_1C_2D_1 \rightarrow & A_2B_2C_2D_1 \rightarrow \\ A_1B_1C_2D_2 \rightarrow & A_1B_2C_2D_2 \rightarrow & A_2B_1C_2D_2 \rightarrow & A_2B_2C_2D_2 \end{array}$$

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 $\mathbf{X}$ constructed than there are degrees of freedom for the effect. Thus, there are linear dependencies among the columns of $\mathbf{X}$. 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 $\widetilde{\mathbf{e}} = \mathbf{PY}$, where $\mathbf{P}$ is a projection matrix, possibly an oblique projector. A typical element $\widetilde{e}_i$ with variance v_i and estimated variance $\widehat{v}_i$ is said to be *standardized* as

$$\frac{\widetilde{e}_i}{\sqrt{\mathrm{Var}[\widetilde{e}_i]}} = \frac{\widetilde{e}_i}{\sqrt{v_i}}$$

and *studentized* as

$$\frac{\widetilde{e}_i}{\sqrt{\widehat{v}_i}}$$

External studentization uses an estimate of $\mathrm{Var}[\widetilde{e}_i]$ that does not involve the ith 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., $\widetilde{e}_i / \sqrt{\widehat{\mathrm{Var}}[Y_i]}$, are referred to as Pearson-type residuals.

Marginal and Conditional Residuals

The marginal and conditional means in the linear mixed model are $\mathrm{E}[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta}$ and $\mathrm{E}[\mathbf{Y}|\boldsymbol{\gamma}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}$, respectively. Accordingly, the vector $\mathbf{r}_m$ of marginal residuals is defined as

$$\mathbf{r}_m = \mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}$$

and the vector $\mathbf{r}_c$ of conditional residuals is

$$\mathbf{r}_c = \mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}} - \mathbf{Z}\widehat{\boldsymbol{\gamma}} = \mathbf{r}_m - \mathbf{Z}\widehat{\boldsymbol{\gamma}}$$

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

$$\widehat{\mathrm{Var}}[\mathbf{r}_m] = \widehat{\mathbf{V}} - \mathbf{Q}$$
$$\widehat{\mathrm{Var}}[\mathbf{r}_c] = \mathbf{K}(\widehat{\mathbf{V}} - \mathbf{Q})\mathbf{K}'$$

For an individual observation the raw, studentized, and Pearson-type residuals computed by the MIXED procedure are given in Table 78.25.

Table 78.25 Residual Types Computed by the MIXED Procedure

Type of Residual	Marginal	Conditional
Raw	$r_{mi} = Y_i - \mathbf{x}_i'\widehat{\boldsymbol{\beta}}$	$r_{ci} = r_{mi} - \mathbf{z}_i'\widehat{\boldsymbol{\gamma}}$
Studentized	$r_{mi}^{student} = \frac{r_{mi}}{\sqrt{\widehat{\mathrm{Var}}[r_{mi}]}}$	$r_{ci}^{student} = \frac{r_{ci}}{\sqrt{\widehat{\mathrm{Var}}[r_{ci}]}}$
Pearson	$r_{mi}^{pearson} = \frac{r_{mi}}{\sqrt{\widehat{\mathrm{Var}}[Y_i]}}$	$r_{ci}^{pearson} = \frac{r_{ci}}{\sqrt{\widehat{\mathrm{Var}}[Y_i \vert \boldsymbol{\gamma}]}}$

When the OUTPM= option is specified in addition to the RESIDUAL option in the MODEL statement, $r_{mi}^{student}$ and $r_{mi}^{pearson}$ are added to the data set as variables Resid, StudentResid, and PearsonResid, respectively. When the OUTP= option is specified, $r_{ci}^{student}$ and $r_{ci}^{pearson}$ 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 $\mathrm{Var}[\mathbf{Y}] = \mathbf{V}$ and $\mathbf{C}'\mathbf{C} = \mathbf{V}$, then $\mathbf{C}'^{-1}\mathbf{Y}$ 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 $\widehat{\mathbf{C}}$ denote the Cholesky root of $\widehat{\mathbf{V}}$, so that $\widehat{\mathbf{C}}'\widehat{\mathbf{C}} = \widehat{\mathbf{V}}$, and define

$$\mathbf{Y}_c = \widehat{\mathbf{C}}'^{-1}\mathbf{Y}$$
$$\mathbf{r}_{m(c)} = \widehat{\mathbf{C}}'^{-1}\mathbf{r}_m$$

By analogy with other scalings, the inverse Cholesky decomposition can also be applied to the residual vector, $\widehat{\mathbf{C}}'^{-1}\mathbf{r}_m$, although $\mathbf{V}$ is not the variance-covariance matrix of $\mathbf{r}_m$.

To diagnose whether the covariance structure of the model has been specified correctly can be difficult based on $\mathbf{Y}_c$, since the inverse Cholesky transformation affects the expected value of $\mathbf{Y}_c$. You can draw on $\mathbf{r}_{m(c)}$ 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, $\mathbf{Y}_c$ is added as variable ScaledDep and $\mathbf{r}_{m(c)}$ 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 6216 for relevant expressions and definitions. The parameter vector $\boldsymbol{\theta}$ denotes all unknown parameters in the $\mathbf{R}$ and $\mathbf{G}$ 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 $\boldsymbol{\beta}$, obtained from all observations except those in the set U is denoted $\widehat{\boldsymbol{\beta}}_{(U)}$. In case of a matrix $\mathbf{A}$, the notation $\mathbf{A}_{(U)}$ represents the matrix with the rows in U removed; these rows are collected in $\mathbf{A}_U$. If $\mathbf{A}$ is symmetric, then notation $\mathbf{A}_{(U)}$ implies removal of rows and columns. The vector Y_U comprises the responses of the data points being removed, and $\mathbf{V}_{(U)}$ is the variance-covariance matrix of the remaining observations. When $k = 1$, lowercase notation emphasizes that single points are removed, such as $\mathbf{A}_{(u)}$.

Managing the Covariance Parameters

An important component of influence diagnostics in the mixed model is the estimated variance-covariance matrix $\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$. To make the dependence on the vector of covariance parameters explicit, write it as $\mathbf{V}(\boldsymbol{\theta})$. If one parameter, σ^2, is profiled or factored out of $\mathbf{V}$, the remaining parameters are denoted as $\boldsymbol{\theta}^*$. Notice that in a model where $\mathbf{G}$ is diagonal and $\mathbf{R} = \sigma^2\mathbf{I}$, the parameter vector $\boldsymbol{\theta}^*$ contains the ratios of each variance component and σ^2 (see Wolfinger, Tobias, and Sall 1994). When 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 $\widehat{\boldsymbol{\theta}}_{(U)} \equiv \widehat{\boldsymbol{\theta}}$.

2. If the residual variance is profiled, then $\widehat{\boldsymbol{\theta}}^*_{(U)} \equiv \widehat{\boldsymbol{\theta}}^*$ and $\widehat{\sigma}^2_{(U)} \neq \widehat{\sigma}^2$. Influence statistics such as Cook's D and internally studentized residuals are based on $\mathbf{V}(\widehat{\boldsymbol{\theta}})$, whereas externally studentized residuals and the DFFITS statistic are based on $\mathbf{V}(\widehat{\boldsymbol{\theta}}_U) = \sigma^2_{(U)}\mathbf{V}(\widehat{\boldsymbol{\theta}}^*)$. In a random components model with uncorrelated errors, for example, the computation of $\mathbf{V}(\widehat{\boldsymbol{\theta}}_U)$ involves scaling of $\widehat{\mathbf{G}}$ and $\widehat{\mathbf{R}}$ by the full-data estimate $\widehat{\sigma}^2$ and multiplying the result with the reduced-data estimate $\widehat{\sigma}^2_{(U)}$.

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, $\mathbf{V}(\widehat{\boldsymbol{\theta}}_U)$ is evaluated at the reduced-data parameter estimates but computed for the entire data set. The matrix $\mathbf{V}_{(U)}(\widehat{\boldsymbol{\theta}}_{(U)})$, 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 $\boldsymbol{\theta}$ is updated, whether the residual variance is profiled or not. The matrices to be distinguished here are $\mathbf{V}(\widehat{\boldsymbol{\theta}})$, $\mathbf{V}(\widehat{\boldsymbol{\theta}}_{(U)})$, and $\mathbf{V}_{(U)}(\widehat{\boldsymbol{\theta}}_{(U)})$, with unambiguous notation.

Predicted Values, PRESS Residual, and PRESS Statistic

An unconditional predicted value is $\widehat{y}_i = \mathbf{x}_i'\widehat{\boldsymbol{\beta}}$, where the vector $\mathbf{x}_i$ is the ith row of $\mathbf{X}$. The (raw) residual is given as $\widehat{\epsilon}_i = y_i - \widehat{y}_i$, and the PRESS *residual* is

$$\widehat{\epsilon}_{i(U)} = y_i - \mathbf{x}_i'\widehat{\boldsymbol{\beta}}_{(U)}$$

The PRESS *statistic* is the sum of the squared PRESS residuals,

$$PRESS = \sum_{i \in U} \widehat{\epsilon}^2_{i(U)}$$

where the sum is over the observations in U.

If EFFECT=, SIZE=, or KEEP= is not specified, PROC MIXED computes the PRESS residual for each observation selected through SELECT= (or all observations if SELECT= is not given). If EFFECT=, SIZE=, or 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 $\mathbf{V}$, or through an oblique projector. The MIXED procedure follows the latter path in the computation of influence diagnostics. The leverage value reported for the ith observation is the ith diagonal entry of the matrix

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{V}(\widehat{\boldsymbol{\theta}})^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}(\widehat{\boldsymbol{\theta}})^{-1}$$

which is the weight of the observation in contributing to its own predicted value, $\mathbf{H} = d\widehat{\mathbf{Y}}/d\mathbf{Y}$.

While $\mathbf{H}$ 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, $\widehat{\mathbf{Y}} = \mathbf{H}\mathbf{Y}$, and in a model with intercept and $\mathbf{V} = \sigma^2\mathbf{I}$, the leverage values

$$h_{ii} = \mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-}\mathbf{x}_i$$

are $h_{ii}^{l} = 1/n \leq h_{ii} \leq 1 = h_{ii}^{u}$ and $\sum_{i=1}^{n} h_{ii} = \text{rank}(\mathbf{X})$. 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 $\mathbf{H}$ equals the rank of $\mathbf{X}$.

If v_{ij} denotes the element in row i, column j of $\mathbf{V}^{-1}$, then for a model containing only an intercept the diagonal elements of $\mathbf{H}$ are

$$h_{ii} = \frac{\sum_{j=1}^{n} v_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{ij}}$$

Because $\sum_{j=1}^{n} v_{ij}$ is a sum of elements in the ith 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 $\mathbf{X} = \mathbf{I}$, $h_{ii} = 1.0$.

Internally and Externally Studentized Residuals

See the section "Residual Diagnostics" on page 6233 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 $\mathbf{V}(\widehat{\boldsymbol{\theta}}) - \mathbf{Q}(\widehat{\boldsymbol{\theta}})$, where $\mathbf{Q}(\widehat{\boldsymbol{\theta}}) = \mathbf{X}(\mathbf{X}'\mathbf{V}(\widehat{\boldsymbol{\theta}})^{-1}\mathbf{X})^{-}\mathbf{X}'$. Externally studentized residuals require iterative influence analysis or a profiled residual variance. In the former case the studentization is based on $\mathbf{V}(\widehat{\boldsymbol{\theta}}_U)$; in the latter case it is based on $\sigma^2_{(U)}\mathbf{V}(\widehat{\boldsymbol{\theta}}^*)$.

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

$$\boldsymbol{\delta}_{(U)} = \widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(U)}$$

Then the MIXED procedure computes

$$D(\boldsymbol{\beta}) = \boldsymbol{\delta}'_{(U)}\widehat{\text{Var}}[\widehat{\boldsymbol{\beta}}]^{-}\boldsymbol{\delta}_{(U)}/\text{rank}(\mathbf{X})$$

where $\widehat{\text{Var}}[\widehat{\boldsymbol{\beta}}]^{-}$ is the matrix that results from sweeping $(\mathbf{X}'\mathbf{V}(\widehat{\boldsymbol{\theta}})^{-1}\mathbf{X})^{-}$.

If $\mathbf{V}$ is known, Cook's D can be calibrated according to a chi-square distribution with degrees of freedom equal to the rank of $\mathbf{X}$ (Christensen, Pearson, and Johnson 1992). For estimated $\mathbf{V}$ the calibration can be carried out according to an $F(\text{rank}(\mathbf{X}), n - \text{rank}(\mathbf{X}))$ 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 $\boldsymbol{\Gamma}$ is the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$, then MIXED computes

$$D_{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(U)})'\widehat{\boldsymbol{\Gamma}}^{-1}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(U)})$$

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:

$$\text{DFFITS}_i = (\widehat{y}_i - \widehat{y}_{i(u)})/\text{ese}(\widehat{y}_i)$$

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 ITER > 0, the estimate is

$$\text{ese}(\widehat{y}_i) = \sqrt{\mathbf{x}_i'(\mathbf{X}'\mathbf{V}(\widehat{\boldsymbol{\theta}}_{(u)})^{-}\mathbf{X})^{-1}\mathbf{x}_i}$$

When ITER=0 and σ^2 is profiled, then

$$\text{ese}(\widehat{y}_i) = \widehat{\sigma}_{(u)}\sqrt{\mathbf{x}_i'(\mathbf{X}'\mathbf{V}(\widehat{\boldsymbol{\theta}}^*)^{-1}\mathbf{X})^{-}\mathbf{x}_i}$$

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 MDFFIT statistic of Belsley, Kuh, and Welsch (1980, p. 32), is closely related to Cook's D. Consider the case $\mathbf{V} = \sigma^2\mathbf{V}(\boldsymbol{\theta}^*)$ so that

$$\text{Var}[\widehat{\boldsymbol{\beta}}] = \sigma^2(\mathbf{X}'\mathbf{V}(\boldsymbol{\theta}^*)^{-1}\mathbf{X})^{-}$$

and let $\widetilde{\text{Var}}[\widehat{\boldsymbol{\beta}}_{(U)}]$ be an estimate of $\text{Var}[\widehat{\boldsymbol{\beta}}_{(U)}]$ that does not use the observations in U. The MDFFITS statistic is then computed as

$$\text{MDFFITS}(\boldsymbol{\beta}) = \boldsymbol{\delta}_{(U)}'\widetilde{\text{Var}}[\widehat{\boldsymbol{\beta}}_{(U)}]^{-}\boldsymbol{\delta}_{(U)}/\text{rank}(\mathbf{X})$$

If ITER=0 and σ^2 is profiled, then $\widetilde{\text{Var}}[\widehat{\boldsymbol{\beta}}_{(U)}]^{-}$ is obtained by sweeping

$$\widehat{\sigma}^2_{(U)}(\mathbf{X}'_{(U)}\mathbf{V}_{(U)}(\widehat{\boldsymbol{\theta}}^*)^{-}\mathbf{X}_{(U)})^{-}$$

The underlying idea is that if $\boldsymbol{\theta}^*$ were known, then

$$(\mathbf{X}'_{(U)}\mathbf{V}_{(U)}(\boldsymbol{\theta}^*)^{-1}\mathbf{X}_{(U)})^{-}$$

would be $\text{Var}[\widehat{\boldsymbol{\beta}}]/\sigma^2$ in a generalized least squares regression with all but the data in U.

In the case of iterative influence analysis, $\widetilde{\text{Var}}[\widehat{\boldsymbol{\beta}}_{(U)}]$ is evaluated at $\widehat{\boldsymbol{\theta}}_{(U)}$. Furthermore, a MDFFITS-type statistic is then computed for the covariance parameters:

$$\text{MDFFITS}(\boldsymbol{\theta}) = (\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(U)})'\widetilde{\text{Var}}[\widehat{\boldsymbol{\theta}}_{(U)}]^{-1}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(U)})$$

Covariance Ratio and Trace

These statistics depend on the availability of an external estimate of $\mathbf{V}$, or at least of σ^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

$$\text{CovTrace}(\boldsymbol{\beta}) = |\text{trace}(\widehat{\text{Var}}[\widehat{\boldsymbol{\beta}}]^{-}\ \widetilde{\text{Var}}[\widehat{\boldsymbol{\beta}}_{(U)}]) - \text{rank}(\mathbf{X})|$$

$$\text{CovRatio}(\boldsymbol{\beta}) = \frac{\det_{ns}(\widetilde{\text{Var}}[\widehat{\boldsymbol{\beta}}_{(U)}])}{\det_{ns}(\widehat{\text{Var}}[\widehat{\boldsymbol{\beta}}])}$$

where $\det_{ns}(\mathbf{M})$ denotes the determinant of the nonsingular part of matrix $\mathbf{M}$.

In the case of iterative influence analysis these statistics are also computed for the covariance parameter estimates. If q denotes the rank of $\text{Var}[\widehat{\boldsymbol{\theta}}]$, then

$$\text{CovTrace}(\boldsymbol{\theta}) = |\text{trace}(\widehat{\text{Var}}[\widehat{\boldsymbol{\theta}}]^{-}\ \widehat{\text{Var}}[\widehat{\boldsymbol{\theta}}_{(U)}]) - q|$$

$$\text{CovRatio}(\boldsymbol{\theta}) = \frac{\det_{ns}(\widehat{\text{Var}}[\widehat{\boldsymbol{\theta}}_{(U)}])}{\det_{ns}(\widehat{\text{Var}}[\widehat{\boldsymbol{\theta}}])}$$

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 6222. Denote as $\boldsymbol{\psi}$ the collection of all parameters, i.e., the fixed effects $\boldsymbol{\beta}$ and the covariance parameters $\boldsymbol{\theta}$. Twice the difference between the (restricted) log-likelihood evaluated at the full-data estimates $\widehat{\boldsymbol{\psi}}$ and at the reduced-data estimates $\widehat{\boldsymbol{\psi}}_{(U)}$ is known as the (restricted) likelihood distance:

$$\text{RLD}_{(U)} = 2\{l_R(\widehat{\boldsymbol{\psi}}) - l_R(\widehat{\boldsymbol{\psi}}_{(U)})\}$$

$$\text{LD}_{(U)} = 2\{l(\widehat{\boldsymbol{\psi}}) - l(\widehat{\boldsymbol{\psi}}_{(U)})\}$$

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 $\boldsymbol{\psi}$ 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 $\mathbf{V} = \sigma^2\mathbf{I}$, $\mathbf{V}$ is known, or $\mathbf{V}^*$ 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 $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}$. When DDFM=KENWARDROGER or DDFM=KENWARDROGER2, 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 $\mathbf{V}$ on $\boldsymbol{\theta}$ is suppressed in the sequel for brevity.

Updating the Fixed Effects Denote by $\mathbf{U}$ the $(n \times k)$ matrix that is assembled from k columns of the identity matrix. Each column of $\mathbf{U}$ corresponds to the removal of one data point. The point being targeted by the ith column of $\mathbf{U}$ corresponds to the row in which a 1 appears. Furthermore, define

$$\begin{aligned}\boldsymbol{\Omega} &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\\ \mathbf{Q} &= \mathbf{X}\boldsymbol{\Omega}\mathbf{X}'\\ \mathbf{P} &= \mathbf{V}^{-1}(\mathbf{V}-\mathbf{Q})\mathbf{V}^{-1}\end{aligned}$$

The change in the fixed-effects estimates following removal of the observations in U is

$$\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(U)} = \boldsymbol{\Omega}\mathbf{X}'\mathbf{V}^{-1}\mathbf{U}(\mathbf{U}'\mathbf{P}\mathbf{U})^{-1}\mathbf{U}'\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}})$$

Using results in Cook and Weisberg (1982, A2) you can further compute

$$\widetilde{\boldsymbol{\Omega}} = (\mathbf{X}'_{(U)}\mathbf{V}^{-1}_{(U)}\mathbf{X}_{(U)})^{-} = \boldsymbol{\Omega} + \boldsymbol{\Omega}\mathbf{X}'\mathbf{V}^{-1}\mathbf{U}(\mathbf{U}'\mathbf{P}\mathbf{U})^{-1}\mathbf{U}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\Omega}$$

If $\mathbf{X}$ is $(n \times p)$ of rank $m < p$, then $\boldsymbol{\Omega}$ is deficient in rank and the MIXED procedure computes needed quantities in $\widetilde{\boldsymbol{\Omega}}$ by sweeping (Goodnight 1979). If the rank of the $(k \times k)$ matrix $\mathbf{U}'\mathbf{P}\mathbf{U}$ is less than k, the removal of the observations introduces a new singularity, whether $\mathbf{X}$ is of full rank or not. The solution vectors $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\beta}}_{(U)}$ 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 $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}$ and of $(\mathbf{X}'_{(U)}\mathbf{V}^{-1}_{(U)}\mathbf{X}_{(U)})^{-}$. If a new singularity is encountered or a former singularity disappears, no influence statistics are computed.

Residual Variance When σ^2 is profiled out of the marginal variance-covariance matrix, a closed-form estimate of σ^2 that is based on only the remaining observations can be computed provided $\mathbf{V}^* = \mathbf{V}(\widehat{\boldsymbol{\theta}}^*)$ is known. Hurtado (1993, Thm. 5.2) shows that

$$(n-q-r)\widehat{\sigma}^2_{(U)} = (n-q)\widehat{\sigma}^2 - \widehat{\boldsymbol{\epsilon}}'_U(\widehat{\sigma}^2\mathbf{U}'\mathbf{P}\mathbf{U})^{-1}\widehat{\boldsymbol{\epsilon}}_U$$

and $\widehat{\boldsymbol{\epsilon}}_U = \mathbf{U}'\mathbf{V}^{*-1}(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}})$. In the case of maximum likelihood estimation $q = 0$ and for REML estimation $q = \mathrm{rank}(\mathbf{X})$. The constant r equals the rank of $(\mathbf{U}'\mathbf{P}\mathbf{U})$ 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 σ^2 is profiled.

The log likelihood function $l(\widehat{\boldsymbol{\theta}})$ evaluated at the full-data and reduced-data estimates can be written as

$$\begin{aligned}l(\widehat{\boldsymbol{\psi}}) &= -\frac{n}{2}\log(\widehat{\sigma}^2) - \frac{1}{2}\log|\mathbf{V}^*| - \frac{1}{2}(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}})'\mathbf{V}^{*-1}(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}})/\widehat{\sigma}^2 - \frac{n}{2}\log(2\pi)\\ l(\widehat{\boldsymbol{\psi}}_{(U)}) &= -\frac{n}{2}\log(\widehat{\sigma}^2_{(U)}) - \frac{1}{2}\log|\mathbf{V}^*| - \frac{1}{2}(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}}_{(U)})'\mathbf{V}^{*-1}(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}}_{(U)})/\widehat{\sigma}^2_{(U)} - \frac{n}{2}\log(2\pi)\end{aligned}$$

Notice that $l(\widehat{\boldsymbol{\theta}}_{(U)})$ 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

$$\mathrm{LD}_{(U)} = n\log\left\{\frac{\widehat{\sigma}^2_{(U)}}{\widehat{\sigma}^2}\right\} - n + \left(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}}_{(U)}\right)'\mathbf{V}^{*-1}\left(\mathbf{y}-\mathbf{X}\widehat{\boldsymbol{\beta}}_{(U)}\right)/\widehat{\sigma}^2_{(U)}$$

Expressions for $\mathrm{RLD}_{(U)}$ 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

$$\frac{\mathbf{g}_k' \mathbf{H}_k^{-1} \mathbf{g}_k}{|f_k|}$$

where f_k is the value of the objective function at iteration k, $\mathbf{g}_k$ is the gradient (first derivative) of f_k, and $\mathbf{H}_k$ is the Hessian (second derivative) of f_k. If $\mathbf{H}_k$ is singular, then PROC MIXED uses the following relative quantity:

$$\frac{\mathbf{g}_k' \mathbf{g}_k}{|f_k|}$$

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 $\mathbf{G}$ and $\mathbf{H}$ (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 $\mathbf{G}$ and $\mathbf{R}$ (see the section "Estimating Covariance Parameters in the Mixed Model" on page 6222). 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 $2\mathbf{H}^{-1}$, where $\mathbf{H}$ is the Hessian matrix. The $\mathbf{H}$ 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 6222. If the log likelihood is an extremely large number, then PROC MIXED has deemed the estimated $\mathbf{V}$ 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 χ^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 $\mathbf{R} = \sigma^2\mathbf{I}$. This statistic has an asymptotic χ^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:

$$F = \frac{\widehat{\boldsymbol{\beta}}'\mathbf{L}'[\mathbf{L}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{L}']^{-}\mathbf{L}\widehat{\boldsymbol{\beta}}}{r}$$

where $r = \mathrm{rank}(\mathbf{L}(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{L}')$. 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 χ^2 tests of the fixed effects. These are carried out by using the numerator of the F statistic and comparing it with the χ^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 78.26.

