NEW FEATURES IN GLM AND VARCOMP

J. H. Goodnight, SAS Institute

1. Introduction

The GLM procedure and the VARCOMP procedure have several new features that will be available in the 1979 version of SAS. GLM has been expanded to include three new statements:

- ESTIMATE
- CONTRAST
- RANDOM

VARCOMP has two new methods of estimating variance components. One of the new methods, MIVQUEO, is now the default method and is extremely efficient with respect to core and CPU time.

2. GLM ESTIMATE Statement

The ESTIMATE statement is used to specify an L vector for estimating the linear function of parameters LB. All of the elements of the L vector may be given; alternately, only certain portions of the L vector may be given, in which case the remaining elements will be constructed by GLM from context.

The L vector is checked for estimability in either event. The estimate $Lb$, where $b=(X'X)^{-1}X'Y$, is printed along with its associated standard error, the square root of $L(X'X)^{-1}a^2$ and t-test on the analysis of variance printout page. There is no limit to the number of ESTIMATE statements, but they must come after the MODEL statement.

The ESTIMATE statement is specified as:

```
ESTIMATE 'name' effect list of constants effect list of constants ...
```

In the ESTIMATE statement, 'name' is a name in single quotes consisting of 20 or fewer characters used on the printout to identify the estimate. The effects are the names of effects which appear in the MODEL statement; the keyword INTERCEPT may be used as an effect when an intercept is fitted in the model. The list of constants are the elements of the L vector associated with the preceding effect.

The option below can appear in the ESTIMATE statement. If it appears, a slash (/) should precede it.

```
E
```

The entire L vector will be printed if E is specified.

Examples For the regression model

```
MODEL Y=X1 X2 X3;
```

the associated parameters are $b_0$, $b_1$, $b_2$, and $b_3$ (where $b_0$ represents the intercept). To estimate $3b_1 + 2b_2$, the following L vector is needed:

```
L=[0 3 2 0]
```

The corresponding ESTIMATE statement would be:

```
ESTIMATE '3B1 + 2B2' X1 3 X2 2;
```

To estimate $3b_0 - 2b_3$ the following L vector is needed:

```
L=[1 1 0 -2]
```

The corresponding ESTIMATE statement would be:

```
ESTIMATE '3B0-2B3'
INTERCEPT 1 X1 1 X3 -2;
```

For models involving class variables such as

```
MODEL Y = A B A*B;
```

with the associated parameters:

```
[1 $\alpha_1$ $\alpha_2$ $\alpha_3$ $\beta_1$ $\alpha_{11}$ $\alpha_{12}$ $\alpha_{21}$ $\alpha_{22}$ $\alpha_{31}$ $\alpha_{32}$]
```

To estimate the least squares mean for $a_1$, the following L vector is needed:

```
L=[1 1 0 0 .5 .5 .5 .5 0 0 0 0]
```

and the following ESTIMATE statement could be used.

```
ESTIMATE 'LSM(A1)' INTERCEPT 1 A 1 3 .5 A*B .5 .5;
```

Note in the above statement that only one element of L was specified following the A effect, even though A has three levels. Whenever the list of constants following an effect name is shorter than the effect's number of levels, zeros are used as the remaining constants. In the event that the list of constants is longer than the number of levels for the effect, the extra constants are ignored.

To estimate the A linear effect in the above model, assuming equally spaced levels for A, the following L could be used:

```
L=[0 -1 0 1 0 0 -.5 -.5 0 0 0 .5]
```

The ESTIMATE statement for the above L can be written as

```
ESTIMATE 'A LINEAR' A -1 0 1;
```

In the event the elements of L are not specified for an effect which "contains" a specified effect, then the elements of the specified effect are equitably distributed over the levels of the higher-order effect. The distribution of lower-order coefficients to higher-order effect coefficients follows the same general rules as in the LSMEANS statement, and is similar to that used to construct FPE IV L's. In the previous example,
the -1 associated with \( \beta_3 \) is distributed among the \( \alpha_2 \) parameters in a similar fashion. in the event that an unspecified effect contains several specified effects, distribution of coefficients to the higher-order effect is additive.

