Chapter 32
Three Way Mixed Model

C. R. Henderson

1984 - Guelph

Some of the principles of preceding chapters are illustrated in this chapter using an unbalanced 3 way mixed model. The method used here is one of several alternatives that appeals to me at this time. However, I would make no claims that it is "best".

1 The Example

Suppose we have a 3 way classification with factors denoted by A, B, C. The levels of A are random and those of B and C are fixed. Accordingly a traditional mixed model would contain factors and interactions as follows, a, b, c, ab, ac, bc, abc with b, c, and bc fixed, and the others random. The subclass numbers are as follows.

	BC subclasses								
A	11	12	13	21	22	23	31	32	33
1	5	2	3	6	0	3	2	5	0
2	1	2	4	0	5	2	3	6	0
3	0	4	8	2	3	5	7	0	0

The associated ABC subclass means are

$$\begin{pmatrix} 3 & 5 & 2 & 4 & - & 8 & 9 & 2 & - \\ 5 & 6 & 7 & - & 8 & 5 & 2 & 6 & - \\ - & 9 & 8 & 4 & 3 & 7 & 5 & - & - \end{pmatrix}.$$

Because there are no observations on bc_{33}, estimates and tests of b, c, and $b \times c$ that mimic the filled subclass case cannot be accomplished using unbiased estimators. Accordingly, we might use some prior on squares and products of bc_{jk} and obtain biased estimators. Let us assume the following prior values, $\sigma_e^2/\sigma_a^2 = 2$, $\sigma_e^2/\sigma_{ab}^2 = 3$, $\sigma_e^2/\sigma_{ac}^2 = 4$, $\sigma_e^2/$pseudo $\sigma_{bc}^2 = 6$, $\sigma_e^2/\sigma_{abc}^2 = 5$.

2 Estimation And Prediction

The OLS equations that include missing observations have 63 unknowns as follows

a	$1-3$	**ac**	$19-27$
b	$4-6$	**bc**	$28-36$
c	$7-9$	**abc**	$37-63$
ab	$10-18$		

$\overline{\mathbf{W}}$ is a 20 $\times$ 63 matrix with 1's in the following columns of the 20 rows. The other elements are 0.

Levels of a	b	c	Cols. with 1
1	1	1	1,4,7,10,19,28,37
1	1	2	1,4,8,10,20,29,38
1	1	3	1,4,9,10,21,30,39
1	2	1	1,5,7,11,19,31,40
1	2	3	1,5,9,11,21,33,42
1	3	1	1,6,7,12,19,34,43
1	3	2	1,6,8,12,20,35,44
2	1	1	2,4,7,13,22,28,46
2	1	2	2,4,8,13,23,29,47
2	1	3	2,4,9,13,24,30,48
2	2	2	2,5,8,14,23,32,50
2	2	3	2,5,9,14,24,33,51
2	3	1	2,6,7,15,22,34,52
2	3	2	2,6,8,15,23,35,53
3	1	2	3,4,8,16,26,29,56
3	1	3	3,4,9,16,27,30,57
3	2	1	3,5,7,17,25,31,58
3	2	2	3,5,8,17,26,32,59
3	2	3	3,5,9,17,27,33,60
3	3	1	3,6,7,18,25,34,61

Let $\mathbf{N}$ be a 20 $\times$ 20 diagonal matrix with filled subclass numbers in the diagonal, that is $\mathbf{N}$ = diag(5,2,. . . ,5,7). Then the OLS coefficient matris is $\overline{\mathbf{W}}'\mathbf{N}\overline{\mathbf{W}}$, and the right hand side vector is $\mathbf{W}'\mathbf{N}\overline{\mathbf{y}}$, where $\overline{\mathbf{y}} = (3\ \ 5\ \ldots\ 7\ \ 5)'$. The right hand side vector is (107, 137, 187, 176, 150, 105, 111, 153, 167, 31, 48, 28, 45, 50, 42, 100, 52, 35, 57, 20, 30, 11, 88, 38, 43, 45, 99, 20, 58, 98, 32, 49, 69, 59, 46, 0, 15, 10, 6, 24, 0, 24, 18, 10, 0, 5, 12, 28, 0, 40, 10, 6, 36, 0, 0, 36, 64, 8, 9, 35, 35, 0, 0).

