Chapter 10
Quadratic Estimation of Variances

C. R. Henderson

1984 - Guelph

Estimation of $\mathbf{G}$ and $\mathbf{R}$ is a crucial part of estimation and tests of significance of estimable functions of $\boldsymbol{\beta}$ and of prediction of $\mathbf{u}$. Estimators and predictors with known desirable properties exist when $\mathbf{G}$ and $\mathbf{R}$ are known, but realistically that is never the case. Consequently we need to have good estimates of them if we are to obtain estimators and predictors that approach BLUE and BLUP. This chapter is concerned with a particular class of estimators namely translation invariant, unbiased, quadratic estimators. First a model will be described that appears to include all linear models proposed for animal breeding problems.

1 A General Model For Variances And Covariances

The model with which we have been concerned is

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu} + \mathbf{e}. \\ Var(\mathbf{u}) &= \mathbf{G}, Var(\mathbf{e}) = \mathbf{R}, Cov(\mathbf{u},\ \mathbf{e}') = \mathbf{0}. \end{aligned}$$

The dimensions of vectors and matrices are

$\mathbf{y} : n \times 1,\ \mathbf{X} : n \times p,\ \boldsymbol{\beta} : p \times 1,\ \mathbf{Z} : n \times q,\ \mathbf{u} : q \times 1,\ \mathbf{e} : n \times 1,\ \mathbf{G} : q \times q,$ and $\mathbf{R} : n \times n.$

Now we characterize $\mathbf{u}$ and $\mathbf{e}$ in more detail. Let

$$\mathbf{Zu} = \sum\nolimits_{i=1}^{b} \mathbf{Z}_i \mathbf{u}_i. \tag{1}$$

$\mathbf{Z}_i$ has dimension $n \times q_i$, and $\mathbf{u}_i$ is $q_i \times 1$.

$$\begin{aligned} \sum\nolimits_{i=1}^{b} q_i &= q, \\ Var(\mathbf{u}_i) &= \mathbf{G}_{ii} g_{ii}. \quad (2) \\ Cov(\mathbf{u}_i, \mathbf{u}_j') &= \mathbf{G}_{ij} g_{ij}. \quad (3) \end{aligned}$$

g_{ii} represents a variance and g_{ij} a covariance. Let

$$\begin{aligned} \mathbf{e}' &= (\mathbf{e}_1'\ \mathbf{e}_2' \ldots \mathbf{e}_c'). \\ Var(\mathbf{e}_i) &= \mathbf{R}_{ii} r_{ii}. \quad (4) \\ Cov(\mathbf{e}_i, \mathbf{e}_j') &= \mathbf{R}_{ij} r_{ij}. \quad (5) \end{aligned}$$

r_{ii} and r_{ij} represent variances and covariances respectively. With this model $Var(\mathbf{y})$ is

$$\mathbf{V} = \sum_{i=1}^{b} \sum_{j=1}^{b} \mathbf{Z}_i \mathbf{G}_{ij} \mathbf{Z}_j' \mathbf{g}_{ij} + \mathbf{R}, \tag{6}$$

$$Var(\mathbf{u}) = \mathbf{G} = \begin{pmatrix} \mathbf{G}_{11}g_{11} & \mathbf{G}_{12}g_{12} & \cdots & \mathbf{G}_{1b}g_{1b} \\ \mathbf{G}_{12}'g_{12} & \mathbf{G}_{22}g_{22} & \cdots & \mathbf{G}_{2b}g_{2b} \\ \vdots & \vdots & & \vdots \\ \mathbf{G}_{1b}'g_{1b} & \mathbf{G}_{2b}'g_{22} & \cdots & \mathbf{G}_{bb}g_{bb} \end{pmatrix}, \tag{7}$$

and

$$Var(\mathbf{e}) = \mathbf{R} = \begin{pmatrix} \mathbf{R}_{11}r_{11} & \mathbf{R}_{12}r_{12} & \cdots & \mathbf{R}_{1c}r_{1c} \\ \mathbf{R}_{12}'r_{12} & \mathbf{R}_{22}r_{22} & \cdots & \mathbf{R}_{2c}r_{2c} \\ \vdots & \vdots & & \vdots \\ \mathbf{R}_{1c}'r_{1c} & \mathbf{R}_{2c}'r_{2c} & \cdots & \mathbf{R}_{cc}r_{cc} \end{pmatrix}. \tag{8}$$