Note: Numerous syntactical expressions were examined for the ESTIMATE statement, including many that involved specifying the effect and level information associated with each coefficient. For models involving higher-level effects, the requirement of specifying level information is syntactically intractable. Consequently, the simpler form of the ESTIMATE statement described above was implemented. The syntax of this ESTIMATE statement puts a burden on the user to know a priori what the order of the parameter list associated with each effect is.

3. GLM CONTRAST Statement

The CONTRAST statement provides a mechanism for specifying an L vector or matrix for testing the hypothesis \( L \beta = 0 \). If the hypothesis is testable, the SS(HO:LB=0) is computed and printed in the analysis of variance table. The SS is computed as \( (Lb)'(L(X'X)-L)'^{-1}(Lb) \), where \( b=(X'X)^{-1}X'Y \).

The rules for specifying a contrast are the same as those for specifying an estimate, except that when L has more than one row a comma is used to separate the rows. The CONTRAST statement is specified as:

```
CONTRAST 'name'
    first row of effects and constants,
    second row of effects and constants,
    etc. / options;
```

The effect to be used as a denominator in the F test may be specified. If a RANDOM statement is used, the expected mean square of the contrast will be printed. There is no limit to the number of CONTRAST statements, but they must come after the MODEL statement.

Example For the model

```
MODEL Y=A B;
```

with A at 5 levels and B at 2 levels, the parameter vector is

\[
[ \alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5 \beta_1 \beta_2 ]
\]

To test the hypothesis that the pooled A linear and A quadratic effect is zero, the following L matrix may be used:

\[
L = \begin{bmatrix}
0 & -2 & -1 & 0 & 1 & 2 & 0 & 0 \\
0 & 2 & -1 & -2 & -1 & 2 & 0 & 0 \\
\end{bmatrix}
\]

The corresponding CONTRAST statement would be

```
CONTRAST 'A LINEAR & QUADRATIC'
```

A -2 -1 0 1 2, A 2 -1 -2 -1 2;

If the first level of A were a control level, and you wanted a test of control vs. others, the following statement could be used:

```
CONTRAST 'CONTROL VS OTHERS'
A -1 .25 .25 .25 .25;
```

The L matrix should be of full row rank. However, if it is not, the degrees of freedom associated with the hypotheses will be reduced to row rank(L). The SS computed in this situation will be equivalent to the SS computed using an L matrix with any row deleted—that is a linear combination of previous rows.

Three options are available in the CONTRAST statement.

- **E**
  The entire L vector will be printed if E is specified.
  
  **E=effect**
  Use the E= option to specify an effect in the model to use as an error term. If none is specified, the error MS is used.
  
  **ETYPE=n**
  The ETYPEN option specifies the type (1, 2, 3, or 4) of the E= effect. If E= is specified and ETYPEN is not, the highest type computed in the analysis is used.

4. GLM RANDOM Statement

The RANDOM statement specifies which effects in the model are random. When a RANDOM statement is used, the expected value of each TYPE I, II, III, IV, or contrast MS used in the analysis will be printed. The 79 version of GLM does not make use of the information pertaining to expected mean squares in any way.

The RANDOM statement is specified as:

```
RANDOM list of effects / option;
```

Only one RANDOM statement may be used and it must come after the MODEL statement.

The list of effects in the RANDOM statement should contain one or more of the pure classification effects (main effects, crossed, or nested effects) specified in the MODEL statement. The levels of each effect specified are assumed to be normally and independently distributed with common variance. Levels in different effects are assumed independent.

The option below can appear in the RANDOM statement. If it appears, a slash (/) should precede it.
Q requests a complete printout of all quadratic forms in the fixed effects, which appear in the expected mean squares.

