Chapter 17
The Two Way Classification

C. R. Henderson

1984 - Guelph

This chapter is concerned with a linear model in which

$$y_{ijk} = \mu + a_i + b_j + \gamma_{ij} + e_{ijk}. \tag{1}$$

For this to be a model we need to specify whether **a** is fixed or random, **b** is fixed or random, and accordingly whether $\boldsymbol{\gamma}$ is fixed or random. In the case of random subvectors we need to specify the variance-covariance matrix, and that is determined in part by whether the vector sampled is finite or infinite.

1 The Two Way Fixed Model

We shall be concerned first with a model in which **a** and **b** are both fixed, and as a consequence so is $\boldsymbol{\gamma}$. For convenience let

$$\mu_{ij} = \mu + a_i + b_j + \gamma_{ij}. \tag{2}$$

Then it is easy to prove that the only estimable linear functions are linear functions of μ_{ij} that are associated with filled subclasses ($n_{ij} > 0$). Further notations and definitions are:

$$\text{Row mean} = \bar{\mu}_{i.}. \tag{3}$$

Its estimate is sometimes called a least squares mean, but I agree with Searle *et al*. (1980) that this is not a desirable name.

$$\begin{aligned} \text{Column mean} &= \bar{\mu}_{.j}. && (4) \\ \text{Row effect} &= \bar{\mu}_{i.} - \bar{\mu}_{..}. && (5) \\ \text{Column effect} &= \bar{\mu}_{.j} - \bar{\mu}_{..}. && (6) \\ \text{General mean} &= \bar{\mu}_{..}. && (7) \\ \text{Interaction effect} &= \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..}. && (8) \end{aligned}$$

From the fact that only μ_{ij} for filled subclasses are estimable, missing subclasses result in the parameters of (17.3) ... (17.8) being non-estimable.

$\bar{\mu}_{i'.}$ is not estimable if any $n_{i'j} = 0$.
$\bar{\mu}_{.j'}$ is not estimable if any $n_{ij'} = 0$.
$\bar{\mu}_{..}$ is not estimable if one or more $n_{ij} = 0$.

All row effects, columns effects, and interaction effects are non-estimable if one or more $n_{ij} = 0$. Due to these non-estimability considerations, mimicking of either the balanced or the filled subclass estimation and tests of hypotheses wanted by many experimenters present obvious difficulties. We shall present biased methods that are frequently used and a newer method with smaller mean squared error of estimation given certain assumptions.

2 BLUE For The Filled Subclass Case

Assuming that $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$, it is easy to prove that $\hat{\mu}_{ij} = \bar{y}_{ij.}$. Then it follows that BLUE of the i^{th} row mean in the filled subclass case is

$$\frac{1}{c} \sum_{j=1}^{c} \bar{y}_{ij.}. \tag{9}$$

BLUE of j^{th} column mean is

$$\frac{1}{r} \sum_{i=1}^{r} \bar{y}_{ij.}. \tag{10}$$

$$\begin{aligned} r &= \text{number of rows, and} \\ c &= \text{number of columns.} \end{aligned}$$

BLUE of i^{th} row effect is

$$\frac{1}{c} \sum_j \bar{y}_{ij.} - \frac{1}{rc} \sum_i \sum_j \bar{y}_{ij.}. \tag{11}$$

Thus BLUE of any of (17.3), ... , (17.8) is that same function of $\hat{\mu}_{ij}$, where $\hat{\mu}_{ij} = \bar{y}_{ij.}$.

The variances of any of these functions are simple to compute. Any of them can be expressed as $\sum_i \sum_j k_{ij}\ \mu_{ij}$ with BLUE =

$$\sum_i \sum_j k_{ij}\bar{y}_{ij.}. \tag{12}$$

The variance of this is

$$\sigma_e^2 \sum_i \sum_j k_{ij}^2/n_{ij}. \tag{13}$$

The covariance between BLUE's of linear functions,

$$\sum_i \sum_j k_{ij}\bar{y}_{ij.} \ \text{ and } \ \sum_i \sum_j t_{ij}\bar{y}_{ij.'}$$

is

$$\sigma_e^2 \sum_i \sum_j k_{ij}t_{ij}/n_{ij}. \tag{14}$$

The numbers required for tests of hypotheses are (17.13) and (17.14) and the associated BLUE's. Consider a standard ANOVA, that is, mean squares for rows, columns, $R \times C$. The $R \times C$ sum of squares with $(r-1)(c-1)$ d.f. can be computed by

$$\sum_i \sum_j \frac{y_{ij.}^2}{n_{ij}} \ - \ \text{Reduction under model with no interaction.} \tag{15}$$

The last term of (17.15) can be obtained by a solution to

$$\begin{pmatrix} \mathbf{D}_i & \mathbf{N}_{ij} \\ \mathbf{N}_{ij} & \mathbf{D}_j \end{pmatrix} \begin{pmatrix} \mathbf{a}^o \\ \mathbf{b}^o \end{pmatrix} = \begin{pmatrix} \mathbf{y}_i \\ \mathbf{y}_j \end{pmatrix}. \tag{16}$$

$$\begin{aligned} \mathbf{D}_i &= \text{diag}\ (n_{1.},\ n_{2.}, \ldots, n_{r.}). \\ \mathbf{D}_j &= \text{diag}\ (n_{.1},\ n_{.2}, \ldots, n_{.c}). \\ \mathbf{N}_{ij} &= \text{matrix of all } n_{ij}. \\ \mathbf{y}_i^{'} &= (y_{1..}, \ldots, y_{r..}). \\ \mathbf{y}_j^{'} &= (y_{.1.}, \ldots, y_{.c.}). \end{aligned}$$

Then the reduction is

$$(\mathbf{a}^o)'\ \mathbf{y}_i + (\mathbf{b}^o)'\mathbf{y}_j. \tag{17}$$

Sums of squares for rows and columns can be computed conveniently by the method of weighted squares of means, due to Yates (1934). For rows compute

$$\alpha_i = \frac{1}{c} \sum\nolimits_j \bar{y}_{ij.}\ (i = 1, \ldots, r),\ \text{and} \tag{18}$$

$$k_i^{-1} = \frac{1}{c^2} \sum\nolimits_j \frac{1}{n_{ij}}.$$

Then the row S.S. with $r - 1$ d.f. is

$$\sum\nolimits_i k_i\ \alpha_i^2\ -\ (\sum\nolimits_i k_i \alpha_i)^2 / \sum\nolimits_i k_i. \tag{19}$$

The column S.S. with $c - 1$ d.f. is computed in a similar manner. The "error" mean square for tests of these mean squares is

$$(\mathbf{y}'\mathbf{y} - \sum\nolimits_i \sum\nolimits_j y_{ij.}^2 / n_{ij}) / (n_{..} - rc). \tag{20}$$

An obvious limitation of the weighted squares of means for testing rows is that the test refers to equal weighting of subclasses across columns. This may not be what is desired by the experimenter.

An illustration of a filled subclass 2 way fixed model is a breed by treatment design with the following n_{ij} and $y_{ij.}$.

	Treatments					
	n_{ij}			$y_{ij.}$		
Breeds	1	2	3	1	2	3
1	5	2	1	68	29	19
2	4	2	2	55	30	36
3	5	1	4	61	13	61
4	4	5	4	47	65	75

$$\sum_i \sum_j y_{ij.}^2 / n_{ij} = 8207.5.$$

Let us test the hypothesis that interaction is negligible. The reduction under a model with no interaction can be obtained from a solution to equation (17.21).

$$\begin{pmatrix} 8 & 0 & 0 & 0 & 5 & 2 \\ & 8 & 0 & 0 & 4 & 2 \\ & & 10 & 0 & 5 & 1 \\ & & & 13 & 4 & 5 \\ & & & & 18 & 0 \\ & & & & & 10 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 116 \\ 121 \\ 135 \\ 187 \\ 231 \\ 137 \end{pmatrix}. \tag{21}$$

The solution is (18.5742, 18.5893, 16.3495, 17.4624, -4.8792, -4.0988)$'$. The reduction is 8187.933. Then $R \times C$ S.S. = 8207.5 - 8187.923 = 19.567. S. S. for rows can be formulated as a test of the hypothesis

$$\mathbf{K}'\boldsymbol{\mu} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu_{11} \\ \vdots \\ \mu_{43} \end{pmatrix} = \mathbf{0}.$$