Table 78.26 ODS Tables Produced by PROC MIXED

Table Name	Description	Required Statement / Option
AccRates	Acceptance rates for posterior sampling	PRIOR
AsyCorr	Asymptotic correlation matrix of covariance parameters	PROC MIXED ASYCORR
AsyCov	Asymptotic covariance matrix of covariance parameters	PROC MIXED ASYCOV
Base	Base densities used for posterior sampling	PRIOR

Table 78.26 *continued*

Table Name	Description	Required Statement / Option
Bound	Computed bound for posterior rejection sampling	PRIOR
CholG	Cholesky root of the estimated **G** matrix	RANDOM / GC
CholR	Cholesky root of blocks of the estimated **R** matrix	REPEATED / RC
CholV	Cholesky root of blocks of the estimated **V** matrix	RANDOM / VC
ClassLevels	Level information from the CLASS statement	Default output
Coef	**L** matrix coefficients	E option in MODEL, CONTRAST, ESTIMATE, or LSMEANS
Contrasts	Results from the CONTRAST statements	CONTRAST
ConvergenceStatus	Convergence status	Default
CorrB	Approximate correlation matrix of fixed-effects parameter estimates	MODEL / CORRB
CovB	Approximate covariance matrix of fixed-effects parameter estimates	MODEL / COVB
CovParms	Estimated covariance parameters	Default output
Diffs	Differences of LS-means	LSMEANS / DIFF (or PDIFF)
Dimensions	Dimensions of the model	Default output
Estimates	Results from ESTIMATE statements	ESTIMATE
FitStatistics	Fit statistics	Default
G	Estimated **G** matrix	RANDOM / G
GCorr	Correlation matrix from the estimated **G** matrix	RANDOM / GCORR
HLM1	Type 1 Hotelling-Lawley-McKeon tests of fixed effects	MODEL / HTYPE=1 and REPEATED / HLM TYPE=UN
HLM2	Type 2 Hotelling-Lawley-McKeon tests of fixed effects	MODEL / HTYPE=2 and REPEATED / HLM TYPE=UN
HLM3	Type 3 Hotelling-Lawley-McKeon tests of fixed effects	REPEATED / HLM TYPE=UN
HLPS1	Type 1 Hotelling-Lawley-Pillai-Samson tests of fixed effects	MODEL / HTYPE=1 and REPEATED / HLPS TYPE=UN
HLPS2	Type 2 Hotelling-Lawley-Pillai-Samson tests of fixed effects	MODEL / HTYPE=1 and REPEATED / HLPS TYPE=UN
HLPS3	Type 3 Hotelling-Lawley-Pillai-Samson tests of fixed effects	REPEATED / HLPS TYPE=UN
Influence	Influence diagnostics	MODEL / INFLUENCE
InfoCrit	Information criteria	PROC MIXED IC
InvCholG	Inverse Cholesky root of the estimated **G** matrix	RANDOM / GCI

Table 78.26 *continued*

Table Name	Description	Required Statement / Option
InvCholR	Inverse Cholesky root of blocks of the estimated **R** matrix	REPEATED / RCI
InvCholV	Inverse Cholesky root of blocks of the estimated **V** matrix	RANDOM / VCI
InvCovB	Inverse of approximate covariance matrix of fixed-effects parameter estimates	MODEL / COVBI
InvG	Inverse of the estimated **G** matrix	RANDOM / GI
InvR	Inverse of blocks of the estimated **R** matrix	REPEATED / RI
InvV	Inverse of blocks of the estimated **V** matrix	RANDOM / VI
IterHistory	Iteration history	Default output
LComponents	Single-degree-of-freedom estimates that correspond to rows of the **L** matrix for fixed effects	MODEL / LCOMPONENTS
LRT	Likelihood ratio test	Default output
LSMeans	LS-means	LSMEANS
MMEq	Mixed model equations	PROC MIXED MMEQ
MMEqSol	Mixed model equations solution	PROC MIXED MMEQSOL
ModelInfo	Model information	Default output
NObs	Number of observations read and used	Default output
ParmSearch	Parameter search values	PARMS
Posterior	Posterior sampling information	PRIOR
Ranks	Ranks of design matrices **X** and (**XZ**)	PROC MIXED RANKS
R	Blocks of the estimated **R** matrix	REPEATED / R
RCorr	Correlation matrix from blocks of the estimated **R** matrix	REPEATED / RCORR
Search	Posterior density search table	PRIOR / PSEARCH
Slices	Tests of LS-means slices	LSMEANS / SLICE=
SolutionF	Fixed-effects solution vector	MODEL / S
SolutionR	Random-effects solution vector	RANDOM / S
Tests1	Type 1 tests of fixed effects	MODEL / HTYPE=1
Tests2	Type 2 tests of fixed effects	MODEL / HTYPE=2
Tests3	Type 3 tests of fixed effects	Default output
Type1	Type 1 analysis of variance	PROC MIXED METHOD=TYPE1
Type2	Type 2 analysis of variance	PROC MIXED METHOD=TYPE2
Type3	Type 3 analysis of variance	PROC MIXED METHOD=TYPE3
Trans	Transformation of covariance parameters	PRIOR / PTRANS
V	Blocks of the estimated **V** matrix	RANDOM / V

Table 78.26 *continued*

Table Name	Description	Required Statement / Option
VCorr	Correlation matrix from blocks of the estimated **V** matrix	RANDOM / VCORR

In Table 78.26, "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, V, CholV, InvCholV, VCorr, and InvV tables.

In Table 78.26, 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 78.27 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 78.27, *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 78.27 and variables in your input data set. You should avoid using input variables with the same names as the variables in these tables.

Table 78.27 Variable Names for the ODS Tables Produced in PROC MIXED

Table Name	Variables
AsyCorr	Row, CovParm, CovP1–CovP*n*
AsyCov	Row, CovParm, CovP1–CovP*n*
Base	Type, Parm1–Parm*n*
Bound	Technique, Converge, Iterations, Evaluations, LogBound, CovP1–CovP*n*, TCovP1–TCovP*n*
CholG	*Model specific*, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Row, Col1–Col*n*
CholR	Index, Row, Col1–Col*n*
CholV	Index, Row, Col1–Col*n*
ClassLevels	Class, Levels, Values

Table 78.27 *continued*

Table Name	Variables
Coef	*Model specific*, LMatrix, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Row1–Row*n*
Contrasts	Label, NumDF, DenDF, ChiSquare, FValue, ProbChiSq, ProbF
CorrB	*Model specific*, Effect, Row, Col1–Col*n*
CovB	*Model specific*, Effect, Row, Col1–Col*n*
CovParms	CovParm, Subject, Group, Estimate, StandardError, ZValue, ProbZ, Alpha, Lower, Upper
Diffs	*Model specific*, Effect, Margins, ByLevel, AT variables, Diff, StandardError, DF, tValue, Tails, Probt, Adjustment, Adjp, Alpha, Lower, Upper, AdjLow, AdjUpp
Dimensions	Descr, Value
Estimates	Label, Estimate, StandardError, DF, tValue, Tails, Probt, Alpha, Lower, Upper
FitStatistics	Descr, Value
G	*Model specific*, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Row, Col1–Col*n*
GCorr	*Model specific*, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Row, Col1–Col*n*
HLM1	Effect, NumDF, DenDF, FValue, ProbF
HLM2	Effect, NumDF, DenDF, FValue, ProbF
HLM3	Effect, NumDF, DenDF, FValue, ProbF
HLPS1	Effect, NumDF, DenDF, FValue, ProbF
HLPS2	Effect, NumDF, DenDF, FValue, ProbF
HLPS3	Effect, NumDF, DenDF, FValue, ProbF
Influence	*Dependent on option modifiers*, Effect, Tuple, Obs1–Obs*k*, Level, Iter, Index, Predicted, Residual, Leverage, PressRes, PRESS, Student, RMSE, RStudent, CookD, DFFITS, MDFFITS, CovRatio, CovTrace, CookDCP, MDFFITSCP, CovRatioCP, CovTraceCP, LD, RLD, Parm1–Parm*p*, CovP1–CovP*q*, Notes
InfoCrit	Neg2LogLike, Parms, AIC, AICC, HQIC, BIC, CAIC
InvCholG	*Model specific*, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Row, Col1–Col*n*
InvCholR	Index, Row, Col1–Col*n*
InvCholV	Index, Row, Col1–Col*n*
InvCovB	*Model specific*, Effect, Row, Col1–Col*n*
InvG	*Model specific*, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Row, Col1–Col*n*
InvR	Index, Row, Col1–Col*n*
InvV	Index, Row, Col1–Col*n*
IterHistory	CovP1–CovP*n*, Iteration, Evaluations, M2ResLogLike, M2LogLike, Criterion
LComponents	Effect, TestType, LIndex, Estimate, StdErr, DF, tValue, Probt
LRT	DF, ChiSquare, ProbChiSq

Table 78.27 *continued*

Table Name	Variables
LSMeans	*Model specific*, Effect, Margins, ByLevel, AT variables, Estimate, StandardError, DF, tValue, Probt, Alpha, Lower, Upper, Cov1–Cov*n*, Corr1–Corr*n*
MMEq	*Model specific*, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Row, Col1–Col*n*
MMEqSol	*Model specific*, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Row, Col1–Col*n*
ModelInfo	Descr, Value
Nobs	Label, N, NObsRead, NObsUsed, SumFreqsRead, SumFreqsUsed
ParmSearch	CovP1–CovP*n*, Var, ResLogLike, M2ResLogLike2, LogLike, M2LogLike, LogDetH
Posterior	Descr, Value
R	Index, Row, Col1–Col*n*
RCorr	Index, Row, Col1–Col*n*
Search	Parm, TCovP1–TCovP*n*, Posterior
Slices	*Model specific*, Effect, Margins, ByLevel, AT variables, NumDF, DenDF, FValue, ProbF
SolutionF	*Model specific*, Effect, Estimate, StandardError, DF, tValue, Probt, Alpha, Lower, Upper
SolutionR	*Model specific*, Effect, Subject, Sub1–Sub*n*, Group, Group1–Group*n*, Estimate, StdErrPred, DF, tValue, Probt, Alpha, Lower, Upper
Tests1	Effect, NumDF, DenDF, ChiSquare, FValue, ProbChiSq, ProbF
Tests2	Effect, NumDF, DenDF, ChiSquare, FValue, ProbChiSq, ProbF
Tests3	Effect, NumDF, DenDF, ChiSquare, FValue, ProbChiSq, ProbF
Type1	Source, DF, SS, MS, EMS, ErrorTerm, ErrorDF, FValue, ProbF
Type2	Source, DF, SS, MS, EMS, ErrorTerm, ErrorDF, FValue, ProbF
Type3	Source, DF, SS, MS, EMS, ErrorTerm, ErrorDF, FValue, ProbF
Trans	Prior, TCovP, CovP1–CovP*n*
V	Index, Row, Col1–Col*n*
VCorr	Index, Row, Col1–Col*n*

Some of the variables listed in Table 78.27 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 607 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 606 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 78.28, along with the required statements and options.

Table 78.28 Graphs Produced by PROC MIXED

ODS Graph Name	Plot Description	Statement or Option
Boxplot	Box plots	PLOTS=BOXPLOT
CovRatioPlot	CovRatio statistics for fixed effects or covariance parameters	PLOTS=INFLUENCESTATPANEL(UNPACK) and MODEL / INFLUENCE
CooksDPlot	Cook's *D* for fixed effects or covariance parameters	PLOTS=INFLUENCESTATPANEL(UNPACK) and MODEL / INFLUENCE
DistancePlot	Likelihood or restricted likelihood distance	MODEL / INFLUENCE
InfluenceEstPlot	Panel of deletion estimates	MODEL / INFLUENCE(EST) or PLOTS=INFLUENCEESTPLOT and MODEL / INFLUENCE
InfluenceEstPlot	Parameter estimates after removing observation or sets of observations	PLOTS=INFLUENCEESTPLOT(UNPACK) and MODEL / INFLUENCE
InfluenceStatPanel	Panel of influence statistics	MODEL / INFLUENCE
PearsonBoxPlot	Box plot of Pearson residuals	PLOTS=PEARSONPANEL(UNPACK BOX)
PearsonByPredicted	Pearson residuals vs. predicted	PLOTS=PEARSONPANEL(UNPACK)
PearsonHistogram	Histogram of Pearson residuals	PLOTS=PEARSONPANEL(UNPACK)
PearsonPanel	Panel of Pearson residuals	MODEL / RESIDUAL
PearsonQQplot	*Q-Q* plot of Pearson residuals	PLOTS=PEARSONPANEL(UNPACK)
PressPlot	Plot of PRESS residuals or PRESS statistic	PLOTS=PRESS and MODEL / INFLUENCE
ResidualBoxplot	Box plot of (raw) residuals	PLOTS=RESIDUALPANEL(UNPACK BOX)
ResidualByPredicted	Residuals vs. predicted	PLOTS=RESIDUALPANEL(UNPACK)
ResidualHistogram	Histogram of raw residuals	PLOTS=RESIDUALPANEL(UNPACK)
ResidualPanel	Panel of (raw) residuals	MODEL / RESIDUAL
ResidualQQplot	*Q-Q* plot of raw residuals	PLOTS=RESIDUALPANEL(UNPACK)
ScaledBoxplot	Box plot of scaled residuals	PLOTS=VCIRYPANEL(UNPACK BOX)
ScaledByPredicted	Scaled residuals vs. predicted	PLOTS=VCIRYPANEL(UNPACK)

Table 78.28 *continued*

ODS Graph Name	Plot Description	Statement or Option
ScaledHistogram	Histogram of scaled residuals	PLOTS=VCIRYPANEL(UNPACK)
ScaledQQplot	*Q*-*Q* plot of scaled residuals	PLOTS=VCIRYPANEL(UNPACK)
StudentBoxplot	Box plot of studentized residuals	PLOTS=STUDENTPANEL(UNPACK BOX)
StudentByPredicted	Studentized residuals vs. predicted	PLOTS=STUDENTPANEL(UNPACK)
StudentHistogram	Histogram of studentized residuals	PLOTS=STUDENTPANEL(UNPACK)
StudentPanel	Panel of studentized residuals	MODEL / RESIDUAL
StudentQQplot	*Q*-*Q* plot of studentized residuals	PLOTS=STUDENTPANEL(UNPACK)
VCIRYPanel	Panel of scaled residuals	MODEL / VCIRY

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 477 and "SLICE Statement" on page 506 in Chapter 19, "Shared Concepts and Topics."

Residual Plots

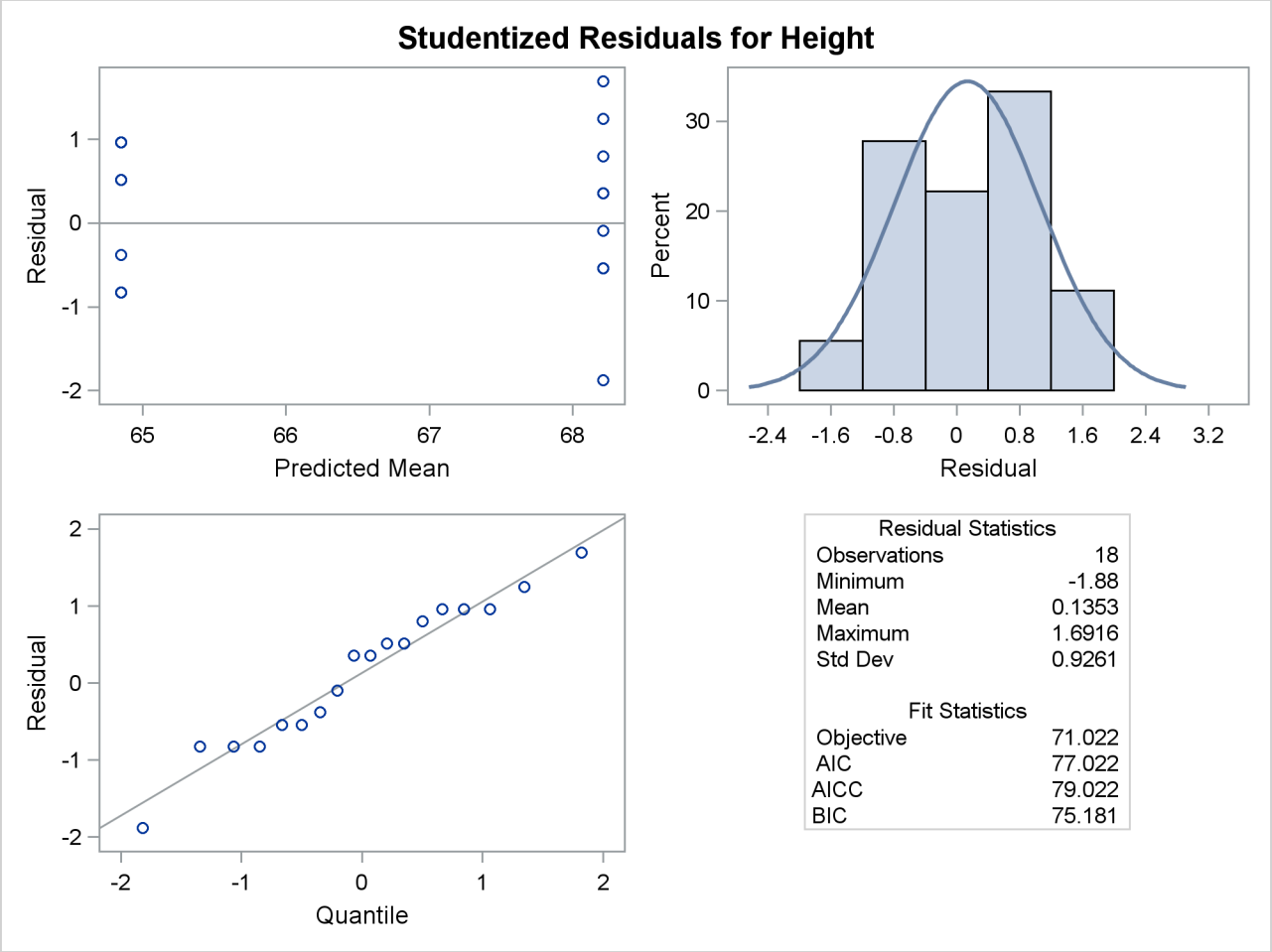
The MIXED procedure can generate panels of residual diagnostics. Each panel consists of a plot of residuals versus predicted values, a histogram with normal density overlaid, a Q-Q plot, and summary residual and fit statistics (Figure 78.15). The plots are produced even if the OUTP= and OUTPM= options 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 6233).

Recall the example in the section "Getting Started: MIXED Procedure" on page 6142. The following statements generate several 2×2 panels of residual graphs:

```
ods graphics on;
proc mixed data=heights plots=studentpanel(marginal conditional);
   class Family Gender;
   model Height = Gender / residual;
   random Family Family*Gender;
run;
ods graphics off;
```

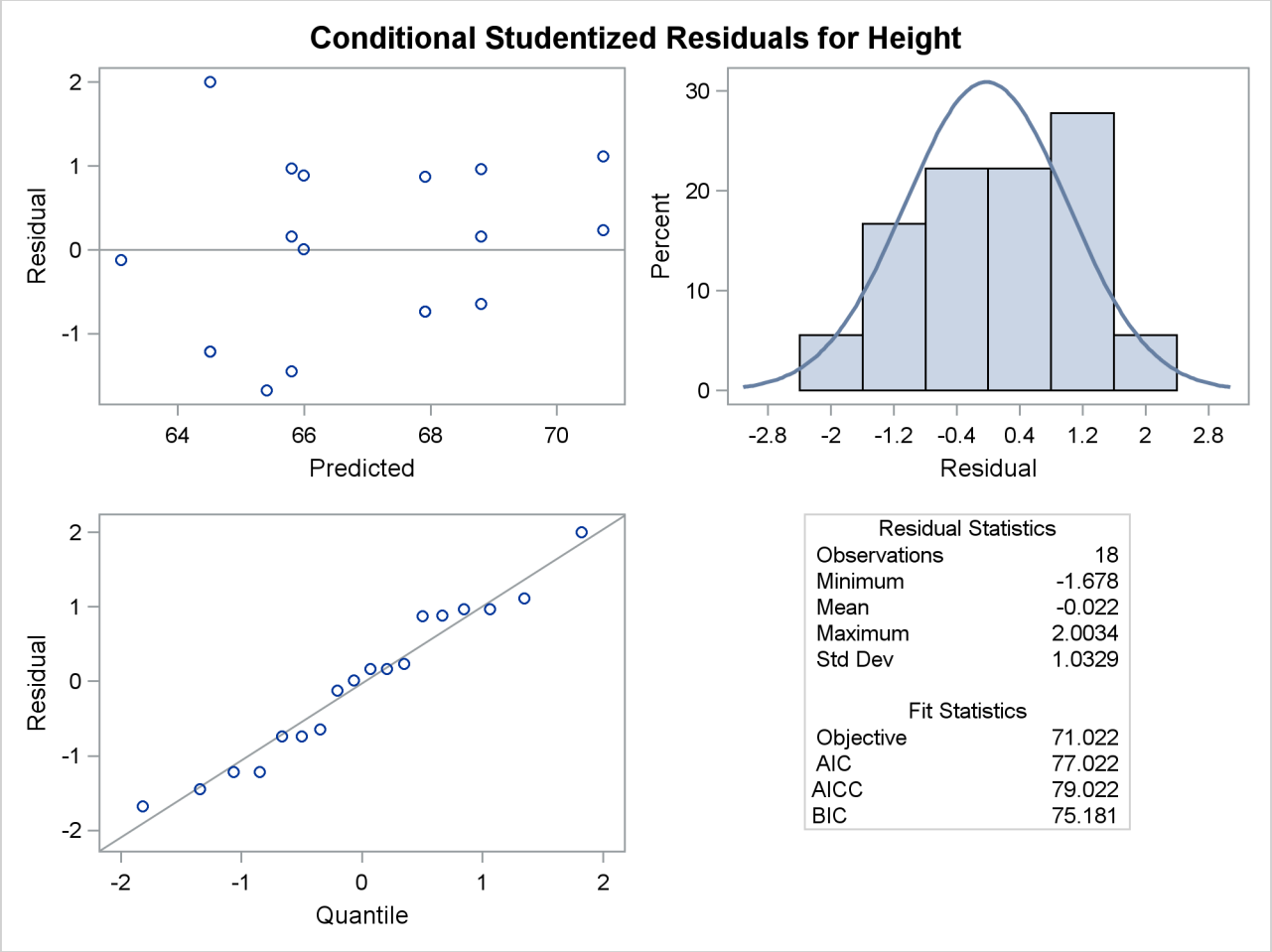
The graphs are created when ODS Graphics is enabled. The panel of the studentized marginal residuals is shown in Figure 78.15, and the panel of the studentized conditional residuals is shown in Figure 78.16.