For the model

$$\text{MODEL } Y = A B(A) C A^{*}C;$$

with B(A) declared as random, the expected mean square of each effect will be printed as

$$\text{VAR(ERROR)} + \text{constant} \cdot \text{VAR}(B(A)) + Q(A, C, A^{*}C).$$

If any fixed effects appear in the expected mean square of an effect, the letter Q followed by the list of fixed effects in the expected value is printed. The actual numeric values of the quadratic form (Q matrix) may be printed using the Q option.

Consider the model

$$Y = X_0 \beta_0 + X_1 \beta_1 + X_2 \beta_2 + \ldots + X_k \beta_k + \epsilon$$

where \( \beta_0 \) represents the fixed effects, and \( \beta_1, \beta_2, \ldots, \beta_k \) represent the random effects. For any \( L \) in the row space of \( X = [X_0 \mid X_1 \mid X_2 \mid \ldots \mid X_k] \), then

$$E(\text{SS}_L) = \text{SS}(C_0 \beta_0) + \text{SS}(C_1 \beta_1) + \ldots + \text{SS}(C_k \beta_k) + \text{rank}(L) \sigma^2$$

where \( C \) is of the same dimensions as \( L \) and partitioned as the \( X \) matrix, i.e.

$$C = [C_0 \mid C_1 \mid \ldots \mid C_k].$$

Furthermore, \( C = M L \) where \( M \) is the inverse of the lower triangular Cholesky decomposition matrix of \( L(X'X)'L' \). \( \text{SSQ}(A) \) is defined as tr(\( A'A \)).

A simple proof may be found in Goodnight and Speed (1978).

5. An Example of New GLM Features

To illustrate the new features of GLM and later VARCH the following data will be used.

<table>
<thead>
<tr>
<th>OBS</th>
<th>A</th>
<th>B</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>237</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>254</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>246</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>178</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>179</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
<td>208</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
<td>178</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>2</td>
<td>187</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>2</td>
<td>146</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>2</td>
<td>145</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>2</td>
<td>141</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>1</td>
<td>186</td>
</tr>
</tbody>
</table>

As used by Hemmerle and Hartley, the model for the above data is

$$Y = A B A^{*}B$$

with B and A*B assumed random. New GLM statements are illustrated below.

```
PROC GLM; CLASSES A B;
MODEL Y = A B A*B;
RANDOM B A*B / Q;
CONTRAST 'A LINEAR' A -1 0 1;
CONTRAST 'A QUADRATIC' A -2 1 1;
ESTIMATE 'LSM(A1)' INTERCEPT 1 A 1 B .5 .5 A*B .5 .5 ;
ESTIMATE 'LSM(A2)' INTERCEPT 1 A 0 1 B .5 .5 A*B 0 0 .5 .5 ;
ESTIMATE 'LSM(A3)' INTERCEPT 1 A 0 0 1 B .5 .5 A*B 0 0 0 0 .5 .5 ;
LSMEANS A / S P ;
```

The output from the above statements follows on the next page.
GLM EXAMPLE SHOWING NEW FEATURES
GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: Y

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SUM OF SQUARES</th>
<th>MEAN SQUARE</th>
<th>F VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL</td>
<td>5</td>
<td>23493.50415667</td>
<td>4696.72083333</td>
<td>59.73</td>
</tr>
<tr>
<td>ERROR</td>
<td>10</td>
<td>786.33333333</td>
<td>78.63333333</td>
<td>PR &gt; F</td>
</tr>
<tr>
<td>CORRECTED TOTAL</td>
<td>15</td>
<td>24259.93750000</td>
<td></td>
<td>0.0001</td>
</tr>
</tbody>
</table>