Now we add the following diagonal matrix to the coefficient matrix, (2, 2, 2, 0, 0, 0, 0, 0, 0, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 6, 6, 6, 6, 6, 6, 6, 6, 6, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5). A resulting mixed model solution is

$$\begin{aligned}
\mathbf{a} &= (-.54801,\ .10555,\ .44246)'. \\
\mathbf{b} &= (6.45206,\ 6.13224,\ 5.77884)'. \\
\mathbf{c} &= (-1.64229,\ -.67094,\ 0)'. \\
\mathbf{ab} &= (-1.21520,\ .46354,\ .38632,\ .14669,\ .29571,\ -.37204, \\
&\quad 1.06850,\ -.75924,\ -.01428)'. \\
\mathbf{ac} &= (.60807,\ -.70385,\ -.17822,\ -.63358,\ .85039,\ -.16403, \\
&\quad .02552,\ -.14653,\ .34225)'. \\
\mathbf{bc} &= (-.12431,\ .47539,\ -.35108,\ -.30013,\ -.05095,\ .35108, \\
&\quad .42444,\ -.42444,\ 0)'. \\
\mathbf{abc} &= (-.26516,\ .34587,\ -.80984,\ -.38914,\ 0,\ .66726,\ 1.14075, \\
&\quad -.90896,\ 0,\ .11598,\ -.38832,\ .36036,\ 0,\ .66901,\ -.49158, \\
&\quad -.62285,\ .39963,\ 0,\ 0,\ .61292,\ .02819,\ .02898,\ -.73014, \\
&\quad .24561,\ -.00857,\ 0,\ 0)'.
\end{aligned}$$

From these results the biased prediction of subclass means are in (32.1).

	B_1			B_2			B_3		
A	C_1	C_2	C_3	C_1	C_2	C_3	C_1	C_2	C_3
1	3.265	4.135	3.350	4.324	4.622	6.888	6.148	2.909	5.439
2	4.420	6.971	6.550	3.957	7.331	6.229	3.038	5.667	5.348
3	6.222	8.234	7.982	3.928	4.217	6.754	5.006	4.965	6.549

(1)

Note that these are different from the $\overline{y}_{ijk}$ for filled subclasses, the latter being BLUE. Also subclass means are predicted for those cases with no observations.

3 Tests Of Hypotheses

Suppose we wish to test the following hypotheses regarding **b**, **c**, and **bc**. Let

$$\mu_{jk} = \mu + b_j + c_k + bc_{jk}.$$

We test $\overline{\mu}_{j.}$ are equal, $\overline{\mu}_{.k}$ are equal, and that all $\overline{\mu}_{jk}$ - $\overline{\mu}_{j.}$ - $\overline{\mu}_{.k}$ + $\overline{\mu}_{..}$ are equal. Of course these functions are not estimable if any jk subclass is missing as is true in our example. Consequently we must resort to biased estimation and accompanying approximate tests based on estimated MSE rather than sampling variances. We assume that our priors are the correct values and proceed for the first test.

$$\mathbf{K}'\boldsymbol{\beta} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 \\ \end{pmatrix.$$
$$\begin{matrix} 0 & -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \end{matrix}\Bigg)\boldsymbol{\beta},$$

where $\boldsymbol{\beta}$ is the vector of μ_{ijk} ordered k in j in i. From (32.1) the estimate of these functions is (6.05893, 3.18058). To find the mean squared errors of this function we first compute the mean squared errors of the $\overline{\mu}_{ijk}$. This is $\overline{\mathbf{W}}\mathbf{C}\overline{\mathbf{W}}' \equiv \mathbf{P}$, where $\overline{\mathbf{W}}$ is the matrix relating ijk subclass means to the 63 elements of our model. $\mathbf{C}$ is a g-inverse of the mixed model coefficient matrix. Then the mean squared error of $\mathbf{K}'\boldsymbol{\beta}$ is

$$\mathbf{K}'\mathbf{P}\mathbf{K} = \begin{pmatrix} 17.49718 & 13.13739 \\ & 16.92104 \end{pmatrix}.$$

Then

$$\boldsymbol{\beta}^{o\prime}\mathbf{K}'(\mathbf{K}'\mathbf{P}\mathbf{K})^{-1}\mathbf{K}'\boldsymbol{\beta}^o = 2.364,$$

and this is distributed approximately as χ^2 with 2 d.f. under the null hypothesis.