Figure 78.15 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 78.16 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 78.17 and Figure 78.18. 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 78.8. 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 78.17 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 78.18 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 78.17 Influence Diagnostics

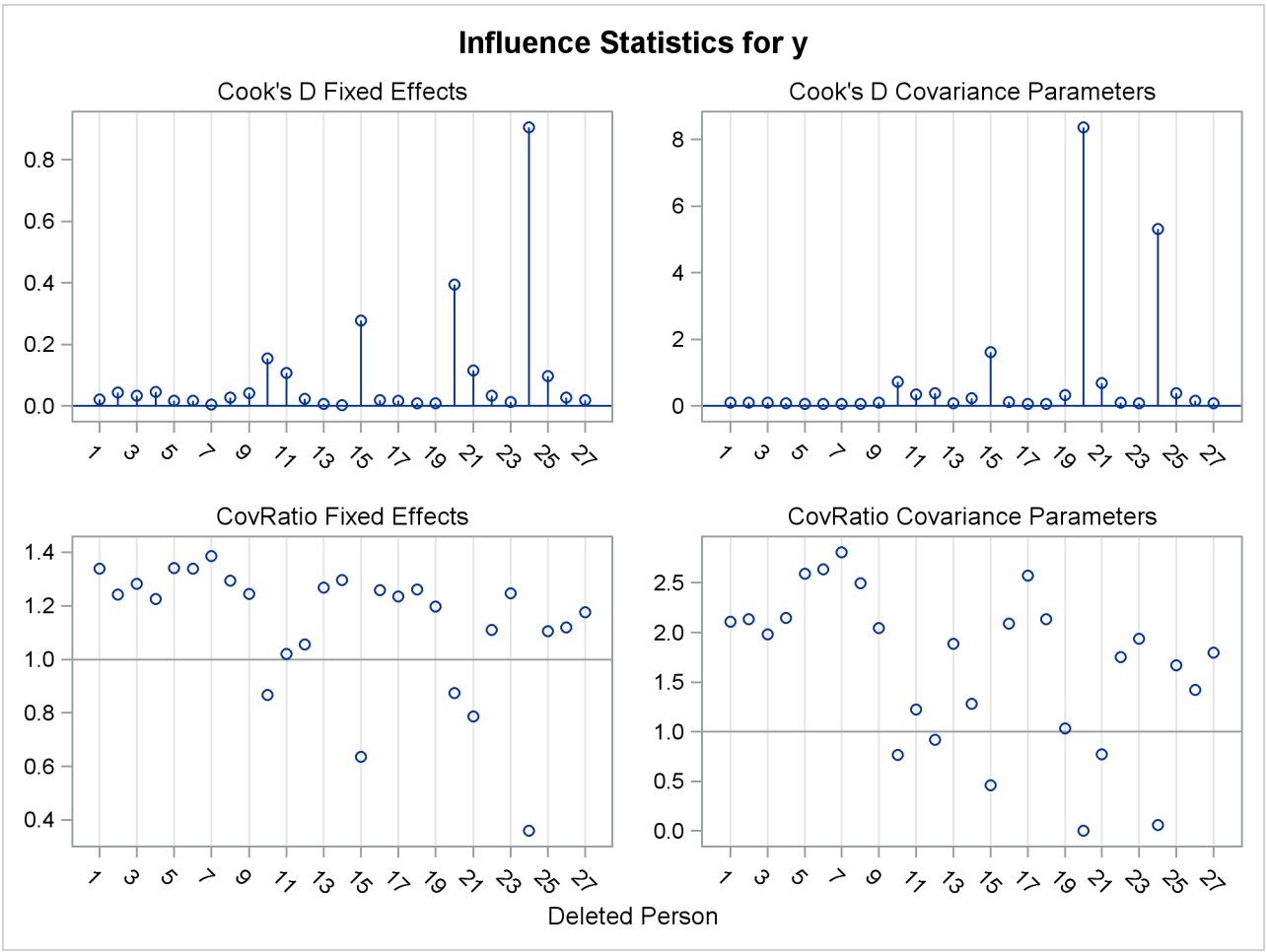
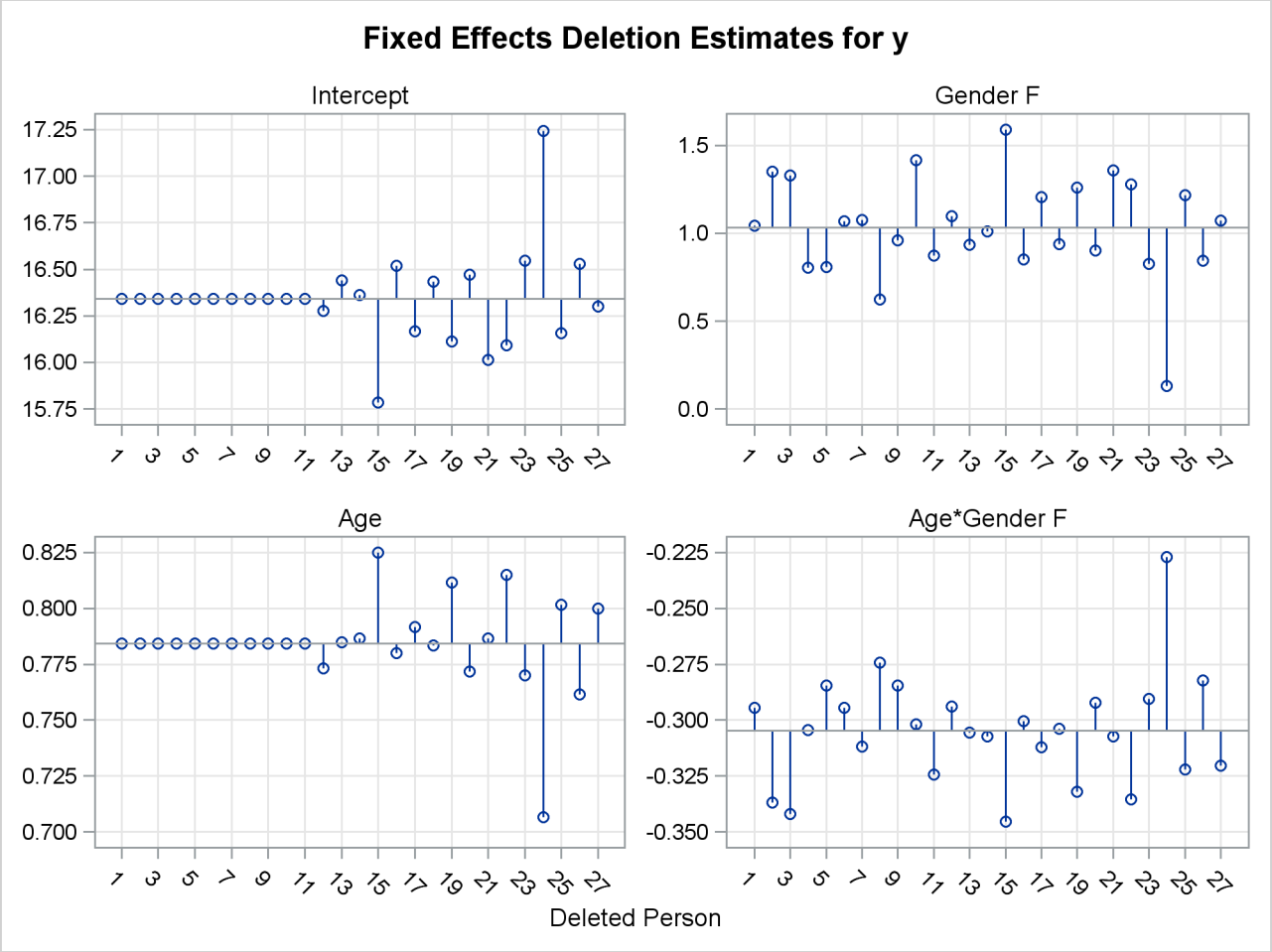


Figure 78.18 Deletion Estimates



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 47, "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.

Table 78.29 Notation for Order Calculations

Symbol	Number
p	Columns of $\mathbf{X}$
g	Columns of $\mathbf{Z}$
N	Observations
q	Covariance parameters
t	Maximum observations per subject
S	Subjects

Using the notation from Table 78.29, 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 $\mathbf{G}$ 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 $\mathbf{R}$ 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 $\mathbf{G}$- and $\mathbf{R}$-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 $\mathbf{G}$ and $\mathbf{R}$. 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 $\mathbf{X}$ changing indicates that there is some linear dependency in the estimate of $\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X}$ that is not found in $\mathbf{X}'\mathbf{X}$. 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 MODEL statement so that the most significant ones are first can help, because PROC MIXED sweeps the estimate of $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ in the order of the 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 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 $\mathbf{X}$, and let g be the number of columns in $\mathbf{Z}$. 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

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 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 $\mathbf{X}$ and $\mathbf{Z}$ 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 $\mathbf{Z}$ matrix with a lot of columns, use the DDFM=BW option in the MODEL statement to eliminate the time required for the containment method.
- If possible, "factor out" a common effect from the effects in the RANDOM statement and make it the SUBJECT= effect. This creates a block-diagonal $\mathbf{G}$ matrix and can often speed calculations.
- If possible, use the same or nested SUBJECT= effects in all RANDOM and REPEATED statements.
- If your data set is very large, you might want to analyze it in pieces. The 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 78.1: 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:

```
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
```

```
2 1 1  30  2 1 2  25
2 2 1  36  2 2 2  28
2 3 1  33  2 3 2  30
3 1 1  32  3 1 2  24
3 2 1  31  3 2 2  27
3 3 1  15  3 3 2  19
4 1 1  30  4 1 2  25
4 2 1  35  4 2 2  30
4 3 1  17  4 3 2  18
;
```

The following statements fit the split-plot model assuming random block effects:

```
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 $\mathbf{X}$ 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 $\mathbf{Z}$ consist of indicator variables corresponding to the levels of the random effects Block and A*Block. The $\mathbf{G}$ matrix is diagonal and contains the variance components of Block and A*Block. The $\mathbf{R}$ matrix is also diagonal and contains the residual variance.

The SAS statements produce Output 78.1.1–Output 78.1.8.

The "Model Information" table in Output 78.1.1 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 78.1.1 Results for Split-Plot Analysis

The Mixed Procedure

Model Information	
Data Set	WORK.SP
Dependent Variable	Y
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

The "Class Level Information" table in Output 78.1.2 lists the levels of all variables specified in the CLASS statement. You can check this table to make sure that the data are correct.

Output 78.1.2 Split-Plot Example (*continued*)

Class Level Information		
Class	**Levels**	**Values**
A	3	1 2 3
B	2	1 2
Block	4	1 2 3 4

The "Dimensions" table in Output 78.1.3 lists the magnitudes of various vectors and matrices. The $\mathbf{X}$ matrix is seen to be 24×12, and the $\mathbf{Z}$ matrix is 24×16.

Output 78.1.3 Split-Plot Example (*continued*)

Dimensions	
Covariance Parameters	3
Columns in X	12
Columns in Z	16
Subjects	1
Max Obs per Subject	24

The "Number of Observations" table in Output 78.1.4 shows that all observations read from the data set are used in the analysis.

Output 78.1.4 Split-Plot Example (*continued*)

Number of Observations	
Number of Observations Read	24
Number of Observations Used	24
Number of Observations Not Used	0

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 78.1.5 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 78.1.5 Split-Plot Analysis (*continued*)

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
0	1	139.81461222	
1	1	119.76184570	0.00000000

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 78.1.6.

Output 78.1.6 Split-Plot Analysis (*continued*)

Covariance Parameter Estimates

Cov Parm	Estimate
Block	62.3958
A*Block	15.3819
Residual	9.3611

The "Fit Statistics" table in Output 78.1.7 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 78.1.7 Split-Plot Analysis (*continued*)

Fit Statistics	
-2 Res Log Likelihood	119.8
AIC (Smaller is Better)	125.8
AICC (Smaller is Better)	127.5
BIC (Smaller is Better)	123.9

Finally, the fixed effects are tested by using Type 3 estimable functions (Output 78.1.8).

Output 78.1.8 Split-Plot Analysis (*continued*)

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
A	2	6	4.07	0.0764
B	1	9	19.39	0.0017
A*B	2	9	4.02	0.0566

The tests match the one obtained from the following PROC GLM statements:

```
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:

```
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;

   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;
run;
```

These statements result in Output 78.1.9.

Output 78.1.9 Inference Space Results

The Mixed Procedure

Estimates

Label	Estimate	Standard 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.

```
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:

```
/* 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 78.2: 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
;

proc mixed data=pr method=ml covtest;
   class Person Gender;
   model y = Gender Age Gender*Age / s;
   repeated / type=un subject=Person r;
run;
```

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 $\mathbf{R}$ matrix is, therefore, block diagonal with 27 blocks, each block consisting of identical 4×4 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 $\mathbf{R}$ be displayed.

The results from this analysis are shown in Output 78.2.1–Output 78.2.9.

Output 78.2.1 Repeated Measures Analysis with Unstructured Covariance Matrix

The Mixed Procedure

Model Information	
Data Set	WORK.PR
Dependent Variable	y
Covariance Structure	Unstructured
Subject Effect	Person
Estimation Method	ML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

In Output 78.2.1, 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 78.2.2 Repeated Measures Analysis (*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

In Output 78.2.2, note that Person has 27 levels and Gender has 2.

Output 78.2.3 Repeated Measures Analysis (*continued*)

Dimensions	
Covariance Parameters	10
Columns in X	6
Columns in Z	0
Subjects	27
Max Obs per Subject	4

In Output 78.2.3, the 10 covariance parameters result from the 4×4 unstructured blocks of $\mathbf{R}$. There is no $\mathbf{Z}$ matrix for this model, and each of the 27 subjects has a maximum of 4 observations.

Output 78.2.4 Repeated Measures Analysis (*continued*)

Number of Observations	
Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

Iteration History			
Iteration	**Evaluations**	**-2 Log Like**	**Criterion**
0	1	478.24175986	
1	2	419.47721707	0.00000152
2	1	419.47704812	0.00000000

Convergence criteria met.

Three Newton-Raphson iterations are required to find the maximum likelihood estimates (Output 78.2.4). 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 78.2.5 Repeated Measures Analysis (*continued*)

Estimated R Matrix for Person 1				
Row	**Col1**	**Col2**	**Col3**	**Col4**
1	5.1192	2.4409	3.6105	2.5222
2	2.4409	3.9279	2.7175	3.0624
3	3.6105	2.7175	5.9798	3.8235
4	2.5222	3.0624	3.8235	4.6180

The 4×4 matrix in Output 78.2.5 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 78.2.6 Repeated Measures Analysis (*continued*)

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	Person	5.1192	1.4169	3.61	0.0002
UN(2,1)	Person	2.4409	0.9835	2.48	0.0131
UN(2,2)	Person	3.9279	1.0824	3.63	0.0001
UN(3,1)	Person	3.6105	1.2767	2.83	0.0047
UN(3,2)	Person	2.7175	1.0740	2.53	0.0114
UN(3,3)	Person	5.9798	1.6279	3.67	0.0001
UN(4,1)	Person	2.5222	1.0649	2.37	0.0179
UN(4,2)	Person	3.0624	1.0135	3.02	0.0025
UN(4,3)	Person	3.8235	1.2508	3.06	0.0022
UN(4,4)	Person	4.6180	1.2573	3.67	0.0001

The "Covariance Parameter Estimates" table in Output 78.2.6 lists the 10 estimated covariance parameters in order; note their correspondence to the first block of **R** displayed in Output 78.2.5. 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:

```
ods output CovParms=cp AsyCov=asy;
proc mixed data=pr method=ml covtest asycov;
   class Person Gender;
   model y = Gender Age Gender*Age / s;
   repeated / type=un subject=Person r;
run;
```

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 78.2.7 Repeated Measures Analysis (*continued*)

Fit Statistics	
-2 Log Likelihood	419.5
AIC (Smaller is Better)	447.5
AICC (Smaller is Better)	452.0
BIC (Smaller is Better)	465.6

Null Model Likelihood Ratio Test		
DF	Chi-Square	Pr > ChiSq
9	58.76	<.0001

The null model likelihood ratio test (LRT) in Output 78.2.7 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 78.2.8 Repeated Measures Analysis (*continued*)

Solution for Fixed Effects						
Effect	Gender	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		15.8423	0.9356	25	16.93	<.0001
Gender	F	1.5831	1.4658	25	1.08	0.2904
Gender	M	0	.	.	.	.
Age		0.8268	0.07911	25	10.45	<.0001
Age*Gender	F	-0.3504	0.1239	25	-2.83	0.0091
Age*Gender	M	0	.	.	.	.

The "Solution for Fixed Effects" table in Output 78.2.8 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 78.2.9 Repeated Measures Analysis (*continued*)

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Gender	1	25	1.17	0.2904
Age	1	25	110.54	<.0001
Age*Gender	1	25	7.99	0.0091

The "Type 3 Tests of Fixed Effects" table in Output 78.2.9 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 78.2.8) 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 χ^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 $\mathbf{R}$, 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:

```
/* 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;
```

To fit a random coefficients model, use the following statements:

```
/* 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, $\mathbf{G}$ is block diagonal with identical 2×2 unstructured blocks for each person. By default, $\mathbf{R}$ becomes $\sigma^2\mathbf{I}$. See Example 78.5 for further information about this model.

Finally, you can fit a compound symmetry structure by using TYPE=CS, as follows:

```
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 78.2.10–Output 78.2.17.

The "Model Information" table in Output 78.2.10 is the same as before except for the change in "Covariance Structure."

Output 78.2.10 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 78.2.11 shows that there are only two covariance parameters in the compound symmetry model; this covariance structure has common variance and common covariance.

Output 78.2.11 Analysis with Compound Symmetry (*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	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 78.2.12).

Output 78.2.12 Analysis with Compound Symmetry (*continued*)

Iteration History			
Iteration	**Evaluations**	**-2 Log Like**	**Criterion**
0	1	478.24175986	
1	1	428.63905802	0.00000000

Convergence criteria met.

Output 78.2.13 displays the estimated $\mathbf{R}$ matrix for the first subject. Note the compound symmetry structure here, which consists of a common covariance with a diagonal enhancement.

Output 78.2.13 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 78.2.14, 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 78.2.14 Analysis with Compound Symmetry (*continued*)

Covariance Parameter Estimates					
Cov Parm	**Subject**	**Estimate**	**Standard Error**	**Z Value**	**Pr Z**
CS	Person	3.0306	0.9552	3.17	0.0015
Residual		1.8746	0.2946	6.36	<.0001

In the case shown in Output 78.2.15, 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 78.2.15 Analysis with Compound Symmetry (*continued*)

Fit Statistics	
-2 Log Likelihood	428.6
AIC (Smaller is Better)	440.6
AICC (Smaller is Better)	441.5
BIC (Smaller is Better)	448.4

Null Model Likelihood Ratio Test		
DF	**Chi-Square**	**Pr > ChiSq**
1	49.60	<.0001

Note that the fixed-effects estimates and their standard errors (Output 78.2.16) are not very different from those in the preceding unstructured example (Output 78.2.8).

Output 78.2.16 Analysis with Compound Symmetry (*continued*)

Solution for Fixed Effects

Effect	Gender	Estimate	Standard 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	0	.	.	.	.
Age		0.7844	0.07654	79	10.25	<.0001
Age*Gender	F	-0.3048	0.1199	79	-2.54	0.0130
Age*Gender	M	0	.	.	.	.

The *F* tests shown in Output 78.2.17 are also similar to those from the preceding unstructured example (Output 78.2.9). Again, the slopes are significantly different but the intercepts are not.

Output 78.2.17 Analysis with Compound Symmetry (*continued*)

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Gender	1	25	0.47	0.5003
Age	1	79	111.10	<.0001
Age*Gender	1	79	6.46	0.0130

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 / ddfm=bw s;
   random int / subject=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 78.2.18–Output 78.2.24. Note that in Output 78.2.18 Gender is listed as a "Group Effect."

Output 78.2.18 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 78.2.19 result from the two compound symmetry structures corresponding to the two levels of Gender.

Output 78.2.19 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 78.2.20 shows, even with the heterogeneity, only one iteration is required for convergence.

Output 78.2.20 Analysis with Heterogeneous Structures (*continued*)

Iteration History			
Iteration	**Evaluations**	**-2 Log Like**	**Criterion**
0	1	478.24175986	
1	1	408.81297228	0.00000000

Convergence criteria met.

The "Covariance Parameter Estimates" table in Output 78.2.21 lists the heterogeneous estimates. Note that both the common covariance and the diagonal enhancement differ between girls and boys.

Output 78.2.21 Analysis with Heterogeneous Structures (*continued*)

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Group**	**Estimate**
Variance	Person	Gender F	0.5900
CS	Person	Gender F	3.8804
Variance	Person	Gender M	2.7577
CS	Person	Gender M	2.4463

As Output 78.2.22 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 χ^2 distribution with two degrees of freedom yields a p-value less than 0.0001, again favoring the heterogeneous model.

Output 78.2.22 Analysis with Heterogeneous Structures (*continued*)

Fit Statistics	
-2 Log Likelihood	408.8
AIC (Smaller is Better)	424.8
AICC (Smaller is Better)	426.3
BIC (Smaller is Better)	435.2

Null Model Likelihood Ratio Test		
DF	**Chi-Square**	**Pr > ChiSq**
3	69.43	<.0001

Note that the fixed-effects estimates shown in Output 78.2.23 are the same as in the homogeneous case, but the standard errors are different.

Output 78.2.23 Analysis with Heterogeneous Structures (*continued*)

Solution for Fixed Effects						
Effect	**Gender**	**Estimate**	**Standard 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 78.2.24 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 78.2.24 Analysis with Heterogeneous Structures (*continued*)

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Gender	1	25	0.55	0.4644
Age	1	79	141.37	<.0001
Age*Gender	1	79	8.22	0.0053

Example 78.3: 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:

```
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 $2\mathbf{H}^{-1}$, where $\mathbf{H}$ 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 78.3.1–Output 78.3.13.

The "Model Information" table in Output 78.3.1 lists details about this variance components model.

Output 78.3.1 Model Information

The Mixed Procedure

Model Information	
Data Set	WORK.HH
Dependent Variable	y
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

The "Class Level Information" table in Output 78.3.2 lists the levels for A and B.

Output 78.3.2 Class Level Information

Class Level Information		
Class	**Levels**	**Values**
a	3	1 2 3
b	2	1 2

The "Dimensions" table in Output 78.3.3 reveals that $\mathbf{X}$ is 16×4 and $\mathbf{Z}$ is 16×8. Since there are no SUBJECT= effects, PROC MIXED considers the data to be effectively from one subject with 16 observations.

Output 78.3.3 Model Dimensions and Number of Observations

Dimensions	
Covariance Parameters	3
Columns in X	4
Columns in Z	8
Subjects	1
Max Obs per Subject	16

Number of Observations	
Number of Observations Read	16
Number of Observations Used	16
Number of Observations Not Used	0

Only a portion of the "Parameter Search" table is shown in Output 78.3.4 because the full listing has 651 rows.