R-SQUARE: C.V. STD DEV Y MEAN
0.967601  4.9419  8.86754382  179.43750000

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>TYPE I SS</th>
<th>F VALUE</th>
<th>PR &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>11736.43750000</td>
<td>74.63</td>
<td>0.0001</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>11448.12564103</td>
<td>145.59</td>
<td>0.0001</td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>299.04102564</td>
<td>1.90</td>
<td>0.1996</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>TYPE IV SS</th>
<th>F VALUE</th>
<th>PR &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>7886.82589744</td>
<td>58.15</td>
<td>0.0001</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>11573.76190476</td>
<td>147.19</td>
<td>0.0001</td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>299.04102564</td>
<td>1.90</td>
<td>0.1996</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CONTRAST</th>
<th>DF</th>
<th>SS</th>
<th>F VALUE</th>
<th>PR &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A LINEAR</td>
<td>1</td>
<td>6657.66666667</td>
<td>84.56</td>
<td>0.0001</td>
</tr>
<tr>
<td>A QUADRATIC</td>
<td>1</td>
<td>1229.76923077</td>
<td>15.64</td>
<td>0.0027</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>ESTIMATE</th>
<th>T FOR HO: PR &gt;</th>
<th>T</th>
<th>STD ERROR OF ESTIMATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSM(A1)</td>
<td>212.08333333</td>
<td>52.40</td>
<td>0.0001</td>
<td>4.04746148</td>
</tr>
<tr>
<td>LSM(A2)</td>
<td>167.58080808</td>
<td>46.27</td>
<td>0.0001</td>
<td>3.62015968</td>
</tr>
<tr>
<td>LSM(A3)</td>
<td>159.41866667</td>
<td>39.39</td>
<td>0.0001</td>
<td>4.04746148</td>
</tr>
</tbody>
</table>

GLM EXAMPLE SHOWING NEW FEATURES
GENERAL LINEAR MODELS PROCEDURE
LEAST SQUARES MEANS

| A | Y | STD ERR | PHOB > | T | PHOB > | T | HO: LSM(E) = LSM(J) =  
|---|---|--------|--------|---|--------|---|I/J | 1 | 2 | 3 |
| 1 | 212.083333 | 4.047461 | 0.0001 | 1 | 0.0001 | 0.0001 |
| 2 | 167.58080808 | 3.62015968 | 0.0001 | 2 | 0.0001 | 0.1574 |
| 3 | 159.41866667 | 4.04746148 | 0.0001 | 3 | 0.0001 | 0.1574 |

NOTE: TO ENSURE OVERALL PROTECTION LEVEL, ONLY PROBABILITIES ASSOCIATED WITH PRE-PLANNED COMPARISONS SHOULD BE USED.
**Example Showing New Features**

**General Linear Models Procedure**

**Dependent Variable: Y**

**Source**

**Type I Expected Mean Square**

- **A**: \( \text{VAR(ERROR)} + 2.725 \text{VAR(A*B)} + 0.1 \text{VAR(B)} + Q(A) \)
- **B**: \( \text{VAR(ERROR)} + 2.63076923 \text{VAR(A*B)} + 7.8 \text{VAR(B)} \)
- **A*B**: \( \text{VAR(ERROR)} + 2.58461538 \text{VAR(A*B)} \)

**Source**

**Type IV Expected Mean Square**

- **A**: \( \text{VAR(ERROR)} + 2.58461538 \text{VAR(A*B)} + Q(A) \)
- **B**: \( \text{VAR(ERROR)} + 2.57142857 \text{VAR(A*B)} + 7.71428571 \text{VAR(B)} \)
- **A*B**: \( \text{VAR(ERROR)} + 2.58461538 \text{VAR(A*B)} \)

**Contrast**

**Expected Mean Square**

- **A Linear**: \( \text{VAR(ERROR)} + 2.4 \text{VAR(A*B)} + Q(A) \)
- **A Quadratic**: \( \text{VAR(ERROR)} + 2.76923077 \text{VAR(A*B)} + Q(A) \)