The $\hat{\mu}_{ij}$ are (13.6, 14.5, 19.0, 13.75, 15.0, 18.0, 12.2, 13.0, 15.25, 11.75, 13.0, 18.75).

$$\begin{aligned} \mathbf{K}'\hat{\boldsymbol{\mu}} &= (3.6 \; 3.25 \; -3.05)'. \\ Var(\mathbf{K}'\hat{\boldsymbol{\mu}}) &= \mathbf{K}' \, [\text{diag } (5 \; 2 \; \ldots 4)]^{-1} \, \mathbf{K}\sigma_e^2 \\ &= \begin{pmatrix} 2.4 & .7 & .7 \\ & 1.95 & .7 \\ & & 2.15 \end{pmatrix} \sigma_e^2. \end{aligned}$$

$$\sigma_e^{-2} \, (\mathbf{K}'\hat{\boldsymbol{\mu}})' \, [Var(\mathbf{K}'\hat{\boldsymbol{\mu}})]^{-1} \; \mathbf{K}'\hat{\boldsymbol{\mu}} = 20.54 \; = \text{SS for rows.}$$

SS for cols. is a test of

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

$$\mathbf{K}'\hat{\boldsymbol{\mu}} = \begin{pmatrix} -19.7 \\ -15.5 \end{pmatrix}.$$

$$Var(\mathbf{K}'\hat{\boldsymbol{\mu}}) = \begin{pmatrix} 2.9 & 2.0 \\ & 4.2 \end{pmatrix} \sigma_e^2.$$

$$\sigma_e^{-2} \, (\mathbf{K}'\hat{\boldsymbol{\mu}})' \, [Var(\mathbf{K}'\hat{\boldsymbol{\mu}})]^{-1} \; \mathbf{K}'\hat{\boldsymbol{\mu}} = 135.12 = \text{SS for Cols.}$$

Next we illustrate weighted squares of means to obtain these same results. Sums of squares for rows uses the values below

	α_i	k_i
1	15.7	5.29412
2	15.5833	7.2
3	13.4833	6.20690
4	14.5	12.85714

$$\begin{aligned} \sum_i k_i \ \alpha_i^2 &= 6885.014. \\ (\sum_i k_i \ \alpha_i)^2 / \sum_i k_i &= 6864.478. \\ \text{Diff.} &= 20.54 \text{ as before.} \end{aligned}$$

Sums of squares for columns uses the values below

	b_j	k_j
1	12.825	17.7778
2	13.875	7.2727
3	17.75	8.

$$\begin{aligned} \sum k_j \ b_j^2 &= 6844.712. \\ (\sum k_j)^2 / \sum k_j &= 6709.590. \\ \text{Diff.} &= 135.12 \text{ as before.} \end{aligned}$$

Another interesting method for obtaining estimates and tests involves setting up least squares equations using Lagrange multipliers to impose the following restrictions

$$\begin{aligned} \sum_i \gamma_{ij} &= 0 \text{ for } i = 1, \ldots, r. \\ \sum_i \gamma_{ij} &= 0 \text{ for } j = 1, \ldots, c. \\ \mu^o &= 0 \end{aligned}$$

A solution is

$$\begin{aligned} \mathbf{b}' &= (0, \ -14, \ -266, \ -144)/120. \\ \mathbf{t}' &= (1645, \ 1771, \ 2236)/120. \\ \boldsymbol{\gamma}' &= (-13, \ -31, \ 44, \ 19, \ 43, \ -62, \ 85, \ 55, \ -140, \ -91, \ -67, \ 158)/120. \end{aligned}$$

Using these values, $\hat{\mu}_{ij}$ are the $\bar{y}_{ij.}$, and the reduction in SS is

$$\sum_i \sum_j y_{ij.}^2 / n_{ij.} = 8207.5.$$

Next the SS for rows is this reduction minus the reduction when $\mathbf{b}^o$ is dropped from the equations restricted as before. A solution in that case is

$$\begin{aligned}\mathbf{t}' &= (12.8133,\ 14.1223,\ 17.3099).\\ \boldsymbol{\gamma}' &= (.4509,\ -.4619,\ .0110,\ .4358,\ -.1241,\ -.3117,\\ &\quad -.0897,\ 1.4953,\ -1.4055,\ -.7970, .9093,\ 1.7063),\end{aligned}$$

and the reduction is 8186.960. The row sums of squares is

$$8207.5 - 8186.960 = 20.54 \text{ as before.}$$

Now drop $\mathbf{t}^o$ from the equations. A solution is

$$\begin{aligned}\hat{\mathbf{b}}' &= (13.9002,\ 15.0562,\ 13.5475,\ 14.4887).\\ \hat{\boldsymbol{\gamma}}' &= (.9648,\ .9390,\ -1.9039,\ .2751,\ .2830,\ -.5581,\\ &\quad -.0825, .1309,\ -.0485,\ -1.1574,\ -1.3530,\ 2.5104),\end{aligned}$$

and the reduction is 8072.377, giving the column sums of squares as

$$8207.5 - 8072.377 = 135.12 \text{ as before.}$$

An interesting way to obtain estimates under the sum to 0 restrictions in $\boldsymbol{\gamma}$ is to solve

$$\bar{\mathbf{X}}_0'\ \bar{\mathbf{X}}_0 \begin{pmatrix} \mathbf{b}^o \\ \mathbf{t}^o \end{pmatrix} = \bar{\mathbf{X}}_0 \bar{\mathbf{y}},$$

where $\bar{\mathbf{X}}_0$ is the submatrix of $\bar{\mathbf{X}}$ referring to $\mathbf{b}$, $\mathbf{t}$ only, and $\bar{\mathbf{y}}$ is a vector of subclass means. These equations are

$$\begin{pmatrix} 3 & 0 & 0 & 0 & 1 & 1 & 1 \\ & 3 & 0 & 0 & 1 & 1 & 1 \\ & & 3 & 0 & 1 & 1 & 1 \\ & & & 3 & 1 & 1 & 1 \\ & & & & 4 & 0 & 0 \\ & & & & & 4 & 0 \\ & & & & & & 4 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ t_1 \\ t_2 \\ t_3 \end{pmatrix} = \begin{pmatrix} 47.1 \\ 46.75 \\ 40.45 \\ 43.5 \\ 51.3 \\ 55.5 \\ 71.0 \end{pmatrix}. \tag{22}$$

A solution is

$$\begin{aligned}\mathbf{b}' &= (0,\ -14,\ -266,\ -144)/120,\\ \mathbf{t}' &= (1645,\ 1771,\ 2236)/120.\end{aligned}$$

This is the same as in the restricted least squares solution. Then

$$\hat{\gamma}_{ij} = \bar{y}_{ij.} - b_i^o - t_j^o,$$

which gives the same result as before. More will be said about these alternative methods in the missing subclass case.

3 The Fixed, Missing Subclass Case

When one or more subclasses is missing, the estimates and tests described in Section 2 cannot be effected. What should be done in this case? There appears to be no agreement among statisticians. It is of course true that any linear functions of μ_{ij} in which $n_{ij} > 0$ can be estimated by BLUE and can be tested, but these may not be of any particular interest to the researcher. One method sometimes used, and this is the basis of a SAS Type 4 analysis, is to select a subset of subclasses, all filled, and then to do a weighted squares of means analysis on this subset. For example, suppose that in a 3×4 design, subclass (1,2) is missing. Then one could discard all data from the second column, leaving a 3×3 design with filled subclasses. This would mean that rows are compared by averaging over columns 1,3,4 and only columns 1,3,4 are compared, these averaged over the 3 rows. One could also discard the first row leaving a 2×4 design. The columns are compared by averaging over only rows 2 and 3, and only rows 2 and 3 are compared, averaging over all 4 columns. Consequently this method is not unique because usually more than one filled subset can be chosen. Further, most experimenters are not happy with the notion of discarding data that may have been costly to obtain.