We illustrate this general model with two different specific models, first a traditional mixed model for variance components estimation, and second a two trait model with missing data. Suppose we have a random sire by fixed treatment model with interaction. The numbers of observations per subclass are

	Sires		
Treatment	1	2	3
1	2	1	2
2	1	3	0

Let the scalar model be

$$y_{ijk} = \mu + t_i + s_j + (ts)_{ij} + e_{ijk}.$$

The s_j have common variance, σ_s^2, and are uncorrelated. The $(ts)_{ij}$ have common variance, σ_{st}^2, and are uncorrelated. The s_j and $(ts)_{ij}$ are uncorrelated. The e_{ijk} have common variance, σ_e^2, and are uncorrelated. The corresponding vector model, for $b = 2$, is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{u}_1 + \mathbf{Z}_2\mathbf{u}_2 + \mathbf{e}.$$

$$\mathbf{X}\boldsymbol{\beta} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ t_1 \\ t_2 \end{pmatrix}, \quad \mathbf{Z}_1\mathbf{u}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix},$$

$$\mathbf{Z}_2\mathbf{u}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ts_{11} \\ ts_{12} \\ ts_{13} \\ ts_{21} \\ ts_{22} \end{pmatrix},$$

and

$$\mathbf{G}_{11}g_{11} = \mathbf{I}_3\ \sigma_s^2,\ \ \mathbf{G}_{22}g_{22} = \mathbf{I}_5\ \sigma_{ts}^2.$$

$\mathbf{G}_{12}g_{12}$ does not exist, $c = 1$, and $\mathbf{R}_{11}r_{11} = \mathbf{I}_9\ \sigma_e^2$.

For a two trait model suppose that we have the following data on progeny of two related sires

Sire	Progeny	Trait 1	Trait 2
1	1	X	X
1	2	X	X
1	3	X	0
2	4	X	X
2	5	X	0

X represents a record and 0 represents a missing record. Let us assume an additive genetic sire model. Order the records by columns, that is animals within traits. Let $\mathbf{u}_1, \mathbf{u}_2$ represent sire values for traits 1 and 2 respectively. These are breeding values divided by 2. Let $\mathbf{e}_1, \mathbf{e}_2$ represent "errors" for traits 1 and 2 respectively. Sire 2 is a son of sire 1, both non-inbred.

$$n = 8,\ \ q_1 = 2,\ \ q_2 = 2.$$

$$\mathbf{Z}_1\mathbf{u}_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{u}_1,\ \ \mathbf{Z}_2\mathbf{u}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{u}_2,$$

$$\mathbf{G}_{11}g_{11} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} g_{11}^*,\ \ \mathbf{G}_{12}g_{12} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} g_{12}^*,$$

$$\mathbf{G}_{22}g_{22} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} g^*_{22},$$

where

$$\begin{pmatrix} g^*_{11} & g^*_{12} \\ g^*_{12} & g^*_{22} \end{pmatrix}$$

is the additive genetic variance-covariance matrix divided by 4. Also,

$$\mathbf{R}_{11}r_{11} = \mathbf{I}_5 r^*_{11},\ \mathbf{R}_{22}r_{22} = \mathbf{I}_3 r^*_{22},\ \mathbf{R}_{12}r_{12} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} r^*_{12},$$

where

$$\begin{pmatrix} r^*_{11} & r^*_{12} \\ r^*_{12} & r^*_{22} \end{pmatrix}$$

is the error variance-covariance matrix for the 2 traits. Then $h^2_1 = 4\ g^*_{11}/(g^*_{11} + r^*_{11})$. Genetic correlation between traits 1 and 2 is $g^*_{12}/(g^*_{11}\ \ g^*_{22})^{1/2}$.