Output 78.3.4 Selected Results of Parameter Search

The Mixed Procedure

CovP1	CovP2	CovP3	Variance	Res Log Like	-2 Res Log Like
17.0000	0.3000	1.0000	80.1400	-52.4699	104.9399
17.0000	0.3050	1.0000	80.0466	-52.4697	104.9393
17.0000	0.3100	1.0000	79.9545	-52.4694	104.9388
17.0000	0.3150	1.0000	79.8637	-52.4692	104.9384
17.0000	0.3200	1.0000	79.7742	-52.4691	104.9381
17.0000	0.3250	1.0000	79.6859	-52.4690	104.9379
17.0000	0.3300	1.0000	79.5988	-52.4689	104.9378
17.0000	0.3350	1.0000	79.5129	-52.4689	104.9377
17.0000	0.3400	1.0000	79.4282	-52.4689	104.9377
17.0000	0.3450	1.0000	79.3447	-52.4689	104.9378
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
20.0000	0.3550	1.0000	78.2003	-52.4683	104.9366
20.0000	0.3600	1.0000	78.1201	-52.4684	104.9368
20.0000	0.3650	1.0000	78.0409	-52.4685	104.9370
20.0000	0.3700	1.0000	77.9628	-52.4687	104.9373
20.0000	0.3750	1.0000	77.8857	-52.4689	104.9377
20.0000	0.3800	1.0000	77.8096	-52.4691	104.9382
20.0000	0.3850	1.0000	77.7345	-52.4693	104.9387
20.0000	0.3900	1.0000	77.6603	-52.4696	104.9392
20.0000	0.3950	1.0000	77.5871	-52.4699	104.9399
20.0000	0.4000	1.0000	77.5148	-52.4703	104.9406

As Output 78.3.5 shows, convergence occurs quickly because PROC MIXED starts from the best value from the grid search.

Output 78.3.5 Iteration History and Convergence Status

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
1	2	104.93416367	0.00000000

Convergence criteria met.

The "Covariance Parameter Estimates" table in Output 78.3.6 lists the variance components estimates. Note that B is much more variable than A*B.

Output 78.3.6 Estimated Covariance Parameters

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr > Z
b	1464.36	2098.01	0.70	0.2426
a*b	26.9581	59.6570	0.45	0.3257
Residual	78.8426	35.3512	2.23	0.0129

The asymptotic covariance matrix in Output 78.3.7 also reflects the large variability of B relative to A*B.

Output 78.3.7 Asymptotic Covariance Matrix of Covariance Parameters

Asymptotic Covariance Matrix of Estimates

Row	Cov Parm	CovP1	CovP2	CovP3
1	b	4401640	1.2831	-273.32
2	a*b	1.2831	3558.96	-502.84
3	Residual	-273.32	-502.84	1249.71

As Output 78.3.8 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 78.3.8 Fit Statistics and Likelihood Ratio Test

Fit Statistics

-2 Res Log Likelihood	104.9
AIC (Smaller is Better)	110.9
AICC (Smaller is Better)	113.6
BIC (Smaller is Better)	107.0

PARMS Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
2	0.00	1.0000

The mixed model equations are analogous to the normal equations in the standard linear model. As Output 78.3.9 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.

Output 78.3.9 Mixed Model Equations

Mixed Model Equations

Row	Effect	a	b	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9	Col10	Col11	Col12
1	Intercept			0.2029	0.06342	0.07610	0.06342	0.1015	0.1015	0.03805	0.02537	0.03805	0.03805	0.02537	0.03805
2	a	1		0.06342	0.06342			0.03805	0.02537	0.03805	0.02537				
3	a	2		0.07610		0.07610		0.03805	0.03805			0.03805	0.03805		
4	a	3		0.06342			0.06342	0.02537	0.03805					0.02537	0.03805
5	b		1	0.1015	0.03805	0.03805	0.02537	0.1022		0.03805		0.03805		0.02537	
6	b		2	0.1015	0.02537	0.03805	0.03805		0.1022		0.02537		0.03805		0.03805
7	a*b	1	1	0.03805	0.03805			0.03805		0.07515					
8	a*b	1	2	0.02537	0.02537				0.02537		0.06246				
9	a*b	2	1	0.03805		0.03805		0.03805				0.07515			
10	a*b	2	2	0.03805		0.03805			0.03805				0.07515		
11	a*b	3	1	0.02537			0.02537	0.02537						0.06246	
12	a*b	3	2	0.03805			0.03805		0.03805						0.07515

Mixed Model Equations	
Row	**Col13**
1	36.4143
2	13.8757
3	12.7469
4	9.7917
5	21.2956
6	15.1187
7	9.3477
8	4.5280
9	7.2676
10	5.4793
11	4.6802
12	5.1115

The solution matrix in Output 78.3.10 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.

Output 78.3.10 Solutions of the Mixed Model Equations

Mixed Model Equations Solution

Row	Effect	a	b	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9	Col10	Col11
1	Intercept			761.84	-29.7718	-29.6578		-731.14	-733.22	-0.4680	0.4680	-0.5257	0.5257	-12.4663
2	a	1		-29.7718	59.5436	29.7718		-2.0764	2.0764	-14.0239	-12.9342	1.0514	-1.0514	12.9342
3	a	2		-29.6578	29.7718	56.2773		-1.0382	1.0382	0.4680	-0.4680	-12.9534	-14.0048	12.4663
4	a	3												
5	b		1	-731.14	-2.0764	-1.0382		741.63	722.73	-4.2598	4.2598	-4.7855	4.7855	-4.2598
6	b		2	-733.22	2.0764	1.0382		722.73	741.63	4.2598	-4.2598	4.7855	-4.7855	4.2598
7	a*b	1	1	-0.4680	-14.0239	0.4680		-4.2598	4.2598	22.8027	4.1555	2.1570	-2.1570	1.9200
8	a*b	1	2	0.4680	-12.9342	-0.4680		4.2598	-4.2598	4.1555	22.8027	-2.1570	2.1570	-1.9200
9	a*b	2	1	-0.5257	1.0514	-12.9534		-4.7855	4.7855	2.1570	-2.1570	22.5560	4.4021	2.1570
10	a*b	2	2	0.5257	-1.0514	-14.0048		4.7855	-4.7855	-2.1570	2.1570	4.4021	22.5560	-2.1570
11	a*b	3	1	-12.4663	12.9342	12.4663		-4.2598	4.2598	1.9200	-1.9200	2.1570	-2.1570	22.8027
12	a*b	3	2	-14.4918	14.0239	14.4918		4.2598	-4.2598	-1.9200	1.9200	-2.1570	2.1570	4.1555

Mixed Model Equations Solution

Row	Col12	Col13
1	-14.4918	159.61
2	14.0239	53.2049
3	14.4918	7.8856
4		
5	4.2598	26.8837
6	-4.2598	-26.8837
7	-1.9200	3.0198
8	1.9200	-3.0198
9	-2.1570	-1.7134
10	2.1570	1.7134
11	4.1555	-0.8115
12	22.8027	0.8115

The A factor is significant at the 5% level (Output 78.3.11).

Output 78.3.11 Tests of Fixed Effects

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
a	2	2	28.00	0.0345

Output 78.3.12 shows that the significance of A appears to be from the difference between its first level and its other two levels.

Output 78.3.12 Least Squares Means for A Effect

Least Squares Means

Effect	a	Estimate	Standard 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 78.3.13 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 78.3.13 Predicted Values

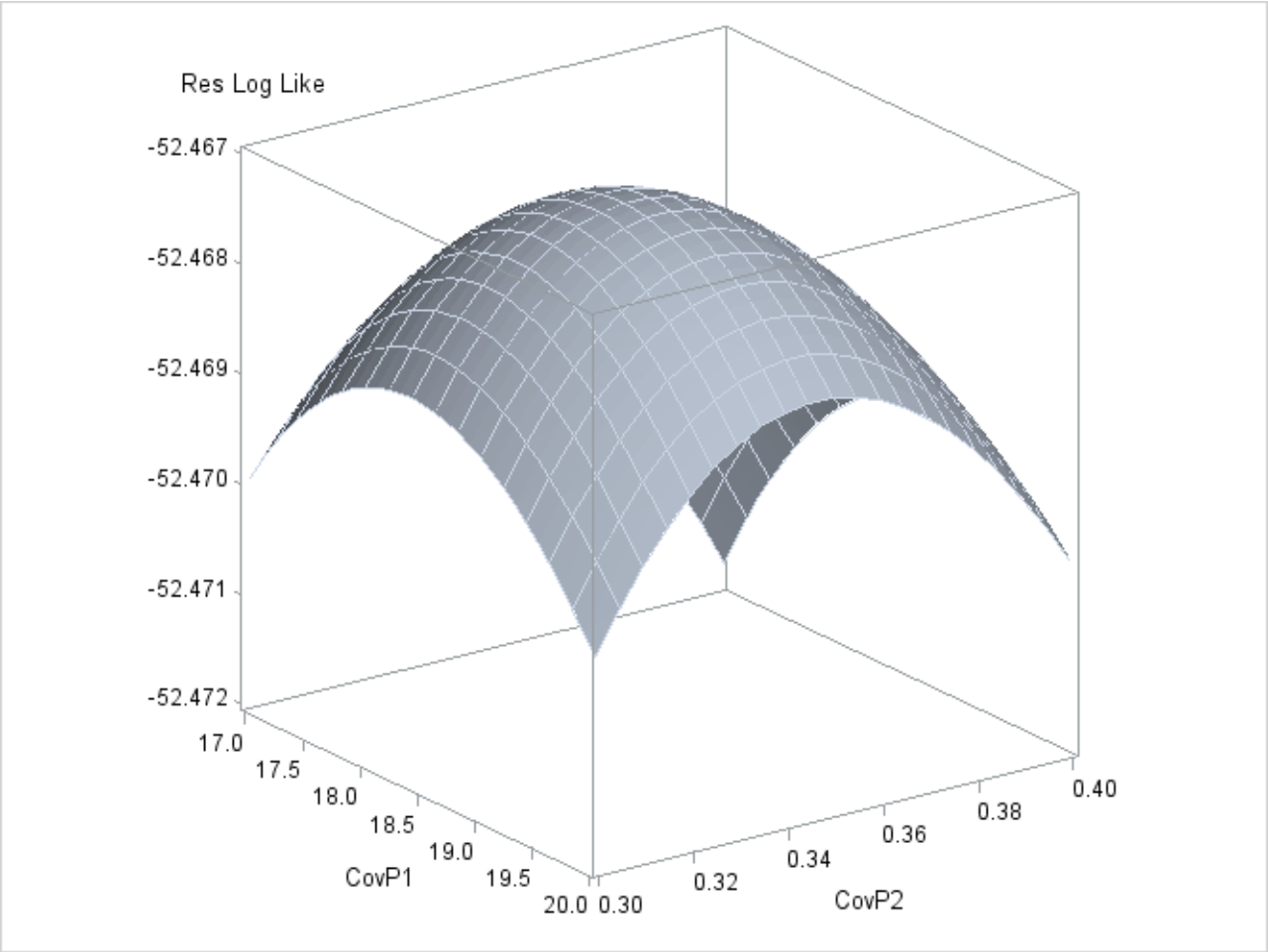
Obs	a	b	y	Pred	StdErrPred	DF	Alpha	Lower	Upper	Resid
1	1	1	237	242.723	4.72563	10	0.05	232.193	253.252	-5.7228
2	1	1	254	242.723	4.72563	10	0.05	232.193	253.252	11.2772
3	1	1	246	242.723	4.72563	10	0.05	232.193	253.252	3.2772
4	1	2	178	182.916	5.52589	10	0.05	170.603	195.228	-4.9159
5	1	2	179	182.916	5.52589	10	0.05	170.603	195.228	-3.9159
6	2	1	208	192.670	4.70076	10	0.05	182.196	203.144	15.3297
7	2	1	178	192.670	4.70076	10	0.05	182.196	203.144	-14.6703
8	2	1	187	192.670	4.70076	10	0.05	182.196	203.144	-5.6703
9	2	2	146	142.330	4.70076	10	0.05	131.856	152.804	3.6703
10	2	2	145	142.330	4.70076	10	0.05	131.856	152.804	2.6703
11	2	2	141	142.330	4.70076	10	0.05	131.856	152.804	-1.3297
12	3	1	186	185.687	5.52589	10	0.05	173.374	197.999	0.3134
13	3	1	183	185.687	5.52589	10	0.05	173.374	197.999	-2.6866
14	3	2	142	133.542	4.72563	10	0.05	123.013	144.072	8.4578
15	3	2	125	133.542	4.72563	10	0.05	123.013	144.072	-8.5422
16	3	2	136	133.542	4.72563	10	0.05	123.013	144.072	2.4578

To plot the likelihood surface by using ODS Graphics, use the following statements:

```
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 78.3.14. The peak of the surface is the REML estimates for the B and A*B variance components.

Output 78.3.14 Plot of Likelihood Surface



Example 78.4: 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:

```
data h;
   input Trait Animal Y;
   datalines;
1 1 6
1 2 8
1 3 7
2 1 9
2 2 5
2 3 .
;
```

Both **G** and **R** are known.

$$\mathbf{G} = \begin{bmatrix} 2 & 1 & 1 & 2 & 1 & 1 \\ 1 & 2 & .5 & 1 & 2 & .5 \\ 1 & .5 & 2 & 1 & .5 & 2 \\ 2 & 1 & 1 & 3 & 1.5 & 1.5 \\ 1 & 2 & .5 & 1.5 & 3 & .75 \\ 1 & .5 & 2 & 1.5 & .75 & 3 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} 4 & 0 & 0 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 4 & 0 & 0 & 1 \\ 1 & 0 & 0 & 5 & 0 & 0 \\ 0 & 1 & 0 & 0 & 5 & 0 \\ 0 & 0 & 1 & 0 & 0 & 5 \end{bmatrix}$$

In order to read **G** into PROC MIXED by using the GDATA= option in the RANDOM statement, perform the following DATA step:

```
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:

```
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 $\mathbf{G}$ and $\mathbf{G}^{-1}$, respectively. The S option requests that the random-effects solution vector be displayed.

Note that the preceding $\mathbf{R}$ matrix is block diagonal if the data are sorted by animals. The REPEATED statement exploits this fact by requesting $\mathbf{R}$ to have unstructured 2×2 blocks corresponding to animals, which are the subjects. The R and RI options request that the estimated 2×2 blocks for the first animal and its inverse be displayed. The PARMS statement lists the parameters of this 2×2 matrix. Note that the parameters from $\mathbf{G}$ 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 78.4.1–Output 78.4.12.

The "Unstructured" covariance structure (Output 78.4.1) applies to both $\mathbf{G}$ and $\mathbf{R}$ here. The levels of Trait and Animal have been specified correctly.

Output 78.4.1 Model and Class Level Information

The Mixed Procedure

Model Information	
Data Set	WORK.H
Dependent Variable	Y
Covariance Structure	Unstructured
Subject Effect	Animal
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	**Levels**	**Values**
Trait	2	1 2
Animal	3	1 2 3

The three covariance parameters indicated in Output 78.4.2 correspond to those from the $\mathbf{R}$ matrix. Those from $\mathbf{G}$ are considered fixed and known because of the GDATA= option.

Output 78.4.2 Model Dimensions and Number of Observations

Dimensions	
Covariance Parameters	3
Columns in X	2
Columns in Z	6
Subjects	1
Max Obs per Subject	5

Output 78.4.2 *continued*

Number of Observations	
Number of Observations Read	6
Number of Observations Used	5
Number of Observations Not Used	1

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 78.4.3).

Output 78.4.3 REML Log Likelihood

Parameter Search				
CovP1	CovP2	CovP3	Res Log Like	-2 Res Log Like
4.0000	1.0000	5.0000	-7.3731	14.7463

The block of **R** corresponding to the first animal and the inverse of this block are shown in Output 78.4.4.

Output 78.4.4 Inverse R Matrix

Estimated R Matrix for Animal 1		
Row	Col1	Col2
1	4.0000	1.0000
2	1.0000	5.0000

Estimated Inv(R) Matrix for Animal 1		
Row	Col1	Col2
1	0.2632	-0.05263
2	-0.05263	0.2105

The **G** matrix as specified in the GDATA= data set and its inverse are shown in Output 78.4.5 and Output 78.4.6.

Output 78.4.5 G Matrix

Estimated G Matrix									
Row	Effect	Trait	Animal	Col1	Col2	Col3	Col4	Col5	Col6
1	Trait*Animal	1	1	2.0000	1.0000	1.0000	2.0000	1.0000	1.0000
2	Trait*Animal	1	2	1.0000	2.0000	0.5000	1.0000	2.0000	0.5000
3	Trait*Animal	1	3	1.0000	0.5000	2.0000	1.0000	0.5000	2.0000
4	Trait*Animal	2	1	2.0000	1.0000	1.0000	3.0000	1.5000	1.5000
5	Trait*Animal	2	2	1.0000	2.0000	0.5000	1.5000	3.0000	0.7500
6	Trait*Animal	2	3	1.0000	0.5000	2.0000	1.5000	0.7500	3.0000

Output 78.4.6 Inverse G Matrix

Estimated Inv(G) Matrix									
Row	Effect	Trait	Animal	Col1	Col2	Col3	Col4	Col5	Col6
1	Trait*Animal	1	1	2.5000	-1.0000	-1.0000	-1.6667	0.6667	0.6667
2	Trait*Animal	1	2	-1.0000	2.0000		0.6667	-1.3333	
3	Trait*Animal	1	3	-1.0000		2.0000	0.6667		-1.3333
4	Trait*Animal	2	1	-1.6667	0.6667	0.6667	1.6667	-0.6667	-0.6667
5	Trait*Animal	2	2	0.6667	-1.3333		-0.6667	1.3333	
6	Trait*Animal	2	3	0.6667		-1.3333	-0.6667		1.3333

The table of covariance parameter estimates in Output 78.4.7 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 78.4.7 R-Side Covariance Parameters

Covariance Parameter Estimates		
Cov Parm	Subject	Estimate
UN(1,1)	Animal	4.0000
UN(2,1)	Animal	1.0000
UN(2,2)	Animal	5.0000

The coefficients of the mixed model equations in Output 78.4.8 agree with Henderson (1984, p. 55). Recall from Output 78.4.1 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 78.4.8 Mixed Model Equations with Y Border

Mixed Model Equations												
Row	Effect	Trait	Animal	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9
1	Trait	1		0.7763	-0.1053	0.2632	0.2632	0.2500	-0.05263	-0.05263		4.6974
2	Trait	2		-0.1053	0.4211	-0.05263	-0.05263		0.2105	0.2105		2.2105
3	Trait*Animal	1	1	0.2632	-0.05263	2.7632	-1.0000	-1.0000	-1.7193	0.6667	0.6667	1.1053
4	Trait*Animal	1	2	0.2632	-0.05263	-1.0000	2.2632		0.6667	-1.3860		1.8421
5	Trait*Animal	1	3	0.2500		-1.0000		2.2500	0.6667		-1.3333	1.7500
6	Trait*Animal	2	1	-0.05263	0.2105	-1.7193	0.6667	0.6667	1.8772	-0.6667	-0.6667	1.5789
7	Trait*Animal	2	2	-0.05263	0.2105	0.6667	-1.3860		-0.6667	1.5439		0.6316
8	Trait*Animal	2	3			0.6667		-1.3333	-0.6667		1.3333	

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 78.4.9).

Output 78.4.9 Solutions of the Mixed Model Equations with Y Border

Mixed Model Equations Solution

Row	Effect	Trait	Animal	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9
1	Trait	1		2.5508	1.5685	-1.3047	-1.1775	-1.1701	-1.3002	-1.1821	-1.1678	6.9909
2	Trait	2		1.5685	4.5539	-1.4112	-1.3534	-0.9410	-2.1592	-2.1055	-1.3149	6.9959
3	Trait*Animal	1	1	-1.3047	-1.4112	1.8282	1.0652	1.0206	1.8010	1.0925	1.0070	0.05450
4	Trait*Animal	1	2	-1.1775	-1.3534	1.0652	1.7589	0.7085	1.0900	1.7341	0.7209	-0.04955
5	Trait*Animal	1	3	-1.1701	-0.9410	1.0206	0.7085	1.7812	1.0095	0.7197	1.7756	0.02230
6	Trait*Animal	2	1	-1.3002	-2.1592	1.8010	1.0900	1.0095	2.7518	1.6392	1.4849	0.2651
7	Trait*Animal	2	2	-1.1821	-2.1055	1.0925	1.7341	0.7197	1.6392	2.6874	0.9930	-0.2601
8	Trait*Animal	2	3	-1.1678	-1.3149	1.0070	0.7209	1.7756	1.4849	0.9930	2.7645	0.1276

The solutions for the fixed and random effects in Output 78.4.10 correspond to the last column in Output 78.4.9. 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 78.4.9).

Output 78.4.10 Solutions for Fixed and Random Effects

Solution for Fixed Effects

Effect	Trait	Estimate	Standard Error	DF	t Value	Pr > |t|
Trait	1	6.9909	1.5971	3	4.38	0.0221
Trait	2	6.9959	2.1340	3	3.28	0.0465

Solution for Random Effects

Effect	Trait	Animal	Estimate	Std Err Pred	DF	t Value	Pr > |t|
Trait*Animal	1	1	0.05450	1.3521	0	0.04	.
Trait*Animal	1	2	-0.04955	1.3262	0	-0.04	.
Trait*Animal	1	3	0.02230	1.3346	0	0.02	.
Trait*Animal	2	1	0.2651	1.6589	0	0.16	.
Trait*Animal	2	2	-0.2601	1.6393	0	-0.16	.
Trait*Animal	2	3	0.1276	1.6627	0	0.08	.

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 78.4.11 Significance Test Comparing Traits

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Trait	2	3	10.59	0.0437

The two estimated traits are significantly different from zero at the 5% level (Output 78.4.11).

Output 78.4.12 displays the predicted values of the observations based on the trait and breeding value estimates—that is, the fixed and random effects.

Output 78.4.12 Predicted Observations

Obs	Trait	Animal	Y	Pred	StdErrPred	DF	Alpha	Lower	Upper	Resid
1	1	1	6	7.04542	1.33027	0	0.05	.	.	-1.04542
2	1	2	8	6.94137	1.39806	0	0.05	.	.	1.05863
3	1	3	7	7.01321	1.41129	0	0.05	.	.	-0.01321
4	2	1	9	7.26094	1.72839	0	0.05	.	.	1.73906
5	2	2	5	6.73576	1.74077	0	0.05	.	.	-1.73576
6	2	3	.	7.12015	2.99088	0	0.05	.	.	.

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 78.5: 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:

```
data rc;
   input Batch Month @@;
   Monthc = Month;
   do i = 1 to 6;
      input Y @@;
      output;
   end;
   datalines;
 1   0  101.2 103.3 103.3 102.1 104.4 102.4
 1   1   98.8  99.4  99.7  99.5     .     .
 1   3   98.4  99.0  97.3  99.8     .     .
 1   6  101.5 100.2 101.7 102.7     .     .
 1   9   96.3  97.2  97.2  96.3     .     .
 1  12   97.3  97.9  96.8  97.7  97.7  96.7
 2   0  102.6 102.7 102.4 102.1 102.9 102.6
 2   1   99.1  99.0  99.9 100.6     .     .
 2   3  105.7 103.3 103.4 104.0     .     .
 2   6  101.3 101.5 100.9 101.4     .     .
 2   9   94.1  96.5  97.2 95.6      .     .
 2  12   93.1  92.8  95.4 92.2    92.2  93.0
 3   0  105.1 103.9 106.1 104.1 103.7 104.6
```