Another possibility is to estimate μ_{ij} for missing subclasses by some biased procedure. For example, one can estimate μ_{ij} such that $E(\hat{\mu}_{ij}) = \mu + a_i + b_j +$ some function of the γ_{ij} associated with filled subclasses. One way of doing this is to set up least squares equations with the following restrictions.

$$\begin{aligned} \sum_j \gamma_{ij} &= 0 \text{ for } i = 1, \ldots, r. \\ \sum_i \gamma_{ij} &= 0 \text{ for } j = 1, \ldots, c. \\ \gamma_{ij} &= 0 \text{ if } n_{ij} = 0. \end{aligned}$$

This is the method used in Harvey's computer package. When equations with these restrictions are solved,

$$\hat{\mu}_{ij} = \mu + a_i^o + b_j^o + \gamma_{ij}^o = \bar{y}_{ij.},$$

when $n_{ij} > 0$ and thus is unbiased. A biased estimator for a missing subclass is $\mu^o + a_i^o + b_j^o$, and this has expectation $\mu + a_i + b_j + \sum_i \sum_j k_{ij}\gamma_{ij}$, where summation in the last term is over filled subclass and $\sum_i \sum_j k_{ij} = 1$. Harvey's package does not compute this but does produce "least squares means" for main effects and some of these are biased.

Thus $\hat{\mu}_{ij}$ is BLUE for filled subclasses and is biased for empty subclasses. In the class of estimators of μ_{ij} with expectation $\mu + a_i + b_j +$ some linear function of μ_{ij} associated with filled subclasses, this method minimizes the contribution of quadratics in $\boldsymbol{\gamma}$ to mean squared error when the squares and products of the elements of $\boldsymbol{\gamma}$ are in accord with no particular pattern of values. This minimization might appear to be a desirable property, but unfortunately the method does not control contributions of σ_e^2 to MSE. If one wishes to minimize the contribution of σ_e^2, but not to control on quadratics in $\boldsymbol{\gamma}$, while still having $E(\hat{\mu}_{ij})$ contain $\mu + a_i + b_j$, the way to accomplish this is to solve least squares equations with $\boldsymbol{\gamma}$ dropped. Then the biased estimators in this case for filled as well as empty subclasses, are

$$\hat{\mu}_{ij} = \mu^o + a_i^o + b_j^o. \tag{23}$$

A third possibility is to assume some prior values of σ_e^2 and squares and products of γ_{ij} and compute as in Section 9.1. Then all $\hat{\mu}_{ij}$ are biased by γ_{ij} but have in their expectations $\mu + a_i + b_j$. Finally one could relax the requirement of $\mu + a_i + b_j$ in the expectation of $\hat{\mu}_{ij}$. In that case one would assume average values of squares and products of the a_i and b_j as well as for the γ_j and use the method described in Section 9.1.

Of these biased methods, I would usually prefer the one in which priors on the $\boldsymbol{\gamma}$, but not on **a** and **b** are used. In most fixed, 2 way models the number of levels of **a** and **b** are too small to obtain a good estimate of the pseudo-variances of **a** and **b**.

We illustrate these methods with a 4×3 design with 2 missing subclasses as follows.

	n_{ij}				$y_{ij.}$			
	1	2	3	4	1	2	3	4
1	5	2	3	2	30	11	13	7
2	4	2	0	5	21	6	–	9
3	3	0	1	4	12	–	3	15

4 A Method Based On Assumption $\gamma_{ij} = 0$ If $n_{ij} = 0$

First we illustrate estimation under sum to 0 model for $\boldsymbol{\gamma}$ and in addition the assumption that $\gamma_{23} = \gamma_{32} = 0$. The simplest procedure for this set of restrictions is to solve for $\mathbf{a}^o$, $\mathbf{b}^o$ in equations (17.24).

$$\begin{pmatrix} 4 & 0 & 0 & 1 & 1 & 1 & 1 \\ & 3 & 0 & 1 & 1 & 0 & 1 \\ & & 3 & 1 & 0 & 1 & 1 \\ & & & 3 & 0 & 0 & 0 \\ & & & & 2 & 0 & 0 \\ & & & & & 2 & 0 \\ & & & & & & 3 \end{pmatrix} \begin{pmatrix} \mathbf{a}^o \\ \mathbf{b}^o \end{pmatrix} = \begin{pmatrix} 19.333 \\ 10.05 \\ 10.75 \\ 15.25 \\ 8.5 \\ 7.333 \\ 9.05 \end{pmatrix}. \tag{24}$$

The first right hand side is $\frac{30}{5} + \frac{11}{2} + \frac{13}{3} + \frac{7}{2} = 19.333$, etc. for others. A solution is (3.964, 2.286, 2.800, 2.067, 1.125, .285, 0). The estimates of μ_{ij} are $\bar{y}_{ij.}$ for filled subclasses and 2.286 + .285 for $\hat{\mu}_{23}$ and 2.800 + 1.125 for $\hat{\mu}_{32}$. If $\hat{\gamma}_{ij}$ are wanted they are

$$\hat{\gamma}_{11} = \bar{y}_{11} - 3.964 - 1.125$$

etc, for filled subclasses, and 0 for $\hat{\gamma}_{23}$ and $\hat{\gamma}_{32}$.

The same results can be obtained, but with much heavier computing by solving least squares equations with restrictions on γ that are $\sum_j \gamma_{ij} = 0$ for all i, $\sum_i \gamma_{ij} = 0$ for all j, and $\gamma_{ij} = 0$ for subclasses with $n_{ij} = 0$. From these equations one can obtain sums of squares that mimic weighted squares of means. A solution to the restricted equations is

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{a}^o &= (3.964,\ 2.286,\ 2.800,\ 2.067)'. \\
\mathbf{b}^o &= (1.125,\ .285,\ 0)'. \\
\boldsymbol{\gamma}^o &= (-.031,\ .411,\ .084,\ -.464,\ .897,\ -.411, \\
& \quad\ 0,\ -.486,\ -.866,\ \ 0,\ -.084,\ .950)'.
\end{aligned}$$

Note that the solution to $\boldsymbol{\gamma}$ conforms to the restrictions imposed. Also note that this solution is the same as the one previously obtained. Further, $\hat{\mu}_{ij} = \mu^o + a_i^o + b_j^o + \gamma_{ij}^o = \bar{y}_{ij.}$ for filled subclasses.

A test of hypothesis that the main effects are equal, that is $\bar{\mu}_{i.} = \bar{\mu}_{i'}$ for all pairs of i, i$'$, can be effected by taking a new solution to the restricted equations with $\mathbf{a}^o$ dropped. Then the SS for rows is

$$(\boldsymbol{\beta}^o)' \text{ RHS} - (\boldsymbol{\beta}_*^o)' \text{ RHS}_*, \tag{25}$$

where $\boldsymbol{\beta}^o$ is a solution to the full set of equations, and this reduction is simply $\sum_i \sum_j y_{ij}^2/n_{ij.}$, $\boldsymbol{\beta}_*^o$ is a solution with $\mathbf{a}$ deleted from the set of equations, and RHS_* is the right hand side. This tests a nontestable hypothesis inasmuch as the main effects are not estimable when subclasses are missing. The test is valid only if γ_{ij} are truly 0 for all missing subclasses, and this is not a testable assumption, Henderson and McAllister (1978). If one is to use a test based on non-estimable functions, as is done in this case, there should be some attempt to evaluate the numerator with respect to quadratics in fixed effects other than those being tested and use this in the denominator. That is, a minimum requirement could seem to be a test of this sort.

$$\begin{aligned}
E(\text{numerator}) &= Q_t(\mathbf{a}) + Q \text{ (fixed effects causing bias in the estimator)} \\
& \quad + \text{linear functions of random variables.}
\end{aligned}$$

Then the denominator should have the same expectation except that $Q_t(\mathbf{a})$, the quadratic in fixed effects being tested, would not be present. In our example the reduction under the full model with restrictions on $\boldsymbol{\gamma}$ is 579.03, and this is the same as the uncorrected subclass sum of squares. A solution with $\boldsymbol{\gamma}$ restricted as before and with $\mathbf{a}$ dropped is

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{b}^o &= (5.123,\ 4.250,\ 4.059,\ 2.790)', \\
\boldsymbol{\gamma}^o &= (.420,\ .129,\ -.119,\ -.430,\ .678,\ -.129, \\
& \quad\ 0,\ -.549,\ -1.098,\ \ 0,\ .119,\ .979)'.
\end{aligned}$$

This gives a reduction of 566.32. Then the sum of squares with 2 df for the numerator is 579.03-566.32, but $\hat{\sigma}_e^2$ is not an appropriate denominator MS, when $\hat{\sigma}_e^2$ is the within subclass mean square, unless γ_{23} and γ_{32} are truly equal to zero, and we cannot test this assumption.