To test C we use

$$\mathbf{K}'\boldsymbol{\beta} = \begin{pmatrix} 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \end{pmatrix.$$
$$\begin{pmatrix} -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \end{pmatrix} \boldsymbol{\beta}.$$

$$\mathbf{K}'\boldsymbol{\beta}^o = (-14.78060,\ -6.03849)',$$

with

$$\text{MSE} = \begin{pmatrix} 17.25559 & 10.00658 \\ & 14.13424 \end{pmatrix}.$$

This gives the test criterion = 13.431, distributed approximately as χ^2 with 2 d.f. under the null hypothesis.

For $B \times C$ interaction we use

$$\mathbf{K}'\boldsymbol{\beta} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix.$$
$$\begin{pmatrix} 0 & -1 & 0 & 1 & 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\ -1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 \\ -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 \end{pmatrix} \boldsymbol{\beta}.$$

This gives

$$\mathbf{K}'\boldsymbol{\beta}^o = (-.83026,\ 5.25381,\ -4.51772,\ .09417)',$$

with

$$\text{MSE} = \begin{pmatrix} 6.37074 & 4.31788 & 4.56453 & 3.64685 \\ & 6.09614 & 3.77847 & 4.70751 \\ & & 6.32592 & 4.23108 \\ & & & 6.31457 \end{pmatrix}.$$

The test criterion is 21.044 distributed approximately as χ^2 with 4 d.f. under the null hypothesis.

Note that in these examples of hypothesis testing the priors used were quite arbitrary. The tests are of little value unless one has good prior estimates. This of course is true for any unbalanced mixed model design.

4 REML Estimation By EM Method

We next illustrate one round of estimation of variances by the EM algorithm. We treat σ^2_{bc} as a variance. The first round of estimation is obtained from the mixed model solution of Section 32.2. For σ^2_e we compute

$$[\mathbf{y}'\mathbf{y} - \text{soln. vector (r.h.s. vector)}]/[\text{n} - \text{rank } (\mathbf{X})].$$

$$\mathbf{y}'\mathbf{y} = 2802.$$

$$\text{Red} = 2674.47.$$

$$\hat{\sigma}^2_e = (2802 - 2674.47)/(78 - 5) = 1.747.$$

$$\hat{\sigma}^2_a = \left(\hat{\mathbf{a}}'\hat{\mathbf{a}} + tr\ 1.747 \begin{pmatrix} .28568 & .10673 & .10759 \\ & .28645 & .10683 \\ & & .28558 \end{pmatrix}\right) /3 = .669$$

$$\hat{\sigma}^2_{ab} = \left(\hat{\mathbf{a}}\hat{\mathbf{b}}'\hat{\mathbf{a}}\hat{\mathbf{b}} + tr\ 1.747 \begin{pmatrix} .2346 & & \cdots \\ & \ddots & \\ & & .26826 \end{pmatrix}\right) /9 = .847.$$

$$\hat{\sigma}^2_{ac} = \left(\hat{\mathbf{a}}\hat{\mathbf{c}}'\hat{\mathbf{a}}\hat{\mathbf{c}} + tr\ 1.747 \begin{pmatrix} .19027 & & \cdots \\ \vdots & \ddots & \\ & & .18846 \end{pmatrix}\right) /9 = .580.$$

$$\hat{\sigma}^2_{bc} = \left(\hat{\mathbf{b}}\hat{\mathbf{c}}'\hat{\mathbf{b}}\hat{\mathbf{c}} + tr\ 1.747 \begin{pmatrix} .14138 & & \cdots \\ \vdots & \ddots & \\ & & .16607 \end{pmatrix}\right) /9 = .357.$$

$$\hat{\sigma}^2_{abc} = \left(\hat{\mathbf{a}}\hat{\mathbf{b}}\hat{\mathbf{c}}'\hat{\mathbf{a}}\hat{\mathbf{b}}\hat{\mathbf{c}} + tr\ 1.747 \begin{pmatrix} .16505 & & \cdots \\ \vdots & \ddots & \\ & & .20000 \end{pmatrix}\right) /127 = .534.$$

The solution for four rounds follows.

	1	2	3	4
σ_e^2	1.747	1.470	1.185	.915
σ_a^2	.169	.468	.330	.231
σ_{ab}^2	.847	.999	1.090	1.102
σ_{ac}^2	.580	.632	.638	.587
σ_{bc}^2	.357	.370	.362	.327
σ_{abc}^2	.534	.743	1.062	1.506

It appears that $\hat{\sigma}_e^2$ and $\hat{\sigma}_{abc}^2$ may be highly confounded, and convergence will be slow. Note that $\hat{\sigma}_e^2 + \hat{\sigma}_{abc}^2$ does not change much.