**Quadratic Forms of Fixed Effects in the Expected Mean Squares**

**Dependent Variable: Y**

**Source: Type I Mean Square for A**

<table>
<thead>
<tr>
<th></th>
<th>A 1</th>
<th>A 2</th>
<th>A 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1</td>
<td>3.43750000</td>
<td>-1.87500000</td>
<td>-1.56250000</td>
</tr>
<tr>
<td>A 2</td>
<td>-1.87500000</td>
<td>3.75000000</td>
<td>-1.87500000</td>
</tr>
<tr>
<td>A 3</td>
<td>-1.56250000</td>
<td>-1.87500000</td>
<td>3.43750000</td>
</tr>
</tbody>
</table>

**Source: Type IV Mean Square for A**

<table>
<thead>
<tr>
<th></th>
<th>A 1</th>
<th>A 2</th>
<th>A 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1</td>
<td>3.32307592</td>
<td>-1.34615385</td>
<td>-1.47692308</td>
</tr>
<tr>
<td>A 2</td>
<td>-1.34615385</td>
<td>3.34615385</td>
<td>-1.47692308</td>
</tr>
<tr>
<td>A 3</td>
<td>-1.47692308</td>
<td>-1.34615385</td>
<td>3.32307592</td>
</tr>
</tbody>
</table>

**Source: Mean Square for Contrast A Linear**

<table>
<thead>
<tr>
<th></th>
<th>A 1</th>
<th>A 2</th>
<th>A 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1</td>
<td>2.40000000</td>
<td>0.00000000</td>
<td>-2.40000000</td>
</tr>
<tr>
<td>A 2</td>
<td>0.00000000</td>
<td>0.00000000</td>
<td>0.00000000</td>
</tr>
<tr>
<td>A 3</td>
<td>-2.40000000</td>
<td>0.00000000</td>
<td>2.40000000</td>
</tr>
</tbody>
</table>

**Source: Mean Square for Contrast A Quadratic**

<table>
<thead>
<tr>
<th></th>
<th>A 1</th>
<th>A 2</th>
<th>A 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1</td>
<td>0.92307592</td>
<td>-1.84615385</td>
<td>0.92307592</td>
</tr>
<tr>
<td>A 2</td>
<td>-1.84615385</td>
<td>3.69230769</td>
<td>-1.84615385</td>
</tr>
<tr>
<td>A 3</td>
<td>0.92307592</td>
<td>-1.84615385</td>
<td>0.92307592</td>
</tr>
</tbody>
</table>
6. New VARCOMP Features

The VARCOMP Procedure for estimating variance components has been significantly improved for the 79 version. Some of the new features are:

- Three methods are now available for estimation.
  1. **TYPE I** equates Type I MS's to their expected values and solves
  2. **MIVQUEO** ... for designs with thousands of levels
  3. **ML** ... Maximum Likelihood

- Level determination and lookup time has been significantly reduced through use of a balanced binary table lookup and insertion technique.

- The ' | ' notation may be used as in GLM, i.e., A|B|C = A*B A*C B*C A*B*C

- The number of fixed effects can be specified with the FIXED=K parameter.

- **TYPE I** and **MIVQUEO** will handle any size problem.

- **ML** requires X'X to be in core.

The following example illustrates the new FIXED option in VARCOMP using the **TYPE I** method which was the only available method in the 76 version. The data is the same as used previously.

**PROC VARCOMP METHOD=TYPEI ; CLASSES A B ;**

**MODEL Y = A B A*B / FIXED=1 ;**

The output from the above statements is shown below.