Similarly a solution when $\mathbf{b}$ is dropped is

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{a}^o &= (5.089, 3.297, 3.741)', \\
\boldsymbol{\gamma}^o &= (.098,\ .254,\ -.355,\ .003,\ 1.035,\ -.254, \\
&\quad 0,\ -.781,\ -1.133,\ \ 0,\ .355,\ .778)'.
\end{aligned}$$

The reduction is 554.81. Then if γ_{23} and $\gamma_{32} = 0$, the numerator sum of squares with 3 df is 579.03-554.81. The sum of squares for interaction with (3-1)(4-1)-2 = 4 df. is 579.03 - reduction with $\boldsymbol{\gamma}$ and the Lagrange multiplier deleted. This latter reduction is 567.81 coming from a solution

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{a}^o &= (3.930,\ 2.296,\ 2.915)',\ \text{and} \\
\boldsymbol{\beta}^o &= (2.118, 1.137, .323, 0)'.
\end{aligned}$$

5 Biased Estimation By Ignoring γ

Another biased estimation method sometimes suggested is to ignore $\boldsymbol{\gamma}$. That is, least squares equations with only μ^o, $\mathbf{a}^o$, $\mathbf{b}^o$ are solved. This is sometimes called the method of fitting constants, Yates (1934). This method has quite different properties than the method of Section 17.4. Both obtain estimators of μ_{ij} with expectations $\mu + a_i + b_j$ + linear functions of γ_{ij}. The method of section 17.4 minimizes the contribution of quadratics in $\boldsymbol{\gamma}$ to MSE, but does a poor job of controlling on the contribution of σ_e^2. In contrast, the method of fitting constants minimizes the contribution of σ_e^2 but does not control quadratics in $\boldsymbol{\gamma}$. The method of the next section is a compromise between these two extremes.

A solution for our example for the method of this section is

$$\begin{aligned}
\mu^o &= 0, \\
\mathbf{a}^o &= (3.930,\ 2.296,\ 2.915)', \\
\mathbf{b}^o &= (2.118,\ 1.137,\ .323,\ \ 0)'.
\end{aligned}$$

Then if we wish $\hat{\mu}_{ij}$ these are $\mu^o + a_i^o + b_j^o$.

A test of row effects often suggested is to compute the reduction in SS under the model with $\boldsymbol{\gamma}$ dropped minus the reduction when $\mathbf{a}$ and $\boldsymbol{\gamma}$ are dropped, the latter being simply $\sum_j y_{.j.}^2/n_{.j.}$. Then this is tested against some denominator. If $\hat{\sigma}_e^2$ is used, the denominator is too small unless $\boldsymbol{\gamma}$ is $\mathbf{0}$. If $R \times C$ for MS is used, the denominator is probably too large. Further, the numerator is not a test of rows averaged in some logical way across columns, but rather each row is averaged differently depending upon the pattern of subclass numbers. That is, $\mathbf{K}'\boldsymbol{\beta}$ is dependent upon the incidence matrix, an obviously undesirable property.

6 Priors On Squares And Products Of γ

The methods of the two preceding sections control in the one case on $\boldsymbol{\gamma}$ and the other on σ_e^2 as contributors to MSE. The method of this section is an attempt to control on both. The logic of the method depends upon the assumption that there is no pattern of values of $\boldsymbol{\gamma}$, such, for example as linear by columns or linear by rows. Then consider the matrix of squares and products of elements of γ_{ij} for all possible permutations of rows and columns. The average values are found to be

$$\begin{aligned}\gamma_{ij}^2 &= \alpha. \\ \gamma_{ij}\gamma_{ij'} &= -\alpha/(c-1). \\ \gamma_{ij}\gamma_{i'j} &= -\alpha/(r-1). \\ \gamma_{ij}\gamma_{i'j'} &= \alpha/(r-1)(c-1). \end{aligned} \tag{26}$$

Note that if we substitute σ_γ^2 for α, this is the same matrix as that for $Var(\boldsymbol{\gamma})$ in the finite random rows and finite random columns model. Then if we have estimates of σ_e^2 and α or an estimate of the relative magnitudes of these parameters, we can proceed to estimate with $\mathbf{a}$ and $\mathbf{b}$ regarded as fixed and $\boldsymbol{\gamma}$ regarded as a pseudo random variable.

We illustrate with our same numerical example. Assume that $\sigma_e^2 = 20$ and $\alpha = 6$. Write the least squares equations that include γ_{23} and γ_{32}, the missing subclasses. Premultiply the last 12 equations by

$$\begin{pmatrix} 6 & -2 & -2 & -2 & -3 & 1 & 1 & 1 & -3 & 1 & 1 & 1 \\ & 6 & -2 & -2 & 1 & -3 & 1 & 1 & 1 & -3 & 1 & 1 \\ & & 6 & -2 & 1 & 1 & -3 & 1 & 1 & 1 & -3 & 1 \\ & & & 6 & 1 & 1 & 1 & -3 & 1 & 1 & 1 & -3 \\ & & & & 6 & -2 & -2 & -2 & -3 & 1 & 1 & 1 \\ & & & & & 6 & -2 & -2 & 1 & -3 & 1 & 1 \\ & & & & & & 6 & -2 & 1 & 1 & -3 & 1 \\ & & & & & & & 6 & 1 & 1 & 1 & -3 \\ & & & & & & & & 6 & -2 & -2 & -2 \\ & & & & & & & & & 6 & -2 & -2 \\ & & & & & & & & & & 6 & -2 \\ & & & & & & & & & & & 6 \end{pmatrix}. \tag{27}$$

Then add 1 to each of the last 12 diagonals. The resulting coefficient matrix is (17.28) ... (17.31). The right hand side vector is (3.05, 1.8, 1.5, 3.15, .85, .8, 2.6, .4, 1.8, -4.8, .95, 1.15, -2.25, .15, -3.55, -1.55, .45, 4.65)′ $\boldsymbol{\beta}' = (a_1 a_2 a_3 b_1 b_2 b_3 \boldsymbol{\gamma}')$. Thus μ and b_4 are deleted, which is equivalent to obtaining a solution with $\mu^o = 0$, $b_4^o = 0$.

Upper left 9×9

$$\begin{pmatrix} .6 & 0 & 0 & .25 & .1 & .15 & .25 & .1 & .15 \\ 0 & .55 & 0 & .2 & .1 & 0 & 0 & 0 & 0 \\ 0 & 0 & .4 & .15 & 0 & .05 & 0 & 0 & 0 \\ .25 & .2 & .15 & .6 & 0 & 0 & .25 & 0 & 0 \\ .1 & .1 & 0 & 0 & .2 & 0 & 0 & .1 & 0 \\ .15 & 0 & .05 & 0 & 0 & .2 & 0 & 0 & .15 \\ .8 & -.25 & -.2 & .45 & -.1 & -.25 & 2.5 & -.2 & -.3 \\ -.4 & .15 & .4 & -.15 & .3 & -.25 & -.5 & 1.6 & -.3 \\ 0 & .55 & .2 & -.15 & -.1 & .75 & -.5 & -.2 & 1.9 \end{pmatrix} \tag{28}$$

Upper right 9×9

$$\begin{pmatrix} .1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .2 & .1 & 0 & .25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .15 & 0 & .05 & .2 \\ 0 & .2 & 0 & 0 & 0 & .15 & 0 & 0 & 0 \\ 0 & 0 & .1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .05 & 0 \\ -.2 & -.6 & .1 & 0 & .25 & -.45 & 0 & .05 & .2 \\ -.2 & .2 & -.3 & 0 & .25 & .15 & 0 & .05 & .2 \\ -.2 & .2 & .1 & 0 & .25 & .15 & 0 & -.15 & .2 \end{pmatrix} \tag{29}$$