Another method for writing $\mathbf{G}$ and $\mathbf{R}$ is the following

$$\mathbf{G} = \mathbf{G}^*_{11}g_{11} + \mathbf{G}^*_{12}g_{12}\ + ... +\ \mathbf{G}^*_{bb}g_{bb}, \tag{9}$$

where

$$\mathbf{G}^*_{11} = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},\ \mathbf{G}^*_{12} = \begin{pmatrix} \mathbf{0} & \mathbf{G}_{12} & \mathbf{0} \\ \mathbf{G}'_{12} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},\ ...,\ \mathbf{G}^*_{bb} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{bb} \end{pmatrix}.$$

Every $\mathbf{G}^*_{ij}$ has order, q, and

$$\mathbf{R} = \mathbf{R}^*_{11}r_{11} + \mathbf{R}^*_{12}r_{12}\ + ... +\ \mathbf{R}^*_{cc}r_{cc}, \tag{10}$$

where

$$\mathbf{R}^*_{11} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},\ \mathbf{R}^*_{12} = \begin{pmatrix} \mathbf{0} & \mathbf{R}_{12} & \mathbf{0} \\ \mathbf{R}'_{12} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},\ \text{etc.}$$

and every $\mathbf{R}^*_{ij}$ has order, n.

2 Quadratic Estimators

Many methods commonly used for estimation of variances and covariances are quadratic, unbiased, and translation invariant. They include among others, ANOVA estimators for balanced designs, unweighted means and weighted squares of means estimators for filled subclass designs, Henderson's methods 1, 2 and 3 for unequal numbers, MIVQUE, and MINQUE. Searle (1968, 1971a) describes in detail some of these methods.

A quadratic estimator is defined as $\mathbf{y}'\mathbf{Q}\mathbf{y}$ where for convenience $\mathbf{Q}$ can be specified as a symmetric matrix. If we derive a quadratic with a non-symmetric matrix, say $\mathbf{P}$, we can convert this to a quadratic with a symmetric matrix by the following identity.

$$\begin{aligned} \mathbf{y}'\mathbf{Q}\mathbf{y} &= (\mathbf{y}'\mathbf{P}\mathbf{y} + \mathbf{y}'\mathbf{P}'\mathbf{y})/2 \\ \text{where} \quad \mathbf{Q} &= (\mathbf{P} + \mathbf{P}')/2. \end{aligned}$$

A translation invariant quadratic estimator satisfies

$$\begin{aligned} \mathbf{y}'\mathbf{Q}\mathbf{y} &= (\mathbf{y} + \mathbf{X}\mathbf{k})'\mathbf{Q}(\mathbf{y} + \mathbf{X}\mathbf{k}) \text{ for any vector, } \mathbf{k}. \\ \mathbf{y}'\mathbf{Q}\mathbf{y} &= \mathbf{y}'\mathbf{Q}\mathbf{y} + 2\mathbf{y}'\mathbf{Q}\mathbf{X}\mathbf{k} + \mathbf{k}'\mathbf{X}'\mathbf{Q}\mathbf{X}\mathbf{k}. \end{aligned}$$

From this it is apparent that for equality it is required that

$$\mathbf{Q}\mathbf{X} = \mathbf{0}. \qquad (11)$$

For unbiasedness we examine the expectation of $\mathbf{y}'\mathbf{Q}\mathbf{y}$ intended to estimate, say g_{gh}.

$$E(\mathbf{y}'\mathbf{Q}\mathbf{y}) = \boldsymbol{\beta}'\mathbf{X}'\mathbf{Q}\mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{b}\sum_{j=1}^{b} tr(\mathbf{Q}\mathbf{Z}_i\mathbf{G}^*_{ij}\mathbf{Z}'_j)g_{ij}$$
$$+ \sum_{i=1}^{c}\sum_{j=i}^{c} tr(\mathbf{Q}\mathbf{R}^*_{ij})r_{ij}.$$