---

### VARCOMP EXAMPLE SHOWING NEW FEATURES

**VARIANCE COMPONENT ESTIMATION PROCEDURE**

**DEPENDENT VARIABLE:** Y

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>TYPE I SS</th>
<th>TYPE I MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>11736.43750000</td>
<td>5868.21875000</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>11448.12564103</td>
<td>11448.12564103</td>
</tr>
<tr>
<td>A*B</td>
<td>2</td>
<td>299.04162564</td>
<td>149.52851282</td>
</tr>
<tr>
<td>ERROR</td>
<td>10</td>
<td>786.33333333</td>
<td>78.63333333</td>
</tr>
<tr>
<td><strong>CORRECTED TOTAL</strong></td>
<td>15</td>
<td><strong>24259.93750000</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>EXPECTED MEAN SQUARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>VAR(ERROR) + 2.725 VAR(A*B) + 0.1 VAR(B) + Q(A)</td>
</tr>
<tr>
<td>B</td>
<td>VAR(ERROR) + 2.63076923 VAR(A*B) + 7.8 VAR(B)</td>
</tr>
<tr>
<td>A*B</td>
<td>VAR(ERROR) + 2.58461538 VAR(A*B)</td>
</tr>
<tr>
<td>ERROR</td>
<td>VAR(ERROR)</td>
</tr>
</tbody>
</table>

**VARIANCE COMPONENT ESTIMATE**

<table>
<thead>
<tr>
<th>COMPONT</th>
<th>ESTIMATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR(b)</td>
<td>1448.37663150</td>
</tr>
<tr>
<td>VAR(A*B)</td>
<td>27.42658730</td>
</tr>
<tr>
<td>VAR(ERROR)</td>
<td>78.63333333</td>
</tr>
</tbody>
</table>
7. The MIVQUEO Method of VARCOMP

MIVQUEO is designed to handle efficiently very large designs. It is Rao's MIVQUE with zero priors. The method was recently advocated by Hartley et al. (1978) for the following reasons.

- Computationally efficient
- Unbiased
- Locally best (at zero)
- Admissible
- Asymptotically consistent
- "Best unbiased" if design is balanced

Goodnight (1978) describes the computing techniques and theoretical background for the method.

The underlying model is

\[ Y = X_0 \beta_0 + \sum_{i=1}^{k} X_i \beta_i + \epsilon \]

where \( \beta_0 \) is fixed, all other effects (NID) random.

The computation of MIVQUEO estimates is described in the following four steps:

**Computing Formulae**

**Step 1:** Form the symmetric matrix:

\[
\begin{bmatrix}
X'X & X'Y \\
Y'X & Y'Y
\end{bmatrix}
\]

**Step 2:** Apply any of the elimination methods... Gauss, Gauss-Jordan, Doolittle, or Cholesky — to the above matrix by pivoting on each diagonal of \( X'X \). This reduces the elements below and to the right of \( X'X \) to the following:

\[
\begin{bmatrix}
X_0'R_0X_0 & X_0'R_0Y \\
X_1'R_0X_1 & X_1'R_0Y \\
\vdots & \vdots \\
X_k'R_0X_k & X_k'R_0Y
\end{bmatrix}
\]

where \( R_0 = I - X_0(X_0'X_0)^{-1}X'_0 \)

**Step 3:** Form the symmetric SSQ matrix from the above:

\[
\begin{bmatrix}
SSQ(X_0'R_0X_0) & \ldots & SSQ(X_0'R_0X_k) & Tr(X_0'R_0X_0) & SSQ(X_0'R_0Y) \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
SSQ(X_k'R_0X_k) & Tr(X_k'R_0X_k) & SSQ(X_k'R_0Y)
\end{bmatrix}
\]

**Step 4:** Reduce the left hand side to an identity using pivoting operations, and providing no linear dependencies are found, the MIVQUEO estimates are obtained.

To demonstrate the computing speed for large designs, several two-way crossed models of the form:

\[ y = \mu + \alpha_i + \beta_j + (\alpha \beta)_{ij} + \epsilon \]

were run. Here \( \mu \) is fixed, and \( \alpha, \beta, (\alpha \beta) \), and \( \epsilon \) are random. This model was chosen because of the ease in which large numbers of levels can be generated for \( \alpha \) and \( \beta \), a critical factor in computing speed. For each set of data generated, there was only one observation per cell, except for those cells involving \( \alpha \), which contained two observations. The problem size below, represents the number of levels of \( \alpha \) and \( \beta \). For a 10x10 there are \( 10 + 10 + 100 \) random rows. Timings for several different region sizes are shown.