Lower left 9×9

$$\begin{pmatrix} -.4 & -.45 & -.4 & -.15 & -.1 & -.25 & -.5 & -.2 & -.3 \\ -.4 & .5 & -.2 & 0 & -.1 & .2 & -.75 & .1 & .15 \\ .2 & -.3 & .4 & 0 & .3 & .2 & .25 & -.3 & .15 \\ 0 & -1.1 & .2 & 0 & -.1 & -.6 & .25 & .1 & -.45 \\ .2 & .9 & -.4 & 0 & -.1 & .2 & .25 & .1 & .15 \\ -.4 & -.25 & .4 & -.45 & .2 & .05 & -.75 & .1 & .15 \\ .2 & .15 & -.8 & .15 & -.6 & .05 & .25 & -.3 & .15 \\ 0 & .55 & -.4 & .15 & .2 & -.15 & .25 & .1 & -.45 \\ .2 & -.45 & .8 & .15 & .2 & .05 & .25 & .1 & .15 \end{pmatrix} \tag{30}$$

Lower right 9×9

$$\begin{pmatrix}
1.6 & .2 & .1 & 0 & -.75 & .15 & 0 & .05 & -.6 \\
.1 & 2.2 & -.2 & 0 & -.5 & -.45 & 0 & .05 & .2 \\
.1 & -.4 & 1.6 & 0 & -.5 & .15 & 0 & .05 & .2 \\
.1 & -.4 & -.2 & 1.0 & -.5 & .15 & 0 & -.15 & .2 \\
-.3 & -.4 & -.2 & 0 & 2.5 & .15 & 0 & .05 & -.6 \\
.1 & -.6 & .1 & 0 & .25 & 1.9 & 0 & -.1 & -.4 \\
.1 & .2 & -.3 & 0 & .25 & -.3 & 1.0 & -.1 & -.4 \\
.1 & .2 & .1 & 0 & .25 & -.3 & 0 & 1.3 & -.4 \\
-.3 & .2 & .1 & 0 & -.75 & -.3 & 0 & -.1 & 2.2
\end{pmatrix} \tag{31}$$

The solution is

$$\begin{aligned}
\mathbf{a}^o &= (3.967,\ 2.312,\ 2.846)'. \\
\mathbf{b}^o &= (2.068,\ 1.111,\ .288,\ \ 0)'.
\end{aligned}$$

$\boldsymbol{\gamma}^o$ displayed as a table is

	1	2	3	4
1	-.026	.230	.050	-.255
2	.614	-.230	0	-.384
3	-.588	0	-.050	.638

Note that the γ^o_{ij} sum to 0 by rows and columns. Now the $\hat{\mu}_{ij} = a^o_i + b^o_j + \gamma^o_{ij}$. The same solution can be obtained more easily by treating $\boldsymbol{\gamma}$ as a random variable with $Var = 12\mathbf{I}$. The value 12 comes from $\frac{rc}{(r-1)(c-1)}\ 6 = \frac{(3)4}{(2)3}\ (6) = 12$. The resulting coefficient matrix (times 60) is in (17.32). The right hand side vector is (3.05, 1.8, 1.5, 3.15, .85, .8, 1.5, .55, .65, .35, 1.05, .3, 0, .45, .6, 0, .15, .75)$'$. μ and b_4 are dropped as before.

$$\begin{pmatrix}
36 & 0 & 0 & 15 & 6 & 9 & 15 & 6 & 9 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & 33 & 0 & 12 & 6 & 0 & 0 & 0 & 0 & 0 & 12 & 6 & 0 & 15 & 0 & 0 & 0 & 0 \\
 & & 24 & 9 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 3 & 12 \\
 & & & 36 & 0 & 0 & 15 & 0 & 0 & 0 & 12 & 0 & 0 & 0 & 9 & 0 & 0 & 0 \\
 & & & & 12 & 0 & 0 & 6 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & 12 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0
\end{pmatrix} \tag{32}$$

$$\text{diag } (20,11,14,11,17,11,5,20,14,5,8,17)$$

The solution is the same as before. This is clearly an easier procedure than using the equations of (17.28). The inverse of the matrix of (17.28) post-multiplied by

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix},$$

where $\mathbf{P}$ = the matrix of (17.27), is not the same as the inverse of the matrix of (17.32) with diagonal $\mathbf{G}$, but if we pre-multiply each of them by $\mathbf{K}'$ and then post-multiply by $\mathbf{K}$, where $\mathbf{K}'$ is the representation of μ_{ij} in terms of $\mathbf{a}$, $\mathbf{b}$, $\boldsymbol{\gamma}$, we obtain the same matrix, which is the mean squared error for the $\hat{\mu}_{ij}$ under the priors used, $\sigma_e^2 = 20$ and $\alpha = 6$. Biased estimates of $\hat{\mu}_{ij}$ are in both methods

$$\begin{pmatrix} 6.009 & 5.308 & 4.305 & 3.712 \\ 4.994 & 3.192 & 2.600 & 1.928 \\ 4.327 & 3.957 & 3.084 & 3.484 \end{pmatrix}.$$

The estimated MSE matrix of this vector is
Upper left 8 × 8

$$\begin{pmatrix} 3.58 & .32 & .18 & .46 & .28 & -.32 & -.87 & -.09 \\ & 8.34 & .15 & .64 & -.43 & 1.66 & -2.29 & -.32 \\ & & 6.01 & .38 & .02 & -.15 & 4.16 & .04 \\ & & & 7.64 & -.29 & -.64 & -1.77 & .49 \\ & & & & 4.40 & .43 & 1.93 & .31 \\ & & & & & 8.34 & 2.29 & .32 \\ & & & & & & 33.72 & 1.54 \\ & & & & & & & 3.63 \end{pmatrix}$$

Upper right 8 × 4

$$\begin{pmatrix} .33 & -.70 & -.55 & -.11 \\ .04 & 5.21 & -.45 & .08 \\ -.33 & -1.33 & 1.97 & -.24 \\ -.37 & -1.47 & -1.14 & .57 \\ .34 & -1.09 & -.06 & -.24 \\ -.04 & 4.79 & .45 & -.08 \\ -1.12 & -4.67 & 7.51 & -1.04 \\ -.25 & -1.04 & -.13 & .22 \end{pmatrix}$$

Lower right 4 × 4

$$\begin{pmatrix} 5.66 & 2.62 & 1.00 & .50 \\ & 33.60 & 4.00 & 2.03 \\ & & 14.08 & .73 \\ & & & 4.44 \end{pmatrix}$$

Suppose we wish an approximate test of the hypothesis that $\bar{\mu}_{i.}$ are equal. In this case we could write $\mathbf{K}'$ as

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{-1}{4} & \frac{-1}{4} & \frac{-1}{4} & \frac{-1}{4} \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{-1}{4} & \frac{-1}{4} & \frac{-1}{4} \end{pmatrix}.$$

Then compute $\mathbf{K}'\mathbf{CK}$, where $\mathbf{C}$ is either the g-inverse of (17.28) post-multiplied by

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix},$$

or the g-inverse of the matrix using diagonal $\mathbf{G}$. This 2×2 matrix gives the MSE for $\sigma_e^2 =$ 20, $\alpha = 6$. Finally premultiply the inverse of this matrix by $(\mathbf{K}'\boldsymbol{\beta}^o)'$ and post-multiply by $\mathbf{K}'\boldsymbol{\beta}^o$. This quantity is distributed approximately as χ_2^2 under the null hypothesis.

7 Priors On Squares And Products Of a, b, And γ

Another possibility for biased estimation is to require only that

$$E(\hat{\mu}_{ij}) = \mu + \text{linear functions of } \mathbf{a}, \mathbf{b},\ \boldsymbol{\gamma}.$$

We do this by assuming prior values of squares and products of $\mathbf{a}$ and of $\mathbf{b}$ as

$$\begin{pmatrix} 1 & & \frac{-1}{r-1} \\ & \ddots & \\ \frac{-1}{r-1} & & 1 \end{pmatrix} \sigma_a^2 \ \text{ and } \begin{pmatrix} 1 & & \frac{-1}{c-1} \\ & \ddots & \\ \frac{-1}{c-1} & & 1 \end{pmatrix} \sigma_b^2,$$

respectively, where σ_a^2 and σ_b^2 are pseudo-variances. The prior on $\boldsymbol{\gamma}$ is the same as in Section 17.6. Then we apply the method for singular G.