```
 3   1  102.2 102.0 100.8  99.8    .      .
 3   3  101.2 101.8 100.8 102.6    .      .
 3   6  101.1 102.0 100.1 100.2    .      .
 3   9  100.9  99.5 102.2 100.8    .      .
 3  12   97.8  98.3  96.9  98.4  96.9  96.5
;

proc mixed data=rc;
   class Batch;
   model Y = Month / s;
   random Int Month / type=un sub=Batch s;
run;
```

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 on page 6292.

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, $\mathbf{G}$ 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 78.5.1–Output 78.5.9. The "Unstructured" covariance structure in Output 78.5.1 applies to $\mathbf{G}$ here.

Output 78.5.1 Model Information in Random Coefficients Analysis

The Mixed Procedure

Model Information	
Data Set	WORK.RC
Dependent Variable	Y
Covariance Structure	Unstructured
Subject Effect	Batch
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Batch is the only classification variable in this analysis, and it has three levels (Output 78.5.2).

Output 78.5.2 Random Coefficients Analysis (*continued*)

Class Level Information		
Class	**Levels**	**Values**
Batch	3	1 2 3

The "Dimensions" table in Output 78.5.3 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 78.5.3 Random Coefficients Analysis (*continued*)

Dimensions	
Covariance Parameters	4
Columns in X	2
Columns in Z per Subject	2
Subjects	3
Max Obs per Subject	28

Number of Observations	
Number of Observations Read	108
Number of Observations Used	84
Number of Observations Not Used	24

As Output 78.5.4 shows, only one iteration is required for convergence.

Output 78.5.4 Random Coefficients Analysis (*continued*)

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
0	1	367.02768461	
1	1	350.32813577	0.00000000

Convergence criteria met.

The Estimate column in Output 78.5.5 lists the estimated elements of the unstructured 2×2 matrix comprising the blocks of **G**. Note that the random coefficients are negatively correlated.

Output 78.5.5 Random Coefficients Analysis (*continued*)

Covariance Parameter Estimates		
Cov Parm	**Subject**	**Estimate**
UN(1,1)	Batch	0.9768
UN(2,1)	Batch	-0.1045
UN(2,2)	Batch	0.03717
Residual		3.2932

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 78.5.6).

Output 78.5.6 Random Coefficients Analysis (*continued*)

Fit Statistics	
-2 Res Log Likelihood	350.3
AIC (Smaller is Better)	358.3
AICC (Smaller is Better)	358.8
BIC (Smaller is Better)	354.7

Null Model Likelihood Ratio Test		
DF	Chi-Square	Pr > ChiSq
3	16.70	0.0008

The fixed-effects estimates represent the estimated means for the random intercept and slope, respectively (Output 78.5.7).

Output 78.5.7 Random Coefficients Analysis (*continued*)

Solution for Fixed Effects					
Effect	Estimate	Standard 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 78.5.8). 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 78.5.8 Random Coefficients Analysis (*continued*)

Solution for Random Effects						
Effect	Batch	Estimate	Std Err Pred	DF	t Value	Pr > |t|
Intercept	1	-1.0010	0.6842	78	-1.46	0.1474
Month	1	0.1287	0.1245	78	1.03	0.3047
Intercept	2	0.3934	0.6842	78	0.58	0.5669
Month	2	-0.2060	0.1245	78	-1.65	0.1021
Intercept	3	0.6076	0.6842	78	0.89	0.3772
Month	3	0.07731	0.1245	78	0.62	0.5365

The F statistic in the "Type 3 Tests of Fixed Effects" table in Output 78.5.9 is the square of the t statistic used in the test of Month in the preceding "Solution for Fixed Effects" table (compare Output 78.5.7 and Output 78.5.9). 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 78.5.9 Random Coefficients Analysis (*continued*)

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Month	1	2	19.41	0.0478

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 78.5.10.

Output 78.5.10 Random Coefficients with Nested Errors Analysis

The Mixed Procedure

Model Information	
Data Set	WORK.RC
Dependent Variable	Y
Covariance Structure	Variance Components
Subject Effect	Batch
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	Levels	Values
Batch	3	1 2 3
Monthc	6	0 1 3 6 9 12

Dimensions	
Covariance Parameters	4
Columns in X	2
Columns in Z per Subject	8
Subjects	3
Max Obs per Subject	28

Number of Observations	
Number of Observations Read	108
Number of Observations Used	84
Number of Observations Not Used	24

Output 78.5.10 *continued*

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	367.02768461	
1	4	277.51945360	.
2	1	276.97551718	0.00104208
3	1	276.90304909	0.00003174
4	1	276.90100316	0.00000004
5	1	276.90100092	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
Intercept	Batch	0
Month	Batch	0.01243
Monthc	Batch	3.7411
Residual		0.7969

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 78.5.11 and Output 78.5.6). 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 78.5.11 Random Coefficients with Nested Errors Analysis (*continued*)

Fit Statistics

-2 Res Log Likelihood	276.9
AIC (Smaller is Better)	282.9
AICC (Smaller is Better)	283.2
BIC (Smaller is Better)	280.2

The better-fitting covariance model affects the standard errors of the fixed-effects parameter estimates more than the estimates themselves (Output 78.5.12).

Output 78.5.12 Random Coefficients with Nested Errors Analysis (*continued*)

Solution for Fixed Effects

Effect	Estimate	Standard 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 78.5.13). You can use these values to compare batches and months.

Output 78.5.13 Random Coefficients with Nested Errors Analysis (*continued*)

Solution for Random Effects

Effect	Batch	Monthc	Estimate	Std Err Pred	DF	t Value	Pr > |t|
Intercept	1		0	.	.	.	.
Month	1		-0.00028	0.09268	66	-0.00	0.9976
Monthc	1	0	0.2191	0.7896	66	0.28	0.7823
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		-0.07571	0.09268	66	-0.82	0.4169
Monthc	2	0	-0.00621	0.7896	66	-0.01	0.9938
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		0.07600	0.09268	66	0.82	0.4152
Monthc	3	0	1.9574	0.7896	66	2.48	0.0157
Monthc	3	1	-0.8850	0.7571	66	-1.17	0.2466
Monthc	3	3	0.3006	0.6865	66	0.44	0.6629
Monthc	3	6	0.7972	0.7328	66	1.09	0.2806
Monthc	3	9	2.0059	0.9300	66	2.16	0.0347
Monthc	3	12	0.002293	1.1992	66	0.00	0.9985

Output 78.5.14 Random Coefficients with Nested Errors Analysis (*continued*)

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Month	1	2	15.78	0.0579

The test of Month is similar to that from the previous model, although it is no longer significant at the 5% level (Output 78.5.14).

Example 78.6: 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 $\mathbf{G}$ and $\mathbf{R}$ can be modeled. One of Stroup's models specifies a diagonal $\mathbf{G}$ containing the variance components for Block, Block*Dir, and Block*Irrig, and a Toeplitz $\mathbf{R}$ 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
 3 Bridger  2.7 2.8 4.0 5.0 5.2 5.2 5.9 6.1 6.0 4.3 3.1 3.1
;

proc mixed;
   class Block Cult Dir Irrig;
   model Y = Cult|Dir|Irrig@2;
   random Block Block*Dir Block*Irrig;
   repeated / type=toep(4) sub=Block*Cult r;
   lsmeans Cult|Irrig;
   estimate 'Bridger vs Luke' Cult 1 -1 0;
   estimate 'Linear Irrig' Irrig -5 -3 -1 1 3 5;
```

```
    estimate 'B vs L x Linear Irrig' Cult*Irrig
             -5 -3 -1 1 3 5 5 3 1 -1 -3 -5;
run;
```

The preceding statements use the bar operator (|) and the at sign (@) to specify all two-factor interactions between Cult, Dir, and Irrig as fixed effects.

The RANDOM statement sets up the **Z** and **G** matrices corresponding to the random effects Block, Block*Dir, and Block*Irrig.

In the REPEATED statement, the TYPE=TOEP(4) option sets up the blocks of the **R** matrix to be Toeplitz with four bands below and including the main diagonal. The subject effect is Block*Cult, and it produces nine 12×12 blocks. The R option requests that the first block of **R** be displayed.

Least squares means (LSMEANS) are requested for Cult, Irrig, and Cult*Irrig, and a few ESTIMATE statements are specified to illustrate some linear combinations of the fixed effects.

The results from this analysis are shown in Output 78.6.1.

The "Covariance Structures" row in Output 78.6.1 reveals the two different structures assumed for **G** and **R**.

Output 78.6.1 Model Information in Line-Source Sprinkler Analysis

The Mixed Procedure

Model Information	
Data Set	WORK.LINE
Dependent Variable	Y
Covariance Structures	Variance Components, Toeplitz
Subject Effect	Block*Cult
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

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 78.6.2).

Output 78.6.2 Class Level Information

Class Level Information		
Class	**Levels**	**Values**
Block	3	1 2 3
Cult	3	Bridger Luke Nugaines
Dir	2	North South
Irrig	6	1 2 3 4 5 6

Even though there is a SUBJECT= effect in the REPEATED statement, the analysis considers all of the data to be from one subject because there is no corresponding SUBJECT= effect in the RANDOM statement (Output 78.6.3).

Output 78.6.3 Model Dimensions and Number of Observations

Dimensions	
Covariance Parameters	7
Columns in X	48
Columns in Z	27
Subjects	1
Max Obs per Subject	108

Number of Observations	
Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

The Newton-Raphson algorithm converges successfully in seven iterations (Output 78.6.4).

Output 78.6.4 Iteration History and Convergence Status

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	226.25427252	
1	4	187.99336173	.
2	3	186.62579299	0.10431081
3	1	184.38218213	0.04807260
4	1	183.41836853	0.00886548
5	1	183.25111475	0.00075353
6	1	183.23809997	0.00000748
7	1	183.23797748	0.00000000

Convergence criteria met.

The first block of the estimated **R** matrix has the TOEP(4) structure, and the observations that are three plots apart exhibit a negative correlation (Output 78.6.5).

Output 78.6.5 Estimated R Matrix for the First Subject

Estimated R Matrix for Block*Cult 1 Bridger												
Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9	Col10	Col11	Col12
1	0.2850	0.007986	0.001452	-0.09253								
2	0.007986	0.2850	0.007986	0.001452	-0.09253							
3	0.001452	0.007986	0.2850	0.007986	0.001452	-0.09253						
4	-0.09253	0.001452	0.007986	0.2850	0.007986	0.001452	-0.09253					
5		-0.09253	0.001452	0.007986	0.2850	0.007986	0.001452	-0.09253				
6			-0.09253	0.001452	0.007986	0.2850	0.007986	0.001452	-0.09253			
7				-0.09253	0.001452	0.007986	0.2850	0.007986	0.001452	-0.09253		
8					-0.09253	0.001452	0.007986	0.2850	0.007986	0.001452	-0.09253	
9						-0.09253	0.001452	0.007986	0.2850	0.007986	0.001452	-0.09253
10							-0.09253	0.001452	0.007986	0.2850	0.007986	0.001452
11								-0.09253	0.001452	0.007986	0.2850	0.007986
12									-0.09253	0.001452	0.007986	0.2850

Output 78.6.6 lists the estimated covariance parameters from both **G** and **R**. The first three are the variance components making up the diagonal **G**, and the final four make up the Toeplitz structure in the blocks of **R**. The Residual row corresponds to the variance of the Toeplitz structure, and it represents the parameter profiled out during the optimization process.

Output 78.6.6 Estimated Covariance Parameters

Covariance Parameter Estimates		
Cov Parm	**Subject**	**Estimate**
Block		0.2194
Block*Dir		0.01768
Block*Irrig		0.03539
TOEP(2)	Block*Cult	0.007986
TOEP(3)	Block*Cult	0.001452
TOEP(4)	Block*Cult	-0.09253
Residual		0.2850

The "–2 Res Log Likelihood" value in Output 78.6.7 is the same as the final value listed in the "Iteration History" table (Output 78.6.4).

Output 78.6.7 Fit Statistics Based on the Residual Log Likelihood

Fit Statistics	
-2 Res Log Likelihood	183.2
AIC (Smaller is Better)	197.2
AICC (Smaller is Better)	198.8
BIC (Smaller is Better)	190.9

Every fixed effect except for Dir and Cult*Irrig is significant at the 5% level (Output 78.6.8).

Output 78.6.8 Tests for Fixed Effects

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Cult	2	68	7.98	0.0008
Dir	1	2	3.95	0.1852
Cult*Dir	2	68	3.44	0.0379
Irrig	5	10	102.60	<.0001
Cult*Irrig	10	68	1.91	0.0580
Dir*Irrig	5	68	6.12	<.0001

The "Estimates" table lists the results from the various linear combinations of fixed effects specified in the ESTIMATE statements (Output 78.6.9). 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.

Output 78.6.9 Estimates

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 78.6.10 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 78.6.10 Least Squares Means for Cult, Irrig, and Their Interaction

Least Squares Means

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	16.81	<.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	0.3697	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	0.3697	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:

```
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 108×108 covariance matrix. See Wolfinger (1993) for further discussion of this example and additional analyses.

Example 78.7: 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.

```
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

$$
\begin{aligned}
Y_{ij} &= \mu + \tau_i + \epsilon_{ij} \\
i &= 1, \ldots, t = 4 \\
j &= 1, \ldots, r = 6 \\
\epsilon_{ij} &= N(0, \sigma_i^2)
\end{aligned}
$$

where Y_{ij} is the amount of fat absorbed by the jth batch of the ith fat type, and τ_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.

```
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 78.7.1.

Output 78.7.1 Heterogeneous Variance Analysis

The Mixed Procedure

Model Information	
Data Set	WORK.ABSORB
Dependent Variable	Absorbed
Covariance Structure	Variance Components
Group Effect	FatType
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Covariance Parameter Estimates

Cov Parm	Group	Estimate
Residual	FatType 1	178.00
Residual	FatType 2	60.4000
Residual	FatType 3	97.6000
Residual	FatType 4	67.6000

Solution for Fixed Effects

Effect	FatType	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:

$$\begin{aligned}
\text{Intercept} &: \mu + \tau_4 \\
\text{FatType1} &: \tau_1 - \tau_4 \\
\text{FatType2} &: \tau_2 - \tau_4 \\
\text{FatType3} &: \tau_3 - \tau_4 \\
\text{FatType4} &: 0
\end{aligned}$$

You can easily verify that these estimates are simple functions of the arithmetic means $\overline{y}_{i.}$ in the groups. For example, $\widehat{\mu + \tau_4} = \overline{y}_{4.} = 162.0$, $\widehat{\tau_1 - \tau_4} = \overline{y}_{1.} - \overline{y}_{4.} = 10.0$, 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 78.7.1). The estimated variance in the first group is two to three times larger than the variance in the other groups.

Output 78.7.2 Asymptotic Variances of Group Variance Estimates

Asymptotic Covariance Matrix of Estimates					
Row	Cov Parm	CovP1	CovP2	CovP3	CovP4
1	Residual	12674			
2	Residual		1459.26		
3	Residual			3810.30	
4	Residual				1827.90

In groups where the residual variance estimate is large, the precision of the estimate is also small (Output 78.7.2).

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 78.7.3):

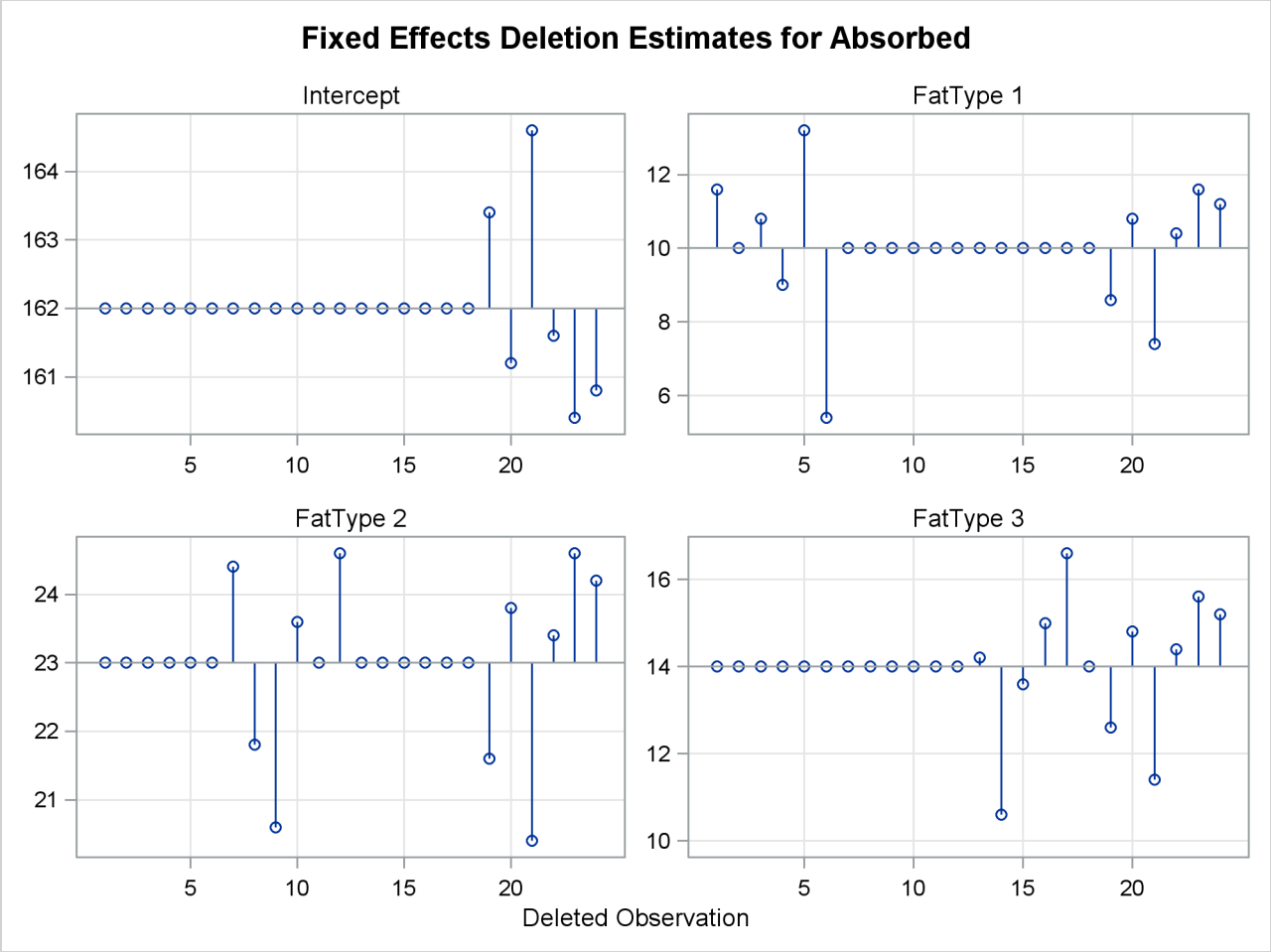
```
proc print data=inf label;
   var parm1-parm5 covp1-covp4;
run;
```

Output 78.7.3 Leave-One-Out Estimates

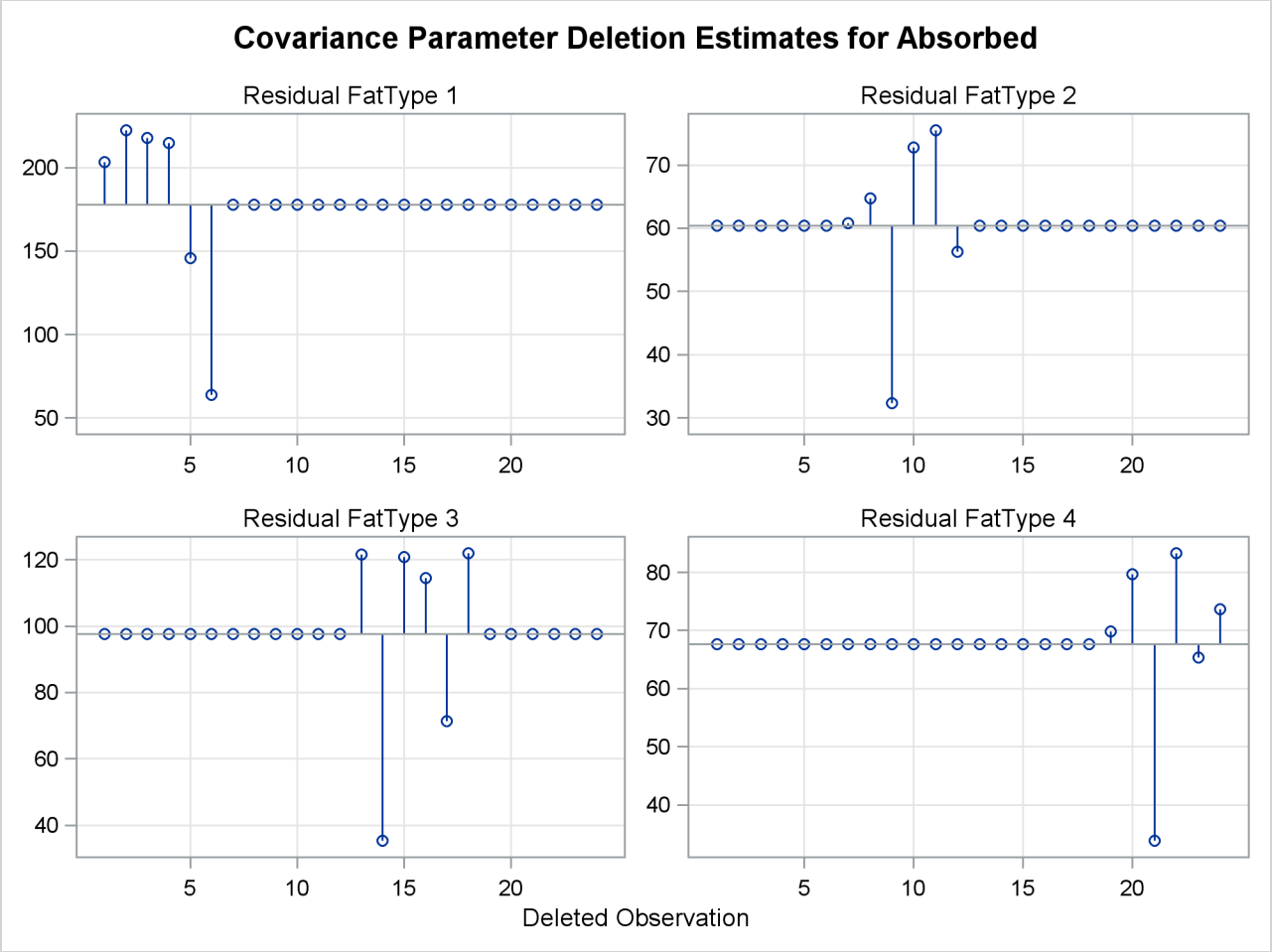
Obs	Intercept	FatType 1	FatType 2	FatType 3	FatType 4	Residual FatType 1	Residual FatType 2	Residual FatType 3	Residual FatType 4
1	162.00	11.600	23.000	14.000	0	203.30	60.400	97.60	67.600
2	162.00	10.000	23.000	14.000	0	222.47	60.400	97.60	67.600
3	162.00	10.800	23.000	14.000	0	217.68	60.400	97.60	67.600
4	162.00	9.000	23.000	14.000	0	214.99	60.400	97.60	67.600
5	162.00	13.200	23.000	14.000	0	145.70	60.400	97.60	67.600
6	162.00	5.400	23.000	14.000	0	63.80	60.400	97.60	67.600
7	162.00	10.000	24.400	14.000	0	178.00	60.795	97.60	67.600
8	162.00	10.000	21.800	14.000	0	178.00	64.691	97.60	67.600
9	162.00	10.000	20.600	14.000	0	178.00	32.296	97.60	67.600
10	162.00	10.000	23.600	14.000	0	178.00	72.797	97.60	67.600
11	162.00	10.000	23.000	14.000	0	178.00	75.490	97.60	67.600
12	162.00	10.000	24.600	14.000	0	178.00	56.285	97.60	67.600
13	162.00	10.000	23.000	14.200	0	178.00	60.400	121.68	67.600
14	162.00	10.000	23.000	10.600	0	178.00	60.400	35.30	67.600
15	162.00	10.000	23.000	13.600	0	178.00	60.400	120.79	67.600
16	162.00	10.000	23.000	15.000	0	178.00	60.400	114.50	67.600
17	162.00	10.000	23.000	16.600	0	178.00	60.400	71.30	67.600
18	162.00	10.000	23.000	14.000	0	178.00	60.400	121.98	67.600
19	163.40	8.600	21.600	12.600	0	178.00	60.400	97.60	69.799
20	161.20	10.800	23.800	14.800	0	178.00	60.400	97.60	79.698
21	164.60	7.400	20.400	11.400	0	178.00	60.400	97.60	33.800
22	161.60	10.400	23.400	14.400	0	178.00	60.400	97.60	83.292
23	160.40	11.600	24.600	15.600	0	178.00	60.400	97.60	65.299
24	160.80	11.200	24.200	15.200	0	178.00	60.400	97.60	73.677

The graphical displays in Output 78.7.4 and Output 78.7.5 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 6249.

Output 78.7.4 Fixed-Effects Deletion Estimates



Output 78.7.5 Covariance Parameter Deletion Estimates



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 78.7.4).

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 78.7.5). 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;
```