We require that the expectation equals g_{gh}. Now if the estimator is translation invariant, the first term in the expectation is 0 because $\mathbf{Q}\mathbf{X} = \mathbf{0}$. Further requirements are that

$$\begin{aligned} tr(\mathbf{Q}\mathbf{Z}\mathbf{G}^*_{ij}\mathbf{Z}') &= 1 \text{ if } i = g \text{ and } j = h \\ &= 0, \text{ otherwise and} \\ tr(\mathbf{Q}\mathbf{R}^*_{ij}) &= 0 \text{ for all } i, j. \end{aligned}$$

3 Variances Of Estimators

Searle(1958) showed that the variance of a quadratic estimator $\mathbf{y}'\mathbf{Q}\mathbf{y}$, that is unbiased and translation invariant is

$$2\ tr(\mathbf{Q}\mathbf{V}\mathbf{Q}\mathbf{V}), \qquad (12)$$

and the covariance between two estimators $\mathbf{y}'\mathbf{Q}_1\mathbf{y}$ and $\mathbf{y}'\mathbf{Q}_2\mathbf{y}$ is

$$2\ tr(\mathbf{Q}_1\mathbf{V}\mathbf{Q}_2\mathbf{V}) \qquad (13)$$

where $\mathbf{y}$ is multivariate normal, and $\mathbf{V}$ is defined in (6). Then it is seen that (12) and (13) are quadratics in the g_{ij} and r_{ij}, the unknown parameters that are estimated. Consequently the results are in terms of these parameters, or they can be evaluated numerically for assumed values of $\mathbf{g}$ and $\mathbf{r}$. In the latter case it is well to evaluate $\mathbf{V}$ numerically for assumed $\mathbf{g}$ and $\mathbf{r}$ and then to proceed with the methods of (12) and (13).

4 Solutions Not In The Parameter Space

Unbiased estimators of variances and covariances with only one exception have positive probabilities of solutions not in the parameter space. The one exception is estimation of error variance from least squares or mixed model residuals. Otherwise estimates of variances can be negative, and functions of estimates of covariances and variances can result in estimated correlations outside the permitted range -1 to 1. In Chapter 12 the condition required for an estimated variance-covariance matrix to be in the parameter space is that there be no negative eigenvalues.

An inevitable price to pay for quadratic unbiasedness is non-zero probability that the estimated variance-covariance matrix will not fall in the parameter space. All such estimates are obtained by solving a set of linear equations obtained by equating a set of quadratics to their expectations. We could, if we knew how, impose side conditions on these equations that would force the solution into the parameter space. Having done this the solution would no longer yield unbiased estimators. What should be done in practice? It is sometimes suggested that we estimate unbiasedly, report all such results and then ultimately we can combine these into a better set of estimates that do fall in the parameter space. On the other hand, if the purpose of estimation is to provide $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ for immediate use in mixed model estimation and prediction, it would be very foolish to use estimates not in the parameter space. For example, suppose that in a sire evaluation situation we estimate σ_e^2/σ_s^2 to be negative and use this in mixed model equations. This would result in predicting a sire with a small number of progeny to be more different from zero than the adjusted progeny mean if $-\hat{\sigma}_e^2/\hat{\sigma}_s^2$ is less than the corresponding diagonal element of the sire. If the absolute value of this ratio is greater than the diagonal element, the sign of $\hat{s}_i$ is reversed as compared to the adjusted progeny mean. These consequences are of course contrary to selection index and BLUP principles.

Another problem in estimation should be recognized. The fact that estimated variance-covariance matrices fall in the parameter space does not necessarily imply that functions of these have that same property. For example, in an additive genetic sire model it is often assumed that $4\hat{\sigma}_s^2/(\hat{\sigma}_s^2 + \hat{\sigma}_e^2)$ is an estimate of h^2. But it is entirely possible that this computed function is greater than one even when $\hat{\sigma}_s^2$ and $\hat{\sigma}_e^2$ are both greater than 0. Of course if $\hat{\sigma}_s^2 < 0$ and $\hat{\sigma}_e^2 > 0$, the estimate of h^2 would be negative. Side conditions to solution of $\hat{\sigma}_s^2$ and $\hat{\sigma}_e^2$ that will insure that $\hat{\sigma}_s^2$, $\hat{\sigma}_e^2$, and h^2 (computed as above) fall in the parameter space are

$$\hat{\sigma}_s^2 > 0, \ \hat{\sigma}_e^2 > 0, \text{ and } \hat{\sigma}_s^2/\hat{\sigma}_e^2 < 1/3.$$

Another point that should be made is that even though $\hat{\sigma}_s^2$ and $\hat{\sigma}_e^2$ are unbiased, $\hat{\sigma}_s^2/\hat{\sigma}_e^2$ is a biased estimator of σ_s^2/σ_e^2, and $4\hat{\sigma}_s^2/(\hat{\sigma}_s^2 + \hat{\sigma}_e^2)$ is a biased estimator of h^2.