To illustrate in our example, let the priors be $\alpha_e^2 = 20$, $\alpha_a^2 = 4$, $\alpha_b^2 = 9$, $\alpha_\gamma^2 = 6$. Then we multiply all equations except the first pertaining to μ by

$$\begin{pmatrix} \mathbf{P}_a\sigma_a^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_b\sigma_b^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_\gamma\sigma_\gamma^2 \end{pmatrix},$$

and add 1's to all diagonals except the first. This yields the equations with coefficient matrix in (17.33) . . . (17.36) and right hand vector = (6.35, 5.60, -1.90, -3.70, 18.75, -8.85, -9.45, -.45, 2.60, .40, 1.80, -4.80, .95, 1.15, -2.25, .15, -3.55, -1.55, .45, 4.65)$'$.

Upper left 10 $\times$ 10

$$\begin{pmatrix}
1.55 & .6 & .55 & .4 & .6 & .2 & .2 & .55 & .25 & .1 \\
.5 & 3.4 & -1.1 & -.8 & .3 & .2 & .5 & -.5 & 1.0 & .4 \\
.2 & -1.2 & 3.2 & -.8 & 0 & .2 & -.4 & .4 & -.5 & -.2 \\
-.7 & -1.2 & -1.1 & 2.6 & -.3 & -.4 & -.1 & .1 & -.5 & -.2 \\
2.55 & 1.2 & .75 & .6 & 6.4 & -.6 & -.6 & -1.65 & 2.25 & -.3 \\
-2.25 & -.6 & -.45 & -1.2 & -1.8 & 2.8 & -.6 & -1.65 & -.75 & .9 \\
-2.25 & 0 & -1.65 & -.6 & -1.8 & -.6 & 2.8 & -1.65 & -.75 & -.3 \\
1.95 & -.6 & 1.35 & 1.2 & -1.8 & -.6 & -.6 & 5.95 & -.75 & -.3 \\
.35 & .8 & -.25 & -.2 & .45 & -.1 & -.25 & .25 & 2.5 & -.2 \\
.15 & -.4 & .15 & .4 & -.15 & .3 & -.25 & .25 & -.5 & 1.6
\end{pmatrix} \tag{33}$$

Upper right 10×10

$$\begin{pmatrix} .15 & .1 & .2 & .1 & 0 & .25 & .15 & 0 & .05 & .2 \\ .6 & .4 & -.4 & -.2 & 0 & -.5 & -.3 & 0 & -.1 & -.4 \\ -.3 & -.2 & .8 & .4 & 0 & 1.0 & -.3 & 0 & -.1 & -.4 \\ -.3 & -.2 & -.4 & -.2 & 0 & -.5 & .6 & 0 & .2 & .8 \\ -.45 & -.3 & 1.8 & -.3 & 0 & -.75 & 1.35 & 0 & -.15 & -.6 \\ -.45 & -.3 & -.6 & .9 & 0 & -.75 & -.45 & 0 & -.15 & -.6 \\ 1.35 & -.3 & -.6 & -.3 & 0 & -.75 & -.45 & 0 & .45 & -.6 \\ -.45 & .9 & -.6 & -.3 & 0 & 2.25 & -.45 & 0 & -.15 & 1.8 \\ -.3 & -.2 & -.6 & .1 & 0 & .25 & -.45 & 0 & .05 & .2 \\ -.3 & -.2 & .2 & -.3 & 0 & .25 & .15 & 0 & .05 & .2 \end{pmatrix} \tag{34}$$

Lower left 10×10

$$\begin{pmatrix} .75 & 0 & .55 & .2 & -.15 & -.1 & .75 & .25 & -.5 & -.2 \\ -1.25 & -.4 & -.45 & -.4 & -.15 & -.1 & -.25 & -.75 & -.5 & -.2 \\ -.1 & -.4 & .5 & -.2 & 0 & -.1 & .2 & -.2 & -.75 & .1 \\ .3 & .2 & -.3 & .4 & 0 & .3 & .2 & -.2 & .25 & -.3 \\ -.9 & 0 & -1.1 & .2 & 0 & -.1 & -.6 & -.2 & .25 & .1 \\ .7 & .2 & .9 & -.4 & 0 & -.1 & .2 & .6 & .25 & .1 \\ -.25 & -.4 & -.25 & .4 & -.45 & .2 & .05 & -.05 & -.75 & .1 \\ -.45 & .2 & .15 & -.8 & .15 & -.6 & .05 & -.05 & .25 & -.3 \\ .15 & 0 & .55 & -.4 & .15 & .2 & -.15 & -.05 & .25 & .1 \\ .55 & .2 & -.45 & .8 & .15 & .2 & .05 & .15 & .25 & .1 \end{pmatrix} \tag{35}$$

Lower right 10×10

$$\begin{pmatrix} 1.9 & -.2 & .2 & .1 & 0 & .25 & .15 & 0 & -.15 & .2 \\ -.3 & 1.6 & .2 & .1 & 0 & -.75 & .15 & 0 & .05 & -.6 \\ .15 & .1 & 2.2 & -.2 & 0 & -.5 & -.45 & 0 & .05 & .2 \\ .15 & .1 & -.4 & 1.6 & 0 & -.5 & .15 & 0 & .05 & .2 \\ -.45 & .1 & -.4 & -.2 & 1.0 & -.5 & .15 & 0 & -.15 & .2 \\ .15 & -.3 & -.4 & -.2 & 0 & 2.5 & .15 & 0 & .05 & -.6 \\ .15 & .1 & -.6 & .1 & 0 & .25 & 1.9 & 0 & -.1 & -.4 \\ .15 & .1 & .2 & -.3 & 0 & .25 & -.3 & 1.0 & -.1 & -.4 \\ -.45 & .1 & .2 & .1 & 0 & .25 & -.3 & 0 & 1.3 & -.4 \\ .15 & -.3 & .2 & .1 & 0 & -.75 & -.3 & 0 & -.1 & 2.2 \end{pmatrix} \tag{36}$$

The solution is

$$\begin{aligned} \hat{\mu} &= 4.014, \\ \hat{\mathbf{a}} &= (.650,\ -.467,\ -.183), \\ \hat{\mathbf{b}} &= (.972,\ .120,\ -.208,\ -.885), \\ \hat{\boldsymbol{\gamma}} &= \begin{pmatrix} .111 & .225 & -.170 & -.166 \\ .489 & -.276 & .303 & -.515 \\ -.599 & .051 & -.133 & .681 \end{pmatrix}. \end{aligned}$$

Note that $\sum_i \hat{a}_i = \sum_j \hat{b}_j = 0$ and the $\hat{\gamma}_{ij}$ sum to 0 by rows and columns. A solution can be obtained by pretending that $\mathbf{a}$, $\mathbf{b}$, $\boldsymbol{\gamma}$ are random variables with $Var(\mathbf{a}) = 3\mathbf{I}$, $Var(\mathbf{b}) = 8\mathbf{I}$, $Var(\boldsymbol{\gamma}) = 12\mathbf{I}$. The coefficient matrix of these is in (17.37) ... (17.39) and the right hand side is (6.35, 3.05, 1.8, 1.5, 3.15, .85, .8, 1.55, 1.5, .55, .65, .35, 1.05, .3, 0, .45, .6, 0, .15, .75)'. The solution is

$$\begin{aligned}
\hat{\mu} &= 4.014, \\
\hat{\mathbf{a}} &= (.325,\ -.233,\ -.092)', \\
\hat{\mathbf{b}} &= (.648,\ .080,\ -.139,\ -.590)', \\
\hat{\boldsymbol{\gamma}} &= \begin{pmatrix} .760 & .590 & .086 & -.136 \\ .580 & -.470 & 0 & -1.043 \\ -.367 & 0 & -.294 & .295 \end{pmatrix}.
\end{aligned}$$

This is a different solution from the one above, but the $\hat{\mu}_{ij}$ are identical for the two. These are as follows, in table form,

$$\begin{pmatrix} 5.747 & 5.009 & 4.286 & 3.613 \\ 5.009 & 3.391 & 3.642 & 2.148 \\ 4.204 & 4.003 & 3.490 & 3.627 \end{pmatrix}.$$

Note that $\sum_i \hat{a}_i = \sum_j \hat{b}_j = 0$, but the $\hat{\gamma}_{ij}$ do not sum to 0 by rows and columns.

$$\begin{aligned}
\textstyle\sum_j \hat{\gamma}_{ij} &= \sigma^2_\gamma\ \hat{a}_i/\sigma^2_a = 4\ \hat{a}_i. \\
\textstyle\sum_i \hat{\gamma}_{ij} &= \sigma^2_\gamma\ \hat{b}_j/\sigma^2_b = 1.5\ \hat{b}_j.
\end{aligned}$$