**Computing Speeds on an IBM 370/168 with Virtual Storage**

<table>
<thead>
<tr>
<th>Problem size</th>
<th>Number of rows of obs.</th>
<th>200K region</th>
<th>300K region</th>
<th>400K region</th>
<th>500K region</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x10</td>
<td>120</td>
<td>.44</td>
<td>.93</td>
<td>.76</td>
<td>.73</td>
</tr>
<tr>
<td>20x20</td>
<td>440</td>
<td>1.94</td>
<td>3.76</td>
<td>3.12</td>
<td>3.01</td>
</tr>
<tr>
<td>30x30</td>
<td>960</td>
<td>11.01</td>
<td>7.04</td>
<td>6.50</td>
<td>5.96</td>
</tr>
<tr>
<td>40x40</td>
<td>1680</td>
<td>28.75</td>
<td>23.09</td>
<td>20.18</td>
<td>18.95</td>
</tr>
<tr>
<td>50x50</td>
<td>2600</td>
<td>65.60</td>
<td>53.96</td>
<td>51.15</td>
<td>49.30</td>
</tr>
</tbody>
</table>

The largest problem tested to date involved 23 random effects with 30,491 random levels. It ran in 221 CPU seconds in a region of 900K.
The MIVQUEO method is illustrated by the following example.

PROC VARCOMP METHOD=MIVQUEO;
CLASSES A B;
MODEL Y = A B A*B / FIXED=1;

### VARCOMP Example Showing New Features

**mIVQUEO(0) Variance Component Estimation Procedure**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>B</th>
<th>A*B</th>
<th>ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>60.8400000000</td>
<td>20.5200000000</td>
<td>7.8800000000</td>
</tr>
<tr>
<td>A*B</td>
<td>20.5200000000</td>
<td>20.5200000000</td>
<td>7.8800000000</td>
</tr>
<tr>
<td>ERROR</td>
<td>7.8800000000</td>
<td>7.8800000000</td>
<td>13.6800000000</td>
</tr>
</tbody>
</table>

**SSQ Matrix**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>ESTIMATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>&amp;</td>
<td>Y</td>
</tr>
<tr>
<td>A*B</td>
<td>69295.30000000</td>
</tr>
<tr>
<td>ERROR</td>
<td>12533.50000000</td>
</tr>
</tbody>
</table>

### 8. Maximum Likelihood Method of VARCOMP

Hemmerle and Hartley (1973) developed the basic formulae used in VARCOMP for Maximum Likelihood estimation. The w-transformation which they describe in the above paper has not only made ML estimation computationally tractable but has also made feasible the computation of REML, MINQUE, MIVQUE and other techniques.

For the model

\[ Y = \sum_{i=1}^{k} X_i \beta_i + \epsilon \]

with \( \beta_0 \) fixed and all other effects (NID) random, the variance-covariance matrix of \( Y \) is \( \sigma^2 H \)

where

\[ H = I + \sum_{i=1}^{k} Y_i X_i^T X_i \]

Letting \( X = [X_0 X_1 ... X_k] \)

the w-transformation maps the matrix

\[ W_0 = \begin{bmatrix} X^T X & X^T \epsilon \\ \epsilon \epsilon^T & \epsilon^T \epsilon \end{bmatrix} \]

into the matrix

\[ W = \begin{bmatrix} X' H^{-1} X & X' H^{-1} \epsilon \\ \epsilon' H^{-1} X & \epsilon' H^{-1} \epsilon \end{bmatrix} \]

The simplified w-transformation developed by Goodnight and Hemmerle (1978) performs the above mapping:

- In place
- Given only \( W_0 \), the \( \epsilon \) vector, and a work vector sufficient to hold one row of \( W_0 \)
- Using only about 12 Fortran or PL/1 statements (includes computation of \( |H|^{-1} \))

Various functions of the elements of \( W \) are employed in each ML step to arrive at new parameter estimates, and the w-transformation is then reapplied.

The following example illustrates the ML method using the previous data set.