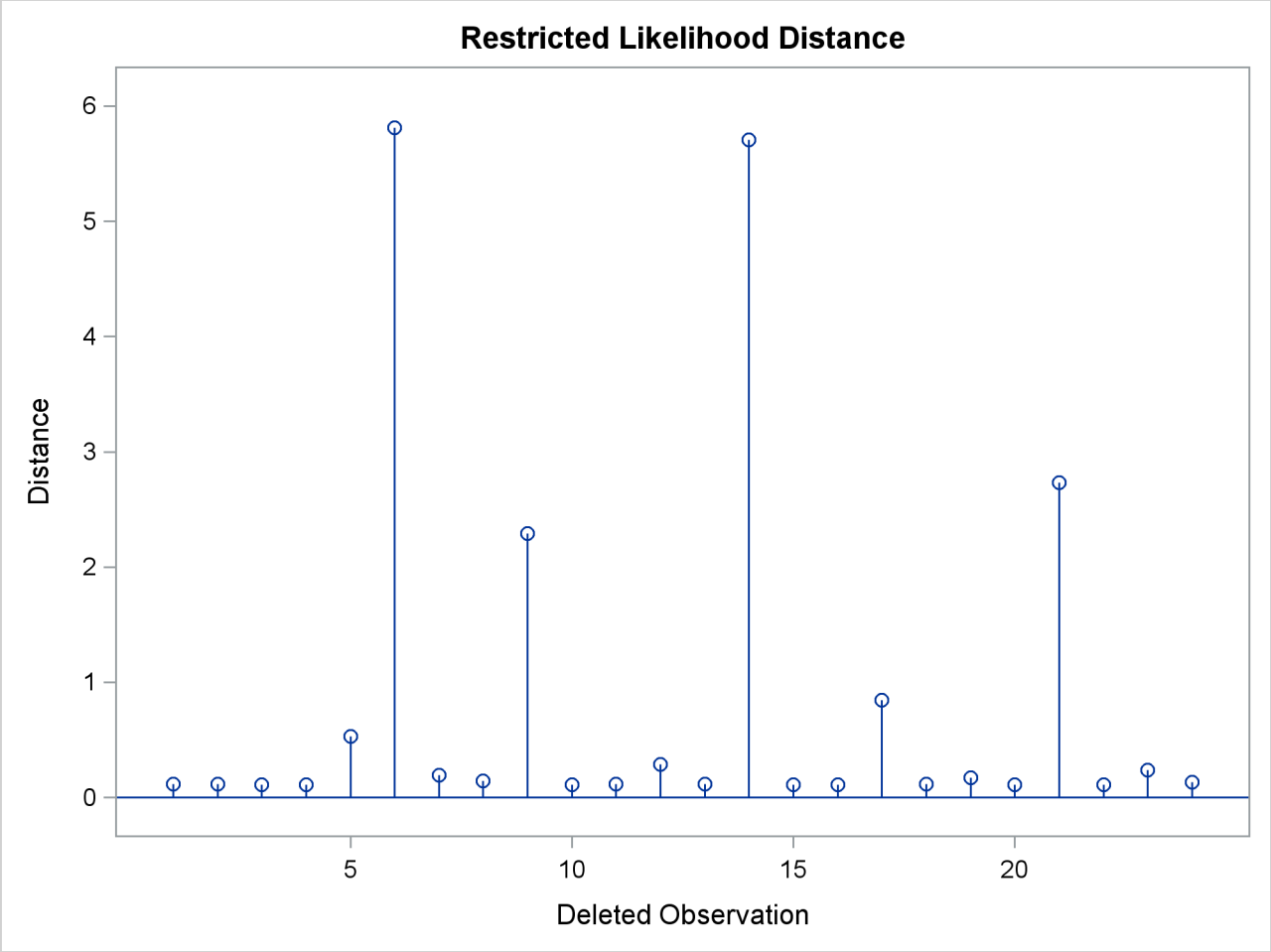
Output 78.7.6 Residual Diagnostics

Obs	Observed Value	Predicted Mean	Residual	PRESS Residual	Internally Studentized Residual	Externally Studentized Residual
1	164	172.0	-8.000	-9.600	-0.6569	-0.6146
2	172	172.0	0.000	0.000	0.0000	0.0000
3	168	172.0	-4.000	-4.800	-0.3284	-0.2970
4	177	172.0	5.000	6.000	0.4105	0.3736
5	156	172.0	-16.000	-19.200	-1.3137	-1.4521
6	195	172.0	23.000	27.600	1.8885	3.1544
7	178	185.0	-7.000	-8.400	-0.9867	-0.9835
8	191	185.0	6.000	7.200	0.8457	0.8172
9	197	185.0	12.000	14.400	1.6914	2.3131
10	182	185.0	-3.000	-3.600	-0.4229	-0.3852
11	185	185.0	0.000	-0.000	0.0000	0.0000
12	177	185.0	-8.000	-9.600	-1.1276	-1.1681
13	175	176.0	-1.000	-1.200	-0.1109	-0.0993
14	193	176.0	17.000	20.400	1.8850	3.1344
15	178	176.0	2.000	2.400	0.2218	0.1993
16	171	176.0	-5.000	-6.000	-0.5544	-0.5119
17	163	176.0	-13.000	-15.600	-1.4415	-1.6865
18	176	176.0	0.000	0.000	0.0000	0.0000
19	155	162.0	-7.000	-8.400	-0.9326	-0.9178
20	166	162.0	4.000	4.800	0.5329	0.4908
21	149	162.0	-13.000	-15.600	-1.7321	-2.4495
22	164	162.0	2.000	2.400	0.2665	0.2401
23	170	162.0	8.000	9.600	1.0659	1.0845
24	168	162.0	6.000	7.200	0.7994	0.7657

Observations 6, 9, 14, and 21 have large studentized residuals (Output 78.7.6). 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 78.7.7. Observations 6, 9, 14, and 21 clearly displace the REML solution more than any other observations.

Output 78.7.7 Restricted Likelihood Distance



The following statements list the restricted likelihood distance and various diagnostics related to the fixed-effects estimates (Output 78.7.8):

```
proc print data=inf label;
   var leverage observed CookD DFFITS CovRatio RLD;
run;
```

Output 78.7.8 Restricted Likelihood Distance and Fixed-Effects Diagnostics

Obs	Leverage	Observed Value	Cook's D	DFFITS	COVRATIO	Restr. Likelihood Distance
1	0.167	164	0.02157	-0.27487	1.3706	0.1178
2	0.167	172	0.00000	-0.00000	1.4998	0.1156
3	0.167	168	0.00539	-0.13282	1.4675	0.1124
4	0.167	177	0.00843	0.16706	1.4494	0.1117
5	0.167	156	0.08629	-0.64938	0.9822	0.5290
6	0.167	195	0.17831	1.41069	0.4301	5.8101
7	0.167	178	0.04868	-0.43982	1.2078	0.1935
8	0.167	191	0.03576	0.36546	1.2853	0.1451
9	0.167	197	0.14305	1.03446	0.6416	2.2909
10	0.167	182	0.00894	-0.17225	1.4463	0.1116
11	0.167	185	0.00000	-0.00000	1.4998	0.1156
12	0.167	177	0.06358	-0.52239	1.1183	0.2856
13	0.167	175	0.00061	-0.04441	1.4961	0.1151
14	0.167	193	0.17766	1.40175	0.4340	5.7044
15	0.167	178	0.00246	0.08915	1.4851	0.1139
16	0.167	171	0.01537	-0.22892	1.4078	0.1129
17	0.167	163	0.10389	-0.75423	0.8766	0.8433
18	0.167	176	0.00000	0.00000	1.4998	0.1156
19	0.167	155	0.04349	-0.41047	1.2390	0.1710
20	0.167	166	0.01420	0.21950	1.4148	0.1124
21	0.167	149	0.15000	-1.09545	0.6000	2.7343
22	0.167	164	0.00355	0.10736	1.4786	0.1133
23	0.167	170	0.05680	0.48500	1.1592	0.2383
24	0.167	168	0.03195	0.34245	1.3079	0.1353

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 78.7.8). 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:

```
proc print data=inf label;
   var iter CookDCP CovRatioCP;
run;
```

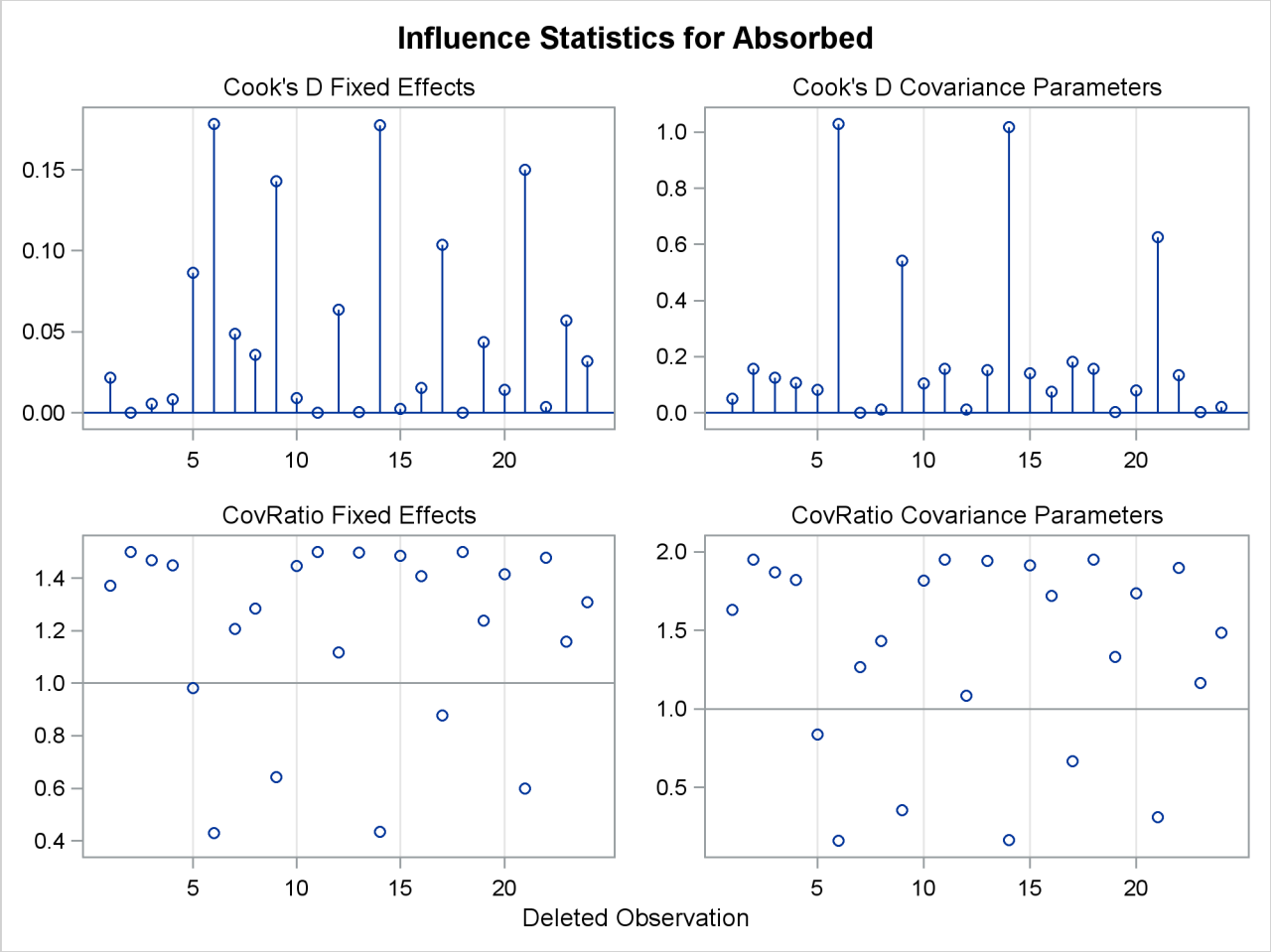
The same conclusions as for the fixed-effects estimates hold for the covariance parameter estimates. Observations 6, 9, 14, and 21 change the estimates and their precision considerably (Output 78.7.9, Output 78.7.10). All iterative updates converged within at most four iterations.

Output 78.7.9 Covariance Parameter Diagnostics

Obs	Iterations	Cook's D CovParms	COVRATIO CovParms
1	3	0.05050	1.6306
2	3	0.15603	1.9520
3	3	0.12426	1.8692
4	3	0.10796	1.8233
5	4	0.08232	0.8375
6	4	1.02909	0.1606
7	1	0.00011	1.2662
8	2	0.01262	1.4335
9	3	0.54126	0.3573
10	3	0.10531	1.8156
11	3	0.15603	1.9520
12	2	0.01160	1.0849
13	3	0.15223	1.9425
14	4	1.01865	0.1635
15	3	0.14111	1.9141
16	3	0.07494	1.7203
17	3	0.18154	0.6671
18	3	0.15603	1.9520
19	2	0.00265	1.3326
20	3	0.08008	1.7374
21	1	0.62500	0.3125
22	3	0.13472	1.8974
23	2	0.00290	1.1663
24	2	0.02020	1.4839

Output 78.7.10 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.

Output 78.7.10 Influence Diagnostics



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 78.8: Influence Analysis for Repeated Measures Data

This example revisits the repeated measures data of Pothoff and Roy (1964) that were analyzed in Example 78.2. 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 78.8.1:

```
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 78.8.1 Default Influence Statistics in Noniterative Analysis

The Mixed Procedure

Influence Diagnostics for Levels of Person

Person	Number of Observations in Level	PRESS Statistic	Cook's D
1	4	10.1716	0.01539
2	4	3.8187	0.03988
3	4	10.8448	0.02891
4	4	24.0339	0.04515
5	4	1.6900	0.01613
6	4	11.8592	0.01634
7	4	1.1887	0.00521
8	4	4.6717	0.02742
9	4	13.4244	0.03949
10	4	85.1195	0.13848
11	4	67.9397	0.09728
12	4	40.6467	0.04438
13	4	13.0304	0.00924
14	4	6.1712	0.00411
15	4	24.5702	0.12727
16	4	20.5266	0.01026
17	4	9.9917	0.01526
18	4	7.9355	0.01070
19	4	15.5955	0.01982
20	4	42.6845	0.01973
21	4	95.3282	0.10075
22	4	13.9649	0.03778
23	4	4.9656	0.01245
24	4	37.2494	0.15094
25	4	4.3756	0.03375
26	4	8.1448	0.03470
27	4	20.2913	0.02523

Each observation in the "Influence Diagnostics for Levels of Person" table in Output 78.8.1 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:

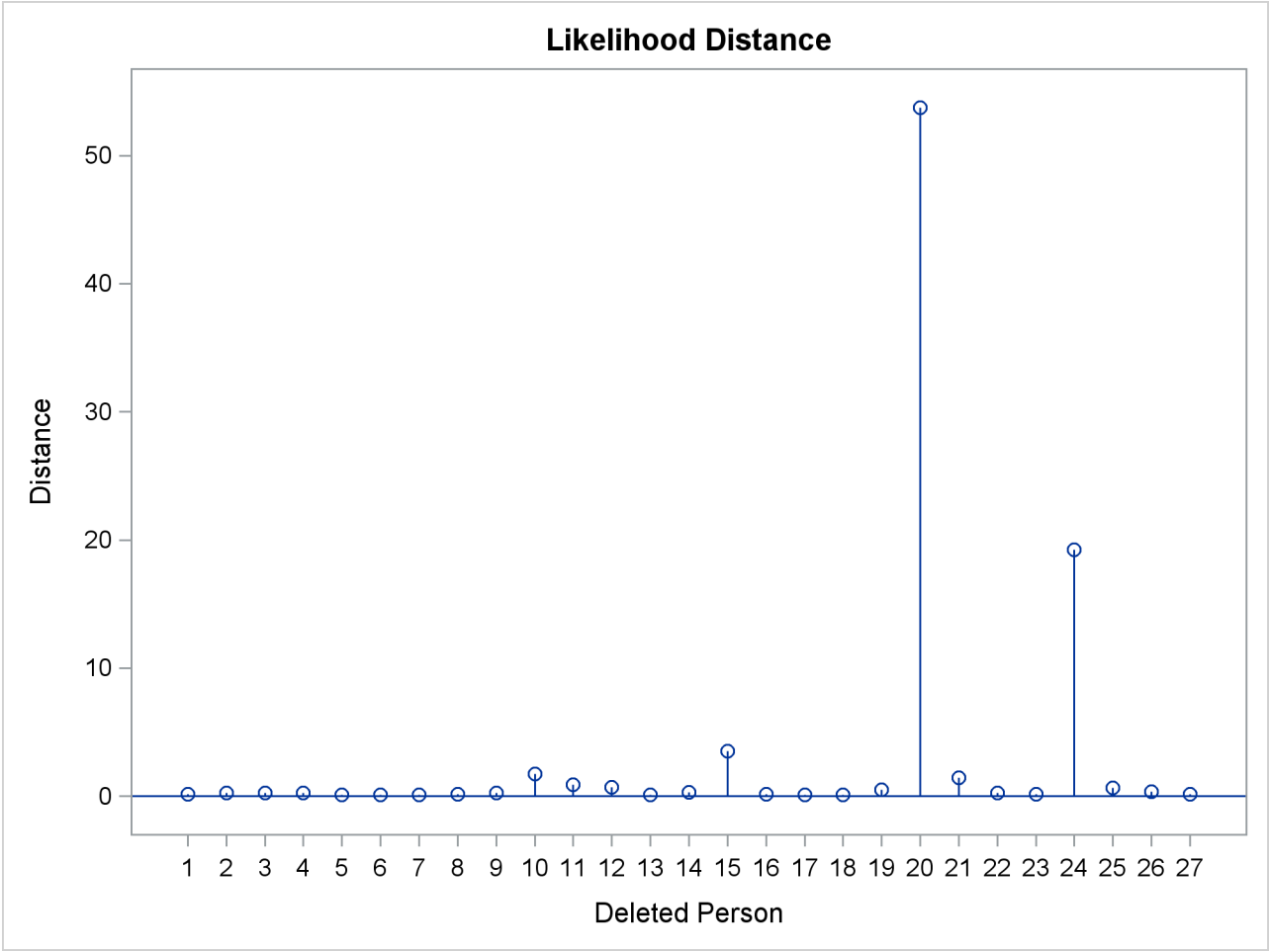
```
ods graphics on;

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;
```

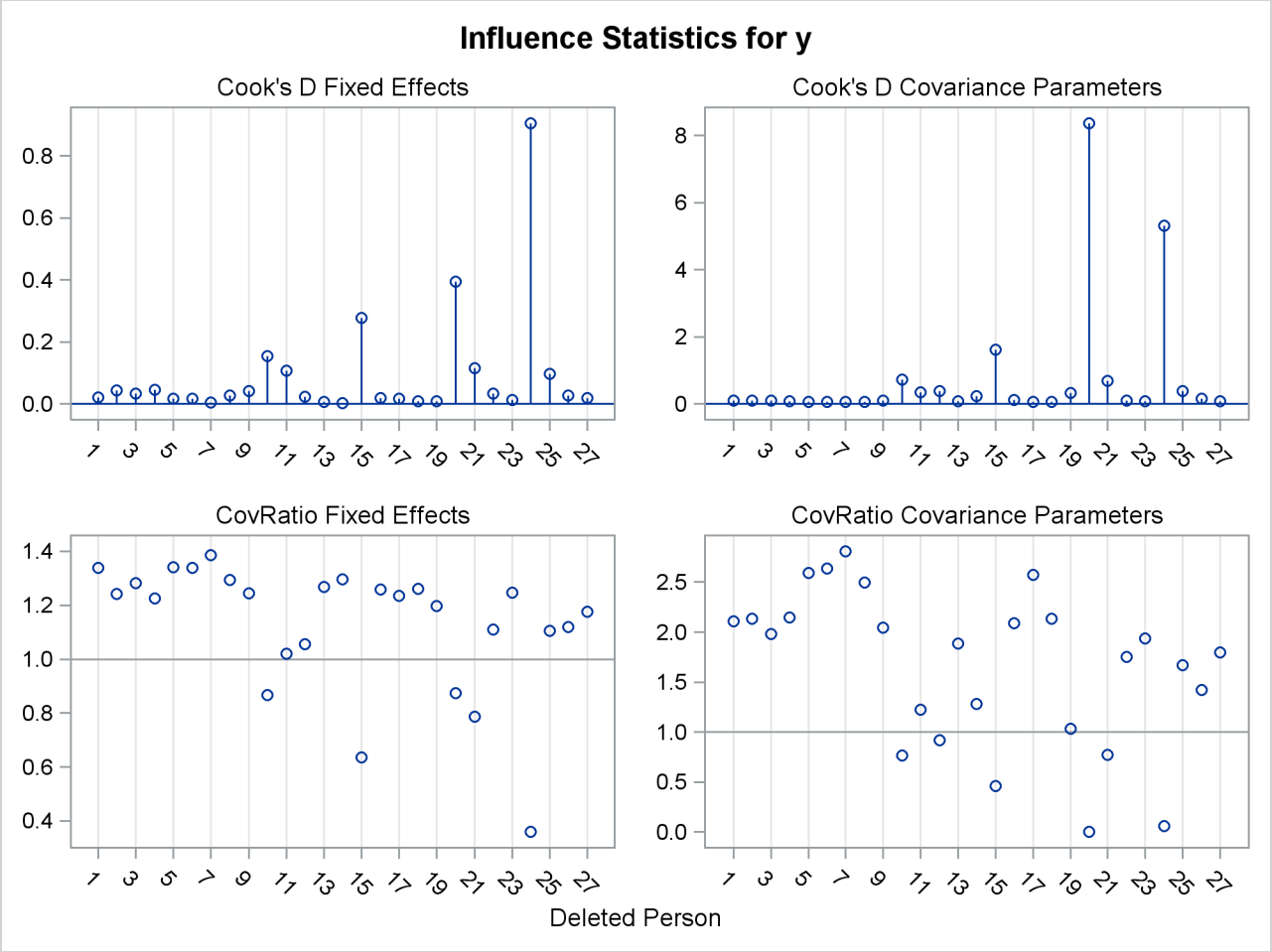
The number of additional iterations following removal of the observations for a particular subject is limited to five. Graphical displays of influence diagnostics are created when ODS Graphics is enabled. For general information about ODS Graphics, see Chapter 21, "Statistical Graphics Using ODS." For specific information about the graphics available in the MIXED procedure, see the section "ODS Graphics" on page 6249.

The MIXED procedure produces a plot of the restricted likelihood distance (Output 78.8.2) and a panel of diagnostics for fixed effects and covariance parameters (Output 78.8.3).

Output 78.8.2 Restricted Likelihood Distance



Output 78.8.3 Influence Diagnostics Panel



As judged by the restricted likelihood distance, subjects 20 and 24 clearly have the most influence on the overall analysis (Output 78.8.2).

Output 78.8.3 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 78.8.3 (CovRatio near 0).

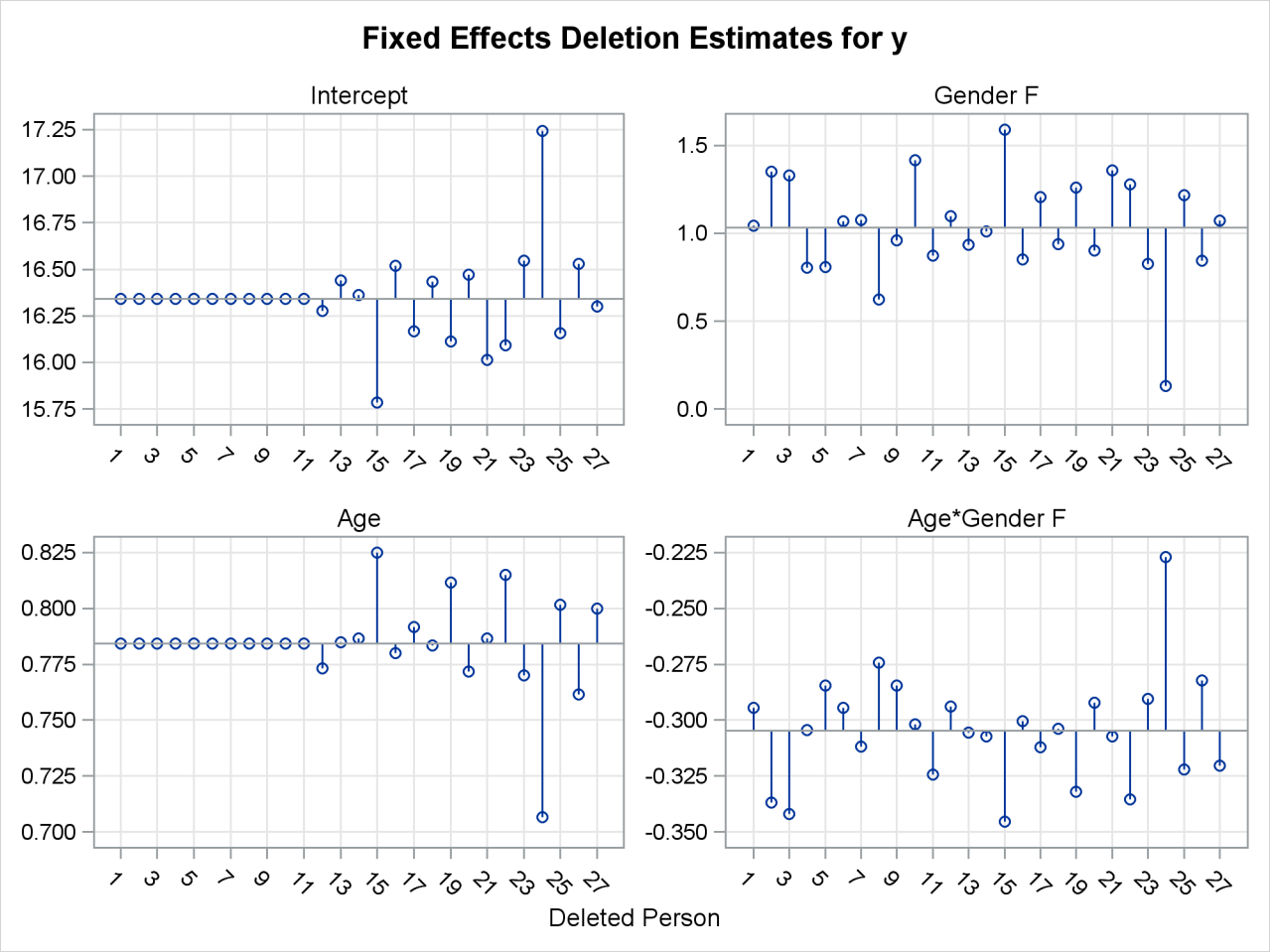
The child who exerts the greatest influence on the fixed effects is subject 24. Maybe surprisingly, this subject affects the variance-covariance matrix of the fixed effects more than subject 20 (small CovRatio in Output 78.8.3).

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 2×2 variance matrix of the random coefficients graphically:

```
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;
```

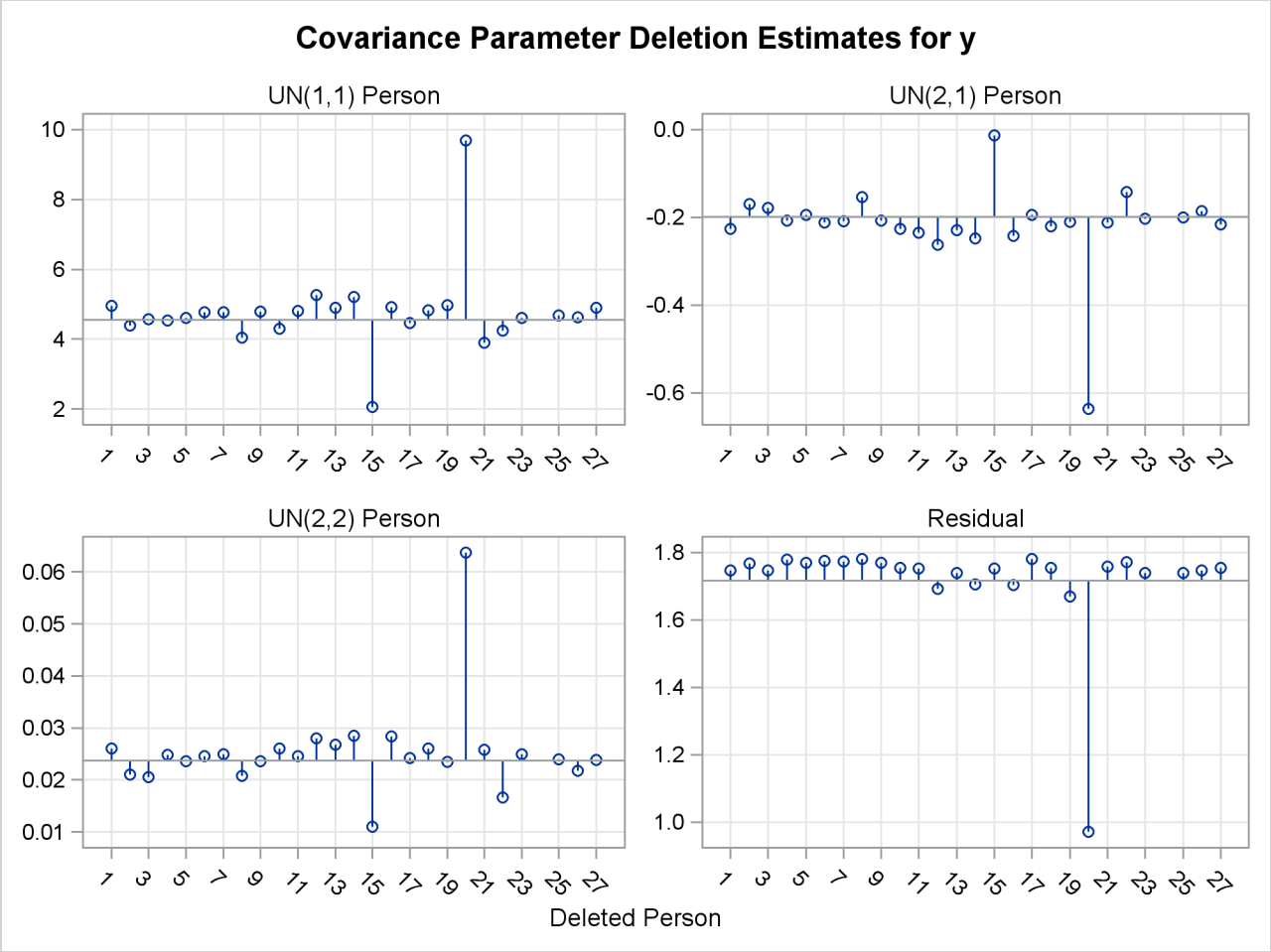
The PLOTS(ONLY)=INFLUENCEESTPLOT option restricts the graphical output from this PROC MIXED run to only the panels of deletion estimates (Output 78.8.4 and Output 78.8.5).

Output 78.8.4 Fixed-Effects Deletion Estimates



In Output 78.8.4 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.

Output 78.8.5 Covariance Parameter Deletion Estimates



The covariance parameter deletion estimates in Output 78.8.5 show several important features.

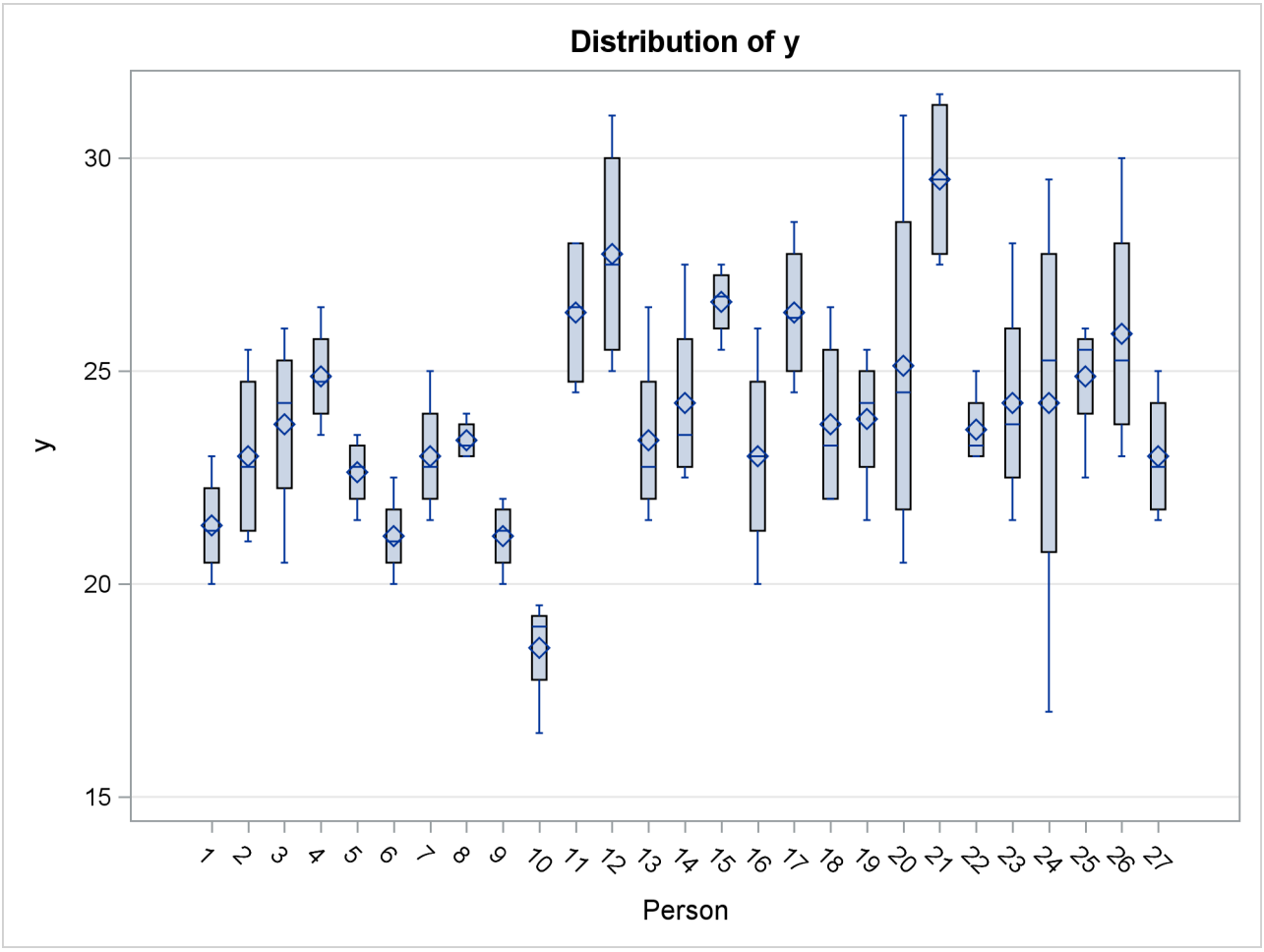
- The panels do not contain information about subject 24. Estimation of the **G** matrix following removal of that child did not yield a positive definite matrix. As a consequence, covariance parameter diagnostics are not produced for this subject.
- Subject 20 has great impact on the four covariance 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:

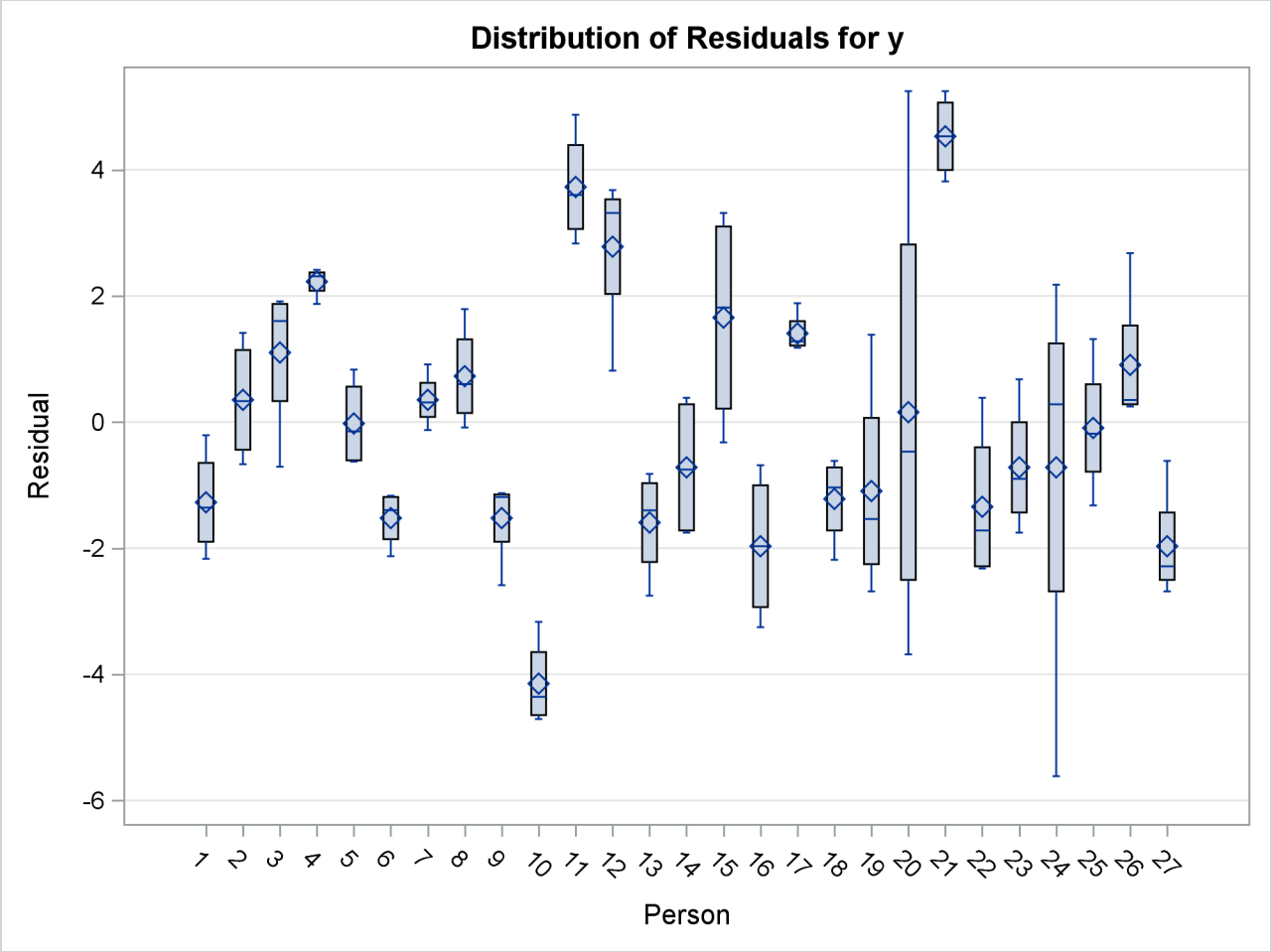
```
ods graphics on;
proc mixed data=pr method=ml
       plot=boxplot(observed marginal conditional subject);
   class person gender;
   model y = gender age gender*age;
   random intercept age / type=un subject=person;
run;
```

The specific boxplot options request a plot of the observed data (Output 78.8.6), the marginal residuals (Output 78.8.7), and the conditional residuals (Output 78.8.8). 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 78.8.6). 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 78.8.7). If child-specific inferences are desired, a model accounting for only Gender, Age, and Gender*Age effects is not adequate for these data.

Output 78.8.6 Distribution of Observed Values

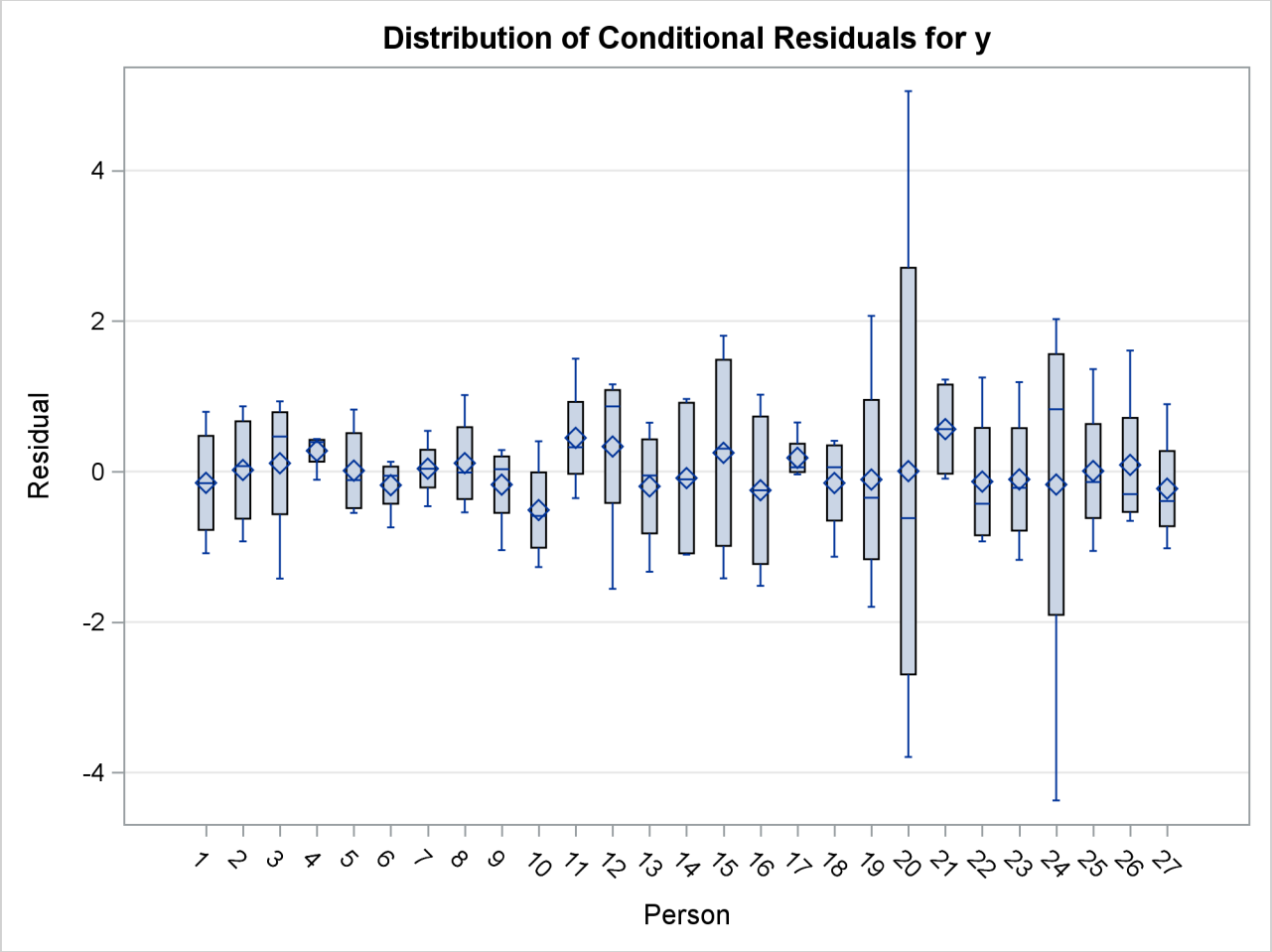


Output 78.8.7 Distribution of Marginal Residuals



The conditional residuals incorporate the EBLUPs for each child and enable you to examine whether the subject-specific model is adequate (Output 78.8.8). 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.

Output 78.8.8 Distribution of Conditional Residuals



Example 78.9: 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 78.1. Recall that variables A and B in the following statements represent the whole-plot and subplot factors, respectively:

```
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 78.9.1. The tests for fixed effects are shown in Output 78.9.2.