The pseudo variances come from (15.18) and (15.19).

$$\begin{aligned}
\sigma^2_{*a} &= \frac{3}{2}\ (4) - \frac{3}{(2)3}\ (6) = 3. \\
\sigma^2_{*b} &= \frac{4}{3}\ (9) - \frac{4}{(2)3}\ (6) = 8. \\
\sigma^2_{*\gamma} &= \frac{(3)4}{(2)3}\ (6) = 12.
\end{aligned}$$

Upper left 8 × 8

$$120^{-1} \begin{pmatrix} 186 & 72 & 66 & 48 & 72 & 24 & 24 & 66 \\ & 112 & 0 & 0 & 30 & 12 & 18 & 12 \\ & & 106 & 0 & 24 & 12 & 0 & 30 \\ & & & 88 & 18 & 0 & 6 & 24 \\ & & & & 87 & 0 & 0 & 0 \\ & & & & & 39 & 0 & 0 \\ & & & & & & 39 & 0 \\ & & & & & & & 81 \end{pmatrix} \quad (37)$$

Lower left 12×8 and (upper right 12×8)'

$$120^{-1} \begin{pmatrix} 30 & 30 & 0 & 0 & 30 & 0 & 0 & 0 \\ 12 & 12 & 0 & 0 & 0 & 12 & 0 & 0 \\ 18 & 18 & 0 & 0 & 0 & 0 & 18 & 0 \\ 12 & 12 & 0 & 0 & 0 & 0 & 0 & 12 \\ 24 & 0 & 24 & 0 & 24 & 0 & 0 & 0 \\ 12 & 0 & 12 & 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 30 & 0 & 30 & 0 & 0 & 0 & 0 & 30 \\ 18 & 0 & 0 & 18 & 18 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 6 & 0 & 0 & 6 & 0 \\ 24 & 0 & 0 & 24 & 0 & 0 & 0 & 24 \end{pmatrix} \tag{38}$$

Lower 12×12

$$= 120^{-1} \text{ diag } (40,\ 22,\ 28,\ 22,\ 34,\ 22,\ 10,\ 40,\ 28,\ 10,\ 16,\ 34) \tag{39}$$

Approximate tests of hypotheses can be effected as described in the previous section. $\mathbf{K}'$ for SS Rows is (times .25)

$$\begin{pmatrix} 0 & 4 & 0 & -4 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 \\ 0 & 0 & 4 & -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{pmatrix}.$$

$\mathbf{K}'$ for SS columns is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 3 & 0 & 0 & 3 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 3 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 3 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & -1 \end{pmatrix} /3.$$

8 The Two Way Mixed Model

The two way mixed model is one in which the elements of the rows (or columns) are a random sample from some population of rows (or columns), and the levels of columns (or rows) are fixed. We shall deal with random rows and fixed columns. There is really more than one type of mixed model, as we shall see, depending upon the variance-covariance matrices, $Var(\mathbf{a})$ and $Var(\boldsymbol{\gamma})$, and consequently $Var(\boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ = vector of elements, $\mu + a_i + b_j + \gamma_{ij}$. The most commonly used model is

$$Var(\boldsymbol{\alpha}) = \begin{pmatrix} \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{C} \end{pmatrix}, \tag{40}$$

where $\mathbf{C}$ is $q \times q$, q being the number of columns. There are p such blocks down the diagonal, where p is the number of rows. $\mathbf{C}$ is a matrix with every diagonal = v and every off-diagonal = c. If the rows were sires and the columns were traits and if $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$, this would imply that the heritability is the same for every trait, 4 $v/(4v + \sigma_e^2)$, and the genetic correlation between any pair of traits is the same, c/v. This set of assumptions should be questioned in most mixed models. Is it logical to assume that $Var(\alpha_{ij}) = Var(\alpha_{ij'})$ and that $Cov(\alpha_{ij},\ \alpha_{ik}) = Cov(\alpha_{ij},\ \alpha_{im})$? Also is it logical to assume that $Var(e_{ijk}) = Var(e_{ij'k})$? Further we cannot necessarily assume that α_{ij} is uncorrelated with $\alpha_{i'j}$. This would not be true if the i^{th} sire is related to the i' sire. We shall deal more specifically with these problems in the context of multiple trait evaluation.

Now let us consider what assumptions regarding

$$Var\begin{pmatrix} \mathbf{a} \\ \boldsymbol{\gamma} \end{pmatrix}$$

will lead to $Var(\boldsymbol{\alpha})$ like (17.40). Two models commonly used in statistics accomplish this. The first is based on the model for unrelated interactions and main effects formulated in Section 15.4.

$$Var(\mathbf{a}) = \mathbf{I}\sigma_a^2,$$

since the number of levels of $\mathbf{a}$ in the population $\rightarrow \infty$, and

$$\begin{aligned}
Var(\gamma_{ij}) &= \sigma_\gamma^2. \\
Cov(\gamma_{ij},\ \gamma_{ij'}) &= -\sigma_\gamma^2/(q-1). \\
Cov(\gamma_{ij},\ \gamma_{i'j}) &= -\sigma_\gamma^2/(\text{one less than population levels of } a) = 0. \\
Cov(\gamma_{ij},\ \gamma_{i'j}) &= -\sigma_\gamma^2/(q-1)\ (\text{one less than population levels of } a) = 0.
\end{aligned}$$

This leads to

$$Var(\boldsymbol{\gamma}) = \begin{pmatrix} \mathbf{P} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{P} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \end{pmatrix} \sigma_y^2, \tag{41}$$

where $\mathbf{P}$ is a matrix with $1'$s in diagonals and $-1/(q-1)$ in all off-diagonals. Under this model

$$\begin{aligned}
Var(\alpha_{ij}) &= \sigma_a^2 + \sigma_\gamma^2. \\
Cov(\alpha_{ij},\ \alpha_{ij'}) &= \sigma_a^2 - \sigma_\gamma^2/(q-1).
\end{aligned} \tag{42}$$

An equivalent model often used that is easier from a computational standpoint, but less logical is

$$\begin{aligned}
Var(\mathbf{a}_*) &= \mathbf{I}\sigma_{*a}^2,\ \text{where } \sigma_{*a}^2 = \sigma_a^2 - \sigma_\gamma^2/(q-1). \\
Var(\gamma_*) &= \mathbf{I}\sigma_{*\gamma}^2,\ \text{where } \sigma_{*\gamma}^2 = q\sigma_\gamma^2/(q-1).
\end{aligned} \tag{43}$$

Note that we have re-labelled the row and interaction effects because these are not the same variables as $\mathbf{a}$ and $\boldsymbol{\gamma}$.

The results of (17.43) come from principles described in Section 15.9. We illustrate these two models (and estimation and prediction methods) with our same two way example. Let

$$\begin{aligned} Var(\mathbf{e}) &= 20\mathbf{I}, Var(\mathbf{a}) = 4\mathbf{I}, \text{ and} \\ Var(\boldsymbol{\gamma}) &= 6 \begin{pmatrix} \mathbf{P} & 0 & 0 \\ 0 & \mathbf{P} & 0 \\ 0 & 0 & \mathbf{P} \end{pmatrix}, \end{aligned}$$

where $\mathbf{P}$ is a 4×4 matrix with 1's for diagonals and $-1/3$ for all off-diagonals. We set up the least squares equations with μ deleted, multiply the first 3 equations by 4 $\mathbf{I}_3$ and the last 12 equations by $Var(\boldsymbol{\gamma})$ described above. Then add 1 to the first 4 and the last 12 diagonal coefficients. This yields equations with coefficient matrix in (17.44) ... (17.47). The right hand side is (12.2, 7.2, 6.0, 3.15, .85, .8, 1.55, 5.9, -1.7, -.9, -3.3, 4.8, -1.2, -3.6, 0, 1.8, -3.0, -1.8, 3.0)'.