PROC VARCOMP METHOD=ML; CLASSES A B;
MODEL Y = A B A*B / FIXED=1;
### VARCOMP Example: Showing New Features

**Maximum Likelihood Variance Component Estimation Procedure**

**Dependent Variable:** Y

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Objective</th>
<th>$\text{VAR}(B)$</th>
<th>$\text{VAR}(A*B)$</th>
<th>$\text{VAR(}\text{ERROR})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>78.38503712</td>
<td>1031.49069751</td>
<td>0.00000074</td>
<td>74.39097179</td>
</tr>
<tr>
<td>1</td>
<td>78.30878578</td>
<td>595.97627809</td>
<td>0.00000001</td>
<td>79.93075397</td>
</tr>
<tr>
<td>2</td>
<td>78.26758062</td>
<td>681.78485449</td>
<td>0.00000000</td>
<td>78.20570320</td>
</tr>
<tr>
<td>3</td>
<td>78.26366863</td>
<td>715.55121529</td>
<td>0.00000000</td>
<td>77.65343754</td>
</tr>
<tr>
<td>4</td>
<td>78.26354981</td>
<td>722.53792585</td>
<td>0.00000000</td>
<td>77.54752507</td>
</tr>
<tr>
<td>5</td>
<td>78.26354716</td>
<td>723.52197847</td>
<td>0.00000000</td>
<td>77.53266174</td>
</tr>
<tr>
<td>6</td>
<td>78.26354712</td>
<td>723.64778973</td>
<td>0.00000000</td>
<td>77.53076595</td>
</tr>
<tr>
<td>7</td>
<td>78.26354712</td>
<td>723.66354481</td>
<td>0.00000000</td>
<td>77.53052723</td>
</tr>
</tbody>
</table>

**Convergence Criterion Met**

Note: Initial estimates of $\gamma_i$ come from MINQUEO estimation. Job time for the above was .30 seconds.

### 9. Summary

The ESTIMATE and CONTRAST statements of GLM have helped to round out the first phase development goals for GLM—that of being a complete fixed effects linear models procedure. The RANDOM statement is the first step into the mixed model phase of development.

The goals of the mixed model phase of development are to compute $F$ tests and standard errors completely automatic, as a function of the error structure defined by the RANDOM statement.

Current GLM research on mixed models centers on the following points:

1. **Denominator construction for approximate "$F" test**
   - Obviously it is possible to construct a denominator MS which has the same expected value under the null hypothesis as does the numerator.
   - Which type or combination of types (I, II, III, IV) do we use?
   - Should we just estimate the variance components and compute the needed linear combination?
   - If we estimate components what method do we use?

   The theoretical problems associated with the above in the general unbalanced mixed model are:
   - The numerator and denominator will invariably be correlated.
   - The denominator will invariably not be distributed as chi-square.
   - The numerator may often not be distributed as a non-central chi-square.
   - Do best tests even exist?

2. **Constructing new $L$ types for numerator and denominator may have some promise.**
   - It can be shown that for any $L$ in the row space of $X$ that $\text{COV}(SS_L, MS\text{error}) = 0$.
   - What conditions on $L_1$ and $L_2$ are required so that $\text{COV}(SS_{L_1}, SS_{L_2}) = 0$?
   - It can be shown that a necessary condition for the above is that $L_1 (X'X)^{-1} L_2 = 0$.
   - Furthermore the 2 conditions:
     - $L_1 L_2' = 0$
     - $L_1 (X'X)^{-1} L_2 = 0$
   - are sufficient for $\text{COV}(SS_{L_1}, SS_{L_2}) = 0$.
   - I am not certain whether $L_1 L_2' = 0$ is a necessary condition, but if it is then meeting the two necessary conditions may be difficult except in balanced designs.

3. **The only remaining solution may be to abandon the classical type of analysis of variance framework and proceed to maximum likelihood.**
   - Research in this area is proceeding with the VARCOMP procedure.
REFERENCES