Output 78.9.1 Coefficients of Type 3 Estimable Functions

The Mixed Procedure

Type 3 Coefficients for A

Effect	A	B	Row1	Row2
Intercept				
A	1		1	
A	2			1
A	3		-1	-1
B		1		
B		2		
A*B	1	1	0.5	
A*B	1	2	0.5	
A*B	2	1		0.5
A*B	2	2		0.5
A*B	3	1	-0.5	-0.5
A*B	3	2	-0.5	-0.5

Type 3 Coefficients for B

Effect	A	B	Row1
Intercept			
A	1		
A	2		
A	3		
B		1	1
B		2	-1
A*B	1	1	0.3333
A*B	1	2	-0.333
A*B	2	1	0.3333
A*B	2	2	-0.333
A*B	3	1	0.3333
A*B	3	2	-0.333

Type 3 Coefficients for A*B

Effect	A	B	Row1	Row2
Intercept				
A	1			
A	2			
A	3			
B		1		
B		2		
A*B	1	1	1	
A*B	1	2	-1	
A*B	2	1		1
A*B	2	2		-1
A*B	3	1	-1	-1
A*B	3	2	1	1

Output 78.9.2 Type 3 Tests in Split-Plot Example

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
A	2	6	4.07	0.0764
B	1	9	19.39	0.0017
A*B	2	9	4.02	0.0566

If $\mu_{i.}$ denotes a whole-plot main effect mean, $\mu_{.j}$ denotes a subplot main effect mean, and μ_{ij} denotes a cell mean, the five components shown in Output 78.9.3 correspond to tests of the following:

- $H_0 : \mu_{1.} = \mu_{3.}$
- $H_0 : \mu_{2.} = \mu_{3.}$
- $H_0 : \mu_{.1} = \mu_{.2}$
- $H_0 : \mu_{11} - \mu_{12} = \mu_{31} - \mu_{32}$
- $H_0 : \mu_{21} - \mu_{22} = \mu_{31} - \mu_{32}$

Output 78.9.3 Type 3 L Components Table

L Components of Type 3 Tests of Fixed Effects

Effect	L Index	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

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:

```
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;
   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 78.9.4.

Output 78.9.4 Results from ESTIMATE Statements

The Mixed Procedure

Estimates

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 78.9.5 Parameter Estimates in Polynomial Model

The Mixed Procedure

Solution for Fixed Effects

Effect	Estimate	Standard 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 78.9.5. 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 78.9.6 Type 1 and Type 3 L Components

L Components of Type 1 Tests of Fixed Effects

Effect	L Index	Estimate	Standard Error	DF	t Value	Pr > |t|
x	1	0.1763	0.01259	16	14.01	<.0001
x*x	1	-0.04886	0.002449	16	-19.95	<.0001
x*x*x	1	-0.00306	0.000489	16	-6.27	<.0001

L Components of Type 3 Tests of Fixed Effects

Effect	L Index	Estimate	Standard Error	DF	t Value	Pr > |t|
x	1	0.3726	0.1426	16	2.61	0.0189
x*x	1	0.04756	0.01558	16	3.05	0.0076
x*x*x	1	-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 78.10: 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 78.10.1.

Output 78.10.1 Analysis of LS-Means with Isotonic Contrasts

The Mixed Procedure

Least Squares Means Estimates
Adjustment for Multiplicity: Simulated

Effect	Label	Estimate	Standard Error	DF	t Value	Tails	Pr > t	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
Temp	avg(1:1)<avg(2:4)	0.8000	0.1906	16	4.20	Upper	0.0003	0.0010	0.05	0.4672	Infty	0.3771	Infty
Temp	avg(1:2)<avg(3:4)	0.7000	0.1651	16	4.24	Upper	0.0003	0.0009	0.05	0.4118	Infty	0.3337	Infty
Temp	avg(1:3)<avg(4:4)	0.4000	0.1906	16	2.10	Upper	0.0260	0.0625	0.05	0.06721	Infty	-0.02291	Infty

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

Akaike, H. (1974). "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* AC-19:716–723.

Akritas, M. G., Arnold, S. F., and Brunner, E. (1997). "Nonparametric Hypotheses and Rank Statistics for Unbalanced Factorial Designs." *Journal of the American Statistical Association* 92:258–265.

Allen, D. M. (1974). "The Relationship between Variable Selection and Data Augmentation and a Method of Prediction." *Technometrics* 16:125–127.

Bates, D. M., and Watts, D. G. (1988). *Nonlinear Regression Analysis and Its Applications*. New York: John Wiley & Sons.

Beckman, R. J., Nachtsheim, C. J., and Cook, R. D. (1987). "Diagnostics for Mixed-Model Analysis of Variance." *Technometrics* 29:413–426.

Belsley, D. A., Kuh, E., and Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley & Sons.

Box, G. E. P., and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. New York: John Wiley & Sons.

Bozdogan, H. (1987). "Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions." *Psychometrika* 52:345–370.

Brown, H., and Prescott, R. (1999). *Applied Mixed Models in Medicine*. New York: John Wiley & Sons.

Brownie, C., Bowman, D. T., and Burton, J. W. (1993). "Estimating Spatial Variation in Analysis of Data from Yield Trials: A Comparison of Methods." *Agronomy Journal* 85:1244–1253.

Brownie, C., and Gumpertz, M. L. (1997). "Validity of Spatial Analysis of Large Field Trials." *Journal of Agricultural, Biological, and Environmental Statistics* 2:1–23.

Brunner, E., Dette, H., and Munk, A. (1997). "Box-Type Approximations in Nonparametric Factorial Designs." *Journal of the American Statistical Association* 92:1494–1502.

Brunner, E., Domhof, S., and Langer, F. (2002). *Nonparametric Analysis of Longitudinal Data in Factorial Experiments*. New York: John Wiley & Sons.

Burdick, R. K., and Graybill, F. A. (1992). *Confidence Intervals on Variance Components*. New York: Marcel Dekker.

Burnham, K. P., and Anderson, D. R. (1998). *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York: Springer-Verlag.

Carlin, B. P., and Louis, T. A. (1996). *Bayes and Empirical Bayes Methods for Data Analysis*. London: Chapman & Hall.

Carroll, R. J., and Ruppert, D. (1988). *Transformation and Weighting in Regression*. London: Chapman & Hall.

Chilès, J. P., and Delfiner, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. New York: John Wiley & Sons.

Christensen, R., Pearson, L. M., and Johnson, W. (1992). "Case-Deletion Diagnostics for Mixed Models." *Technometrics* 34:38–45.

Cook, R. D. (1977). "Detection of Influential Observations in Linear Regression." *Technometrics* 19:15–18.

Cook, R. D. (1979). "Influential Observations in Linear Regression." *Journal of the American Statistical Association* 74:169–174.

Cook, R. D., and Weisberg, S. (1982). *Residuals and Influence in Regression*. New York: Chapman & Hall.

Cressie, N. (1993). *Statistics for Spatial Data*. Rev. ed. New York: John Wiley & Sons.

Crowder, M. J., and Hand, D. J. (1990). *Analysis of Repeated Measures*. New York: Chapman & Hall.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society, Series B* 39:1–38.

Diggle, P. J. (1988). "An Approach to the Analysis of Repeated Measurements." *Biometrics* 44:959–971.

Diggle, P. J., Liang, K.-Y., and Zeger, S. L. (1994). *Analysis of Longitudinal Data*. Oxford: Clarendon Press.

Dunnett, C. W. (1980). "Pairwise Multiple Comparisons in the Unequal Variance Case." *Journal of the American Statistical Association* 75:796–800.

Edwards, D., and Berry, J. J. (1987). "The Efficiency of Simulation-Based Multiple Comparisons." *Biometrics* 43:913–928.

Everitt, B. S. (1995). "The Analysis of Repeated Measures: A Practical Review with Examples." *Journal of the Royal Statistical Society, Series D* 44:113–135. http://www.jstor.org/stable/2348622.

Fai, A. H. T., and Cornelius, P. L. (1996). "Approximate *F*-Tests of Multiple Degree of Freedom Hypotheses in Generalized Least Squares Analyses of Unbalanced Split-Plot Experiments." *Journal of Statistical Computation and Simulation* 54:363–378.

Federer, W. T., and Wolfinger, R. D. (1998). "SAS Code for Recovering Intereffect Information in Experiments with Incomplete Block and Lattice Rectangle Designs." *Agronomy Journal* 90:545–551.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*. New York: John Wiley & Sons.

Fuller, W. A., and Battese, G. E. (1973). "Transformations for Estimation of Linear Models with Nested Error Structure." *Journal of the American Statistical Association* 68:626–632.

Galecki, A. T. (1994). "General Class of Covariance Structures for Two or More Repeated Factors in Longitudinal Data Analysis." *Communications in Statistics—Theory and Methods* 23:3105–3109.

Games, P. A., and Howell, J. F. (1976). "Pairwise Multiple Comparison Procedures with Unequal *n*'s and/or Variances: A Monte Carlo Study." *Journal of Educational Statistics* 1:113–125.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. F. M. (1990). "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling." *Journal of the American Statistical Association* 85:972–985.

Ghosh, M. (1992). "Discussion of Schervish, M., 'Bayesian Analysis of Linear Models'." In *Bayesian Statistics*, vol. 4, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, 432–433. Oxford: Clarendon Press.

Giesbrecht, F. G. (1989). *A General Structure for the Class of Mixed Linear Models*. Southern Cooperative Series Bulletin 343, Louisiana Agricultural Experiment Station, Baton Rouge.

Giesbrecht, F. G., and Burns, J. C. (1985). "Two-Stage Analysis Based on a Mixed Model: Large-Sample Asymptotic Theory and Small-Sample Simulation Results." *Biometrics* 41:477–486.

Golub, G. H., and Van Loan, C. F. (1989). *Matrix Computations*. 2nd ed. Baltimore: Johns Hopkins University Press.

Goodnight, J. H. (1978). *Tests of Hypotheses in Fixed-Effects Linear Models*. Technical Report R-101, SAS Institute Inc., Cary, NC.

Goodnight, J. H. (1979). "A Tutorial on the Sweep Operator." *American Statistician* 33:149–158.

Goodnight, J. H., and Hemmerle, W. J. (1979). "A Simplified Algorithm for the W-Transformation in Variance Component Estimation." *Technometrics* 21:265–268.

Gotway, C. A., and Stroup, W. W. (1997). "A Generalized Linear Model Approach to Spatial Data and Prediction." *Journal of Agricultural, Biological, and Environmental Statistics* 2:157–187.

Greenhouse, S. W., and Geisser, S. (1959). "On Methods in the Analysis of Profile Data." *Psychometrika* 32:95–112.

Gregoire, T. G., Schabenberger, O., and Barrett, J. P. (1995). "Linear Modelling of Irregularly Spaced, Unbalanced, Longitudinal Data from Permanent Plot Measurements." *Canadian Journal of Forest Research* 25:137–156.

Handcock, M. S., and Stein, M. L. (1993). "A Bayesian Analysis of Kriging." *Technometrics* 35:403–410.

Handcock, M. S., and Wallis, J. R. (1994). "An Approach to Statistical Spatial-Temporal Modeling of Meteorological Fields." *Journal of the American Statistical Association* 89:368–390. With discussion.

Hanks, R. J., Sisson, D. V., Hurst, R. L., and Hubbard, K. G. (1980). "Statistical Analysis of Results from Irrigation Experiments Using the Line-Source Sprinkler System." *Soil Science Society American Journal* 44:886–888.

Hannan, E. J., and Quinn, B. G. (1979). "The Determination of the Order of an Autoregression." *Journal of the Royal Statistical Society, Series B* 41:190–195.

Hartley, H. O., and Rao, J. N. K. (1967). "Maximum-Likelihood Estimation for the Mixed Analysis of Variance Model." *Biometrika* 54:93–108.

Harville, D. A. (1977). "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems." *Journal of the American Statistical Association* 72:320–338.

Harville, D. A. (1988). "Mixed-Model Methodology: Theoretical Justifications and Future Directions." In *Proceedings of the Statistical Computing Section*, 41–49. Alexandria, VA: American Statistical Association.

Harville, D. A. (1990). "BLUP (Best Linear Unbiased Prediction), and Beyond." In *Advances in Statistical Methods for Genetic Improvement of Livestock*, edited by D. Gianola and K. Hammond, 239–276. Vol. 18 of Advanced Series in Agricultural Sciences. Berlin: Springer-Verlag.

Harville, D. A., and Jeske, D. R. (1992). "Mean Squared Error of Estimation or Prediction under a General Linear Model." *Journal of the American Statistical Association* 87:724–731.

Hemmerle, W. J., and Hartley, H. O. (1973). "Computing Maximum Likelihood Estimates for the Mixed AOV Model Using the W-Transformation." *Technometrics* 15:819–831.

Henderson, C. R. (1984). *Applications of Linear Models in Animal Breeding*. Guelph, ON: University of Guelph.

Henderson, C. R. (1990). "Statistical Method in Animal Improvement: Historical Overview." In *Advances in Statistical Methods for Genetic Improvement of Livestock*, 1–14. New York: Springer-Verlag.

Hirotsu, C., and Srivastava, M. (2000). "Simultaneous Confidence Intervals Based on One-Sided Max *t* Test." *Statistics and Probability Letters* 49:25–37.

Hsu, J. C. (1992). "The Factor Analytic Approach to Simultaneous Inference in the General Linear Model." *Journal of Computational and Graphical Statistics* 1:151–168.

Huber, P. J. (1967). "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions." *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* 1:221–233.

Hurtado, G. I. (1993). "Detection of Influential Observations in Linear Mixed Models." Ph.D. diss., Department of Statistics, North Carolina State University.

Hurvich, C. M., and Tsai, C.-L. (1989). "Regression and Time Series Model Selection in Small Samples." *Biometrika* 76:297–307.

Huynh, H., and Feldt, L. S. (1970). "Conditions Under Which Mean Square Ratios in Repeated Measurements Designs Have Exact F-Distributions." *Journal of the American Statistical Association* 65:1582–1589.

Jennrich, R. I., and Schluchter, M. D. (1986). "Unbalanced Repeated-Measures Models with Structured Covariance Matrices." *Biometrics* 42:805–820.

Johnson, D. E., Chaudhuri, U. N., and Kanemasu, E. T. (1983). "Statistical Analysis of Line-Source Sprinkler Irrigation Experiments and Other Nonrandomized Experiments Using Multivariate Methods." *Soil Science Society American Journal* 47:309–312.

Jones, R. H., and Boadi-Boateng, F. (1991). "Unequally Spaced Longitudinal Data with AR(1) Serial Correlation." *Biometrics* 47:161–175.

Kackar, R. N., and Harville, D. A. (1984). "Approximations for Standard Errors of Estimators of Fixed and Random Effects in Mixed Linear Models." *Journal of the American Statistical Association* 79:853–862.

Kass, R. E., and Steffey, D. (1989). "Approximate Bayesian Inference in Conditionally Independent Hierarchical Models (Parametric Empirical Bayes Models)." *Journal of the American Statistical Association* 84:717–726.

Kenward, M. G. (1987). "A Method for Comparing Profiles of Repeated Measurements." *Journal of the Royal Statistical Society, Series C* 36:296–308.

Kenward, M. G., and Roger, J. H. (1997). "Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood." *Biometrics* 53:983–997.

Kenward, M. G., and Roger, J. H. (2009). "An Improved Approximation to the Precision of Fixed Effects from Restricted Maximum Likelihood." *Computational Statistics and Data Analysis* 53:2583–2595.

Keselman, H. J., Algina, J., Kowalchuk, R. K., and Wolfinger, R. D. (1998). "A Comparison of Two Approaches for Selecting Covariance Structures in the Analysis of Repeated Measures." *Communications in Statistics—Simulation and Computation* 27:591–604.

Keselman, H. J., Algina, J., Kowalchuk, R. K., and Wolfinger, R. D. (1999). "A Comparison of Recent Approaches to the Analysis of Repeated Measurements." *British Journal of Mathematical and Statistical Psychology* 52:63–78.

Kramer, C. Y. (1956). "Extension of Multiple Range Tests to Group Means with Unequal Numbers of Replications." *Biometrics* 12:307–310.

Laird, N. M., Lange, N. T., and Stram, D. O. (1987). "Maximum Likelihood Computations with Repeated Measures: Application of the EM Algorithm." *Journal of the American Statistical Association* 82:97–105.

Laird, N. M., and Ware, J. H. (1982). "Random-Effects Models for Longitudinal Data." *Biometrics* 38:963–974.

LaMotte, L. R. (1973). "Quadratic Estimation of Variance Components." *Biometrics* 29:311–330.

Liang, K.-Y., and Zeger, S. L. (1986). "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika* 73:13–22.

Lindsey, J. K. (1993). *Models for Repeated Measurements*. Oxford: Clarendon Press.

Lindstrom, M. J., and Bates, D. M. (1988). "Newton-Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data." *Journal of the American Statistical Association* 83:1014–1022.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Schabenberger, O. (2006). *SAS for Mixed Models*. 2nd ed. Cary, NC: SAS Institute Inc.

Little, R. J. A. (1995). "Modeling the Drop-Out Mechanism in Repeated-Measures Studies." *Journal of the American Statistical Association* 90:1112–1121.

Louis, T. A. (1988). "General Methods for Analyzing Repeated Measures." *Statistics in Medicine* 7:29–45.

Macchiavelli, R. E., and Arnold, S. F. (1994). "Variable Order Ante-dependence Models." *Communications in Statistics—Theory and Methods* 23:2683–2699.

Marx, D., and Thompson, K. (1987). *Practical Aspects of Agricultural Kriging*. Bulletin 903, Arkansas Agricultural Experiment Station, Fayetteville.

Matérn, B. (1986). *Spatial Variation*. 2nd ed. New York: Springer-Verlag.

McKeon, J. J. (1974). "F Approximations to the Distribution of Hotelling's T_0^2." *Biometrika* 61:381–383.

McLean, R. A., and Sanders, W. L. (1988). "Approximating Degrees of Freedom for Standard Errors in Mixed Linear Models." In *Proceedings of the Statistical Computing Section*, 50–59. Alexandria, VA: American Statistical Association.

McLean, R. A., Sanders, W. L., and Stroup, W. W. (1991). "A Unified Approach to Mixed Linear Models." *American Statistician* 45:54–64.

Milliken, G. A., and Johnson, D. E. (1992). *Designed Experiments.* Vol. 1 of Analysis of Messy Data. Reprint edition. New York: Chapman & Hall.

Moriguchi, S., ed. (1976). *Statistical Method for Quality Control*. Tokyo: Japan Standards Association. In Japanese.

Murray, D. M. (1998). *Design and Analysis of Group-Randomized Trials*. New York: Oxford University Press.

Myers, R. H. (1990). *Classical and Modern Regression with Applications*. 2nd ed. Belmont, CA: PWS-Kent.

Obenchain, R. L. (1990). *STABLSIM.EXE, Version 9010.* Unpublished C code. Indianapolis: Eli Lilly.

Patel, H. I. (1991). "Analysis of Incomplete Data from a Clinical Trial with Repeated Measurements." *Biometrika* 78:609–619.

Patterson, H. D., and Thompson, R. (1971). "Recovery of Inter-block Information When Block Sizes Are Unequal." *Biometrika* 58:545–554.

Pillai, K. C. S., and Samson, P., Jr. (1959). "On Hotelling's Generalization of T^2." *Biometrika* 46:160–168.

Pothoff, R. F., and Roy, S. N. (1964). "A Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems." *Biometrika* 51:313–326.

Prasad, N. G. N., and Rao, J. N. K. (1990). "The Estimation of Mean Squared Error of Small-Area Estimators." *Journal of the American Statistical Association* 85:163–171.

Pringle, R. M., and Rayner, A. A. (1971). *Generalized Inverse Matrices with Applications to Statistics*. New York: Hafner Publishing.

Rao, C. R. (1972). "Estimation of Variance and Covariance Components in Linear Models." *Journal of the American Statistical Association* 67:112–115.

Ripley, B. D. (1987). *Stochastic Simulation*. New York: John Wiley & Sons.

Robinson, G. K. (1991). "That BLUP Is a Good Thing: The Estimation of Random Effects." *Statistical Science* 6:15–51.

Rubin, D. B. (1976). "Inference and Missing Data." *Biometrika* 63:581–592.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). "Design and Analysis of Computer Experiments." *Statistical Science* 4:409–435.

Schabenberger, O., and Gotway, C. A. (2005). *Statistical Methods for Spatial Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC.

Schervish, M. J. (1992). "Bayesian Analysis of Linear Models." In *Bayesian Statistics*, vol. 4, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, 419–434. Oxford: Clarendon Press.

Schluchter, M. D., and Elashoff, J. D. (1990). "Small-Sample Adjustments to Tests with Unbalanced Repeated Measures Assuming Several Covariance Structures." *Journal of Statistical Computation and Simulation* 37:69–87.

Schwarz, G. (1978). "Estimating the Dimension of a Model." *Annals of Statistics* 6:461–464.

Searle, S. R. (1971). *Linear Models*. New York: John Wiley & Sons.

Searle, S. R. (1982). *Matrix Algebra Useful for Statisticians*. New York: John Wiley & Sons.

Searle, S. R. (1988). "Mixed Models and Unbalanced Data: Wherefrom, Whereat, and Whereto?" *Communications in Statistics—Theory and Methods* 17:935–968.

Searle, S. R., Casella, G., and McCulloch, C. E. (1992). *Variance Components*. New York: John Wiley & Sons.

Self, S. G., and Liang, K.-Y. (1987). "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions." *Journal of the American Statistical Association* 82:605–610.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons.

Singer, J. D. (1998). "Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models." *Journal of Educational and Behavioral Statistics* 23:323–355.

Smith, A. F. M., and Gelfand, A. E. (1992). "Bayesian Statistics without Tears: A Sampling-Resampling Perspective." *American Statistician* 46:84–88.

Snedecor, G. W., and Cochran, W. G. (1980). *Statistical Methods*. 7th ed. Ames: Iowa State University Press.

Steel, R. G. D., Torrie, J. H., and Dickey, D. A. (1997). *Principles and Procedures of Statistics: A Biometrical Approach*. 3rd ed. New York: McGraw-Hill.

Stram, D. O., and Lee, J. W. (1994). "Variance Components Testing in the Longitudinal Mixed Effects Model." *Biometrics* 50:1171–1177.

Stroup, W. W. (1989a). "Predictable Functions and Prediction Space in the Mixed Model Procedure." *Applications of Mixed Models in Agriculture and Related Disciplines,* 39–48. Southern Cooperative Series Bulletin No. 343, Louisiana Agricultural Experiment Station, Baton Rouge.

Stroup, W. W. (1989b). "Use of Mixed Model Procedure to Analyze Spatially Correlated Data: An Example Applied to a Line-Source Sprinkler Irrigation Experiment." *Applications of Mixed Models in Agriculture and Related Disciplines,* 104–122. Southern Cooperative Series Bulletin No. 343, Louisiana Agricultural Experiment Station, Baton Rouge.

Stroup, W. W., Baenziger, P. S., and Mulitze, D. K. (1994). "Removing Spatial Variation from Wheat Yield Trials: A Comparison of Methods." *Crop Science* 86:62–66.

Sullivan, L. M., Dukes, K. A., and Losina, E. (1999). "An Introduction to Hierarchical Linear Modelling." *Statistics in Medicine* 18:855–888.

Swallow, W. H., and Monahan, J. F. (1984). "Monte Carlo Comparison of ANOVA, MIVQUE, REML, and ML Estimators of Variance Components." *Technometrics* 28:47–57.

Tamhane, A. C. (1979). "A Comparison of Procedures for Multiple Comparisons of Means with Unequal Variances." *Journal of the American Statistical Association* 74:471–480.

Tierney, L. (1994). "Markov Chains for Exploring Posterior Distributions." *Annals of Statistics* 22:1701–1762.

Verbeke, G., and Molenberghs, G., eds. (1997). *Linear Mixed Models in Practice: A SAS-Oriented Approach*. New York: Springer.

Verbeke, G., and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer.

Westfall, P. H., Tobias, R. D., Rom, D., Wolfinger, R. D., and Hochberg, Y. (1999). *Multiple Comparisons and Multiple Tests Using the SAS System*. Cary, NC: SAS Institute Inc.

Westfall, P. H., and Young, S. S. (1993). *Resampling-Based Multiple Testing: Examples and Methods for* p*-Value Adjustment*. New York: John Wiley & Sons.

White, H. (1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48:817–838.

Whittle, P. (1954). "On Stationary Processes in the Plane." *Biometrika* 41:434–449.

Winer, B. J. (1971). *Statistical Principles in Experimental Design*. 2nd ed. New York: McGraw-Hill.

Wolfinger, R. D. (1993). "Covariance Structure Selection in General Mixed Models." *Communications in Statistics—Simulation and Computation* 22:1079–1106.

Wolfinger, R. D. (1996). "Heterogeneous Variance-Covariance Structures for Repeated Measures." *Journal of Agricultural, Biological, and Environmental Statistics* 1:205–230.

Wolfinger, R. D. (1997). "An Example of Using Mixed Models and PROC MIXED for Longitudinal Data." *Journal of Biopharmaceutical Statistics* 7:481–500.

Wolfinger, R. D., and Chang, M. (1995). "Comparing the SAS GLM and MIXED Procedures for Repeated Measures." In *Proceedings of the Twentieth Annual SAS Users Group Conference*, 1172–1182. Cary, NC: SAS Institute Inc. http://www.sascommunity.org/sugi/SUGI95/Sugi-95-198%20Wolfinger%20Chang.pdf.

Wolfinger, R. D., Tobias, R. D., and Sall, J. (1991). "Mixed Models: A Future Direction." In *Proceedings of the Sixteenth Annual SAS Users Group Conference*, 1380–1388. Cary, NC: SAS Institute Inc. http://www.sascommunity.org/sugi/SUGI91/Sugi-91-249%20Wolfinger%20Tobias%20Sall.pdf.

Wolfinger, R. D., Tobias, R. D., and Sall, J. (1994). "Computing Gaussian Likelihoods and Their Derivatives for General Linear Mixed Models." *SIAM Journal on Scientific Computing* 15:1294–1310.

Wright, S. P. (1994). *Adjusted* F *Tests for Repeated Measures with the MIXED Procedure*. Knoxville: Statistics Department, University of Tennessee.

Zimmerman, D. L., and Harville, D. A. (1991). "A Random Field Approach to the Analysis of Field-Plot Experiments and Other Spatial Experiments." *Biometrics* 47:223–239.

Subject Index

Syntax Index