Upper left 10 × 10

$$\begin{pmatrix} 3.4 & 0 & 0 & 1.0 & .4 & .6 & .4 & 1.0 & .4 & .6 \\ 0 & 3.2 & 0 & .8 & .4 & 0 & 1.0 & 0 & 0 & 0 \\ 0 & 0 & 2.6 & .6 & 0 & .2 & .8 & 0 & 0 & 0 \\ .25 & .2 & .15 & .6 & 0 & 0 & 0 & .25 & 0 & 0 \\ .1 & .1 & 0 & 0 & .2 & 0 & 0 & 0 & .1 & 0 \\ .15 & 0 & .05 & 0 & 0 & .2 & 0 & 0 & 0 & .15 \\ .1 & .25 & .2 & 0 & 0 & 0 & .55 & 0 & 0 & 0 \\ .8 & 0 & 0 & 1.5 & -.2 & -.3 & -.2 & 2.5 & -.2 & -.3 \\ -.4 & 0 & 0 & -.5 & .6 & -.3 & -.2 & -.5 & 1.6 & -.3 \\ 0 & 0 & 0 & -.5 & -.2 & .9 & -.2 & -.5 & -.2 & 1.9 \end{pmatrix} \quad (44)$$

Upper right 10 × 9

$$\begin{pmatrix} .4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .8 & .4 & 0 & 1.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .6 & 0 & .2 & .8 \\ 0 & .2 & 0 & 0 & 0 & .15 & 0 & 0 & 0 \\ 0 & 0 & .1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .05 & 0 \\ .1 & 0 & 0 & 0 & .25 & 0 & 0 & 0 & .2 \\ -.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (45)$$

Lower left 9 × 10

$$\begin{pmatrix} -.4 & 0 & 0 & -.5 & -.2 & -.3 & .6 & -.5 & -.2 & -.3 \\ 0 & .5 & 0 & 1.2 & -.2 & 0 & -.5 & 0 & 0 & 0 \\ 0 & -.3 & 0 & -.4 & .6 & 0 & -.5 & 0 & 0 & 0 \\ 0 & -1.1 & 0 & -.4 & -.2 & 0 & -.5 & 0 & 0 & 0 \\ 0 & .9 & 0 & -.4 & -.2 & 0 & 1.5 & 0 & 0 & 0 \\ 0 & 0 & .4 & .9 & 0 & -.1 & -.4 & 0 & 0 & 0 \\ 0 & 0 & -.8 & -.3 & 0 & -.1 & -.4 & 0 & 0 & 0 \\ 0 & 0 & -.4 & -.3 & 0 & .3 & -.4 & 0 & 0 & 0 \\ 0 & 0 & .8 & -.3 & 0 & -.1 & 1.2 & 0 & 0 & 0 \end{pmatrix} \tag{46}$$

Lower right 9 × 9

$$\begin{pmatrix} 1.6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.2 & -.2 & 0 & -.5 & 0 & 0 & 0 & 0 \\ 0 & -.4 & 1.6 & 0 & -.5 & 0 & 0 & 0 & 0 \\ 0 & -.4 & -.2 & 1.0 & -.5 & 0 & 0 & 0 & 0 \\ 0 & -.4 & -.2 & 0 & 2.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.9 & 0 & -.1 & -.4 \\ 0 & 0 & 0 & 0 & 0 & -.3 & 1.0 & -.1 & -.4 \\ 0 & 0 & 0 & 0 & 0 & -.3 & 0 & 1.3 & -.4 \\ 0 & 0 & 0 & 0 & 0 & -.3 & 0 & -.1 & 2.2 \end{pmatrix} \tag{47}$$

The solution is

$$\begin{aligned} \hat{\mathbf{a}} &= (.563,\ -.437,\ -.126)'. \\ \hat{\mathbf{b}} &= (5.140,\ 4.218,\ 3.712,\ 2.967)'. \\ \hat{\boldsymbol{\gamma}} &= \begin{pmatrix} .104 & .163 & -.096 & -.170 \\ .421 & -.226 & .219 & -.414 \\ -.524 & .063 & -.122 & .584 \end{pmatrix}. \end{aligned}$$

The $\hat{\gamma}_{ij}$ sum to 0 by rows and columns.

When we employ the model with $Var(\mathbf{a}_*) = 2\mathbf{I}$ and $Var(\boldsymbol{\gamma}_*) = 8\mathbf{I}$, the coefficient matrix is in (17.48) ... (17.50) and the right hand side is (3.05, 1.8, 1.5, 3.15, .85, .8, 1.55, 1.5, .55, .65, .35, 1.05, .3, 0, .45, .6, 0, .15, .75)′.

Upper left 7 × 7

$$\begin{pmatrix} 1.1 & 0 & 0 & .25 & .1 & .15 & .1 \\ & 1.05 & 0 & .2 & .1 & 0 & .25 \\ & & .9 & .15 & 0 & .05 & .2 \\ & & & .6 & 0 & 0 & 0 \\ & & & & .2 & 0 & 0 \\ & & & & & .2 & 0 \\ & & & & & & .55 \end{pmatrix} \tag{48}$$

Lower left 12 × 7 and (upper right 7 × 12)′

$$\begin{pmatrix} .25 & 0 & 0 & .25 & 0 & 0 & 0 \\ .1 & 0 & 0 & 0 & .1 & 0 & 0 \\ .15 & 0 & 0 & 0 & 0 & .15 & 0 \\ .1 & 0 & 0 & 0 & 0 & 0 & .1 \\ 0 & .2 & 0 & .2 & 0 & 0 & 0 \\ 0 & .1 & 0 & 0 & .1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .25 & 0 & 0 & 0 & 0 & .25 \\ 0 & 0 & .15 & .15 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .05 & 0 & 0 & .05 & 0 \\ 0 & 0 & .2 & 0 & 0 & 0 & .2 \end{pmatrix} \qquad (49)$$

Lower right 12 × 12

$$= \text{diag}\ (.375,\ .225,\ .275,\ .225,\ .325,\ .225,\ .125,\ .375,\ .275,\ .125,\ .175,\ .325). \qquad (50)$$

The solution is

$$\begin{aligned} \hat{\mathbf{a}} &= (.282,\ -.219,\ -.063)',\ \text{different from above.} \\ \hat{\mathbf{b}} &= (5.140,\ 4.218,\ 3.712,\ 2.967)',\ \ \text{the same as before.} \\ \hat{\boldsymbol{\gamma}} &= \begin{pmatrix} .385 & .444 & .185 & .112 \\ .202 & -.444 & 0 & -.632 \\ -.588 & 0 & -.185 & .521 \end{pmatrix}, \end{aligned}$$

different from above. Now the $\hat{\gamma}$ sum to 0 by columns, but not by rows. This sum is

$$\sigma^2_{*\gamma}\hat{a}_i/\sigma^2_{*a} = 4\hat{a}_i.$$

As we should expect, the predictions of subclass means are identical in the two solutions. These are

$$\begin{pmatrix} 5.807 & 4.945 & 4.179 & 3.360 \\ 5.124 & 3.555 & 3.493 & 2.116 \\ 4.490 & 4.155 & 3.463 & 3.425 \end{pmatrix}.$$

These are all unbiased, including missing subclasses. This is in contrast to the situation in which both rows and columns are fixed. Note, however, that we should not predict μ_{ij} except for $j = 1, 2, 3, 4$. We could predict μ_{ij} (j=1,2,3,4) for $i > 3$, that is for rows not in the sample. BLUP would be $\hat{b}_j$. Remember, that $\hat{b}_j$ is BLUP of b_j + the mean of all $\mathbf{a}_i$ in the infinite population, and $\mathbf{a}_i$ is BLUP of $\mathbf{a}_i$ minus the mean of all $\mathbf{a}_i$ in the population.

We could, if we choose, obtain biased estimators and predictors by using some prior on the squares and products of $\mathbf{b}$, say

$$\begin{pmatrix} 1 & & \frac{-1}{3} \\ & \ddots & \\ \frac{-1}{3} & & 1 \end{pmatrix} \sigma^2_b,$$

where σ_b^2 is a pseudo-variance.

Suppose we want to estimate the variances. In that case the model with

$$Var(\mathbf{a}_*) = \mathbf{I}\sigma_{*a}^2 \text{ and } Var(\boldsymbol{\gamma}_*) = \mathbf{I}\sigma_{*\gamma}^2$$

is obviously easier to deal with than the pedagogically more logical model with $Var(\boldsymbol{\gamma})$ not a diagonal matrix. If we want to use that model, we can estimate σ_{*a}^2 and $\sigma_{*\gamma}^2$ and then by simple algebra convert those to estimates of σ_a^2 and σ_γ^2 .