5 Form Of Quadratics

Except for MIVQUE and MINQUE most quadratic estimators in models with all $g_{ij} = 0$ for $i \neq j$ and with $\mathbf{R} = \mathbf{I}\sigma_e^2$ can be expressed as linear functions of $\mathbf{y}'\mathbf{y}$ and of reductions in sums of squares that will now be defined.

Let OLS equations in $\boldsymbol{\beta}, \mathbf{u}$ be written as

$$\mathbf{W}'\mathbf{W}\boldsymbol{\alpha}^o = \mathbf{W}'\mathbf{y} \tag{14}$$

where $\mathbf{W} = (\mathbf{X} \ \ \mathbf{Z})$ and

$$\boldsymbol{\alpha}^o = \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix}.$$

Then reduction under the full model is

$$(\boldsymbol{\alpha}^o)'\mathbf{W}'\mathbf{y} \tag{15}$$

Partition with possible re-ordering of columns

$$\mathbf{W} = (\mathbf{W}_1 \ \ \mathbf{W}_2) \tag{16}$$

and correspondingly

$$\boldsymbol{\alpha}^o = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}.$$

$\boldsymbol{\alpha}_1$ should always contain $\boldsymbol{\beta}$ and from 0 to $b-1$ of the $\mathbf{u}_i$. Solve for $\boldsymbol{\alpha}_1^*$ in

$$\mathbf{W}_1'\mathbf{W}_1\boldsymbol{\alpha}_1^* = \mathbf{W}_1'\mathbf{y}. \tag{17}$$

Then reduction under the reduced model is

$$(\boldsymbol{\alpha}_1^*)'\mathbf{W}_1'\mathbf{y}. \tag{18}$$

6 Expectations of Quadratics

Let us derive the expectations of these ANOVA type quadratics.

$$E(\mathbf{y}'\mathbf{y}) = tr \ Var(\mathbf{y}) + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \tag{19}$$

$$= \sum_{i=1}^{b} tr(\mathbf{Z}_i\mathbf{G}_{ii}\mathbf{Z}_i')g_{ii} + n\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{20}$$

In traditional variance components models every $\mathbf{G}_{ii} = \mathbf{I}$. Then

$$E(\mathbf{y}'\mathbf{y}) = \sum_{i=1}^{b} n\ g_{ii} + n\ \sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{21}$$

It can be seen that (15) and (18) are both quadratics in $\mathbf{W}'\mathbf{y}$. Consequently we use $Var(\mathbf{W}'\mathbf{y})$ in deriving expectations. The random part of $\mathbf{W}'\mathbf{y}$ is

$$\sum_i \mathbf{W}'\mathbf{Z}_i\mathbf{u}_i + \mathbf{W}'\mathbf{e}. \tag{22}$$

The matrix of the quadratic in $\mathbf{W}'\mathbf{y}$ for the reduction under the full model is $(\mathbf{W}'\mathbf{W})^-$. Therefore the expectation is

$$\sum_{i=1}^{b} tr(\mathbf{W}'\mathbf{W})^-\mathbf{W}'\mathbf{Z}_i\mathbf{G}_{ii}\mathbf{Z}_l'\mathbf{W}g_{ii}\ +\ \text{rank}\ (\mathbf{W})\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}(\mathbf{W}'\mathbf{W})^-\mathbf{W}'\mathbf{X}\boldsymbol{\beta}. \tag{23}$$

When all $\mathbf{G}_{ii} = \mathbf{I}$, (23) reduces to

$$\sum_{i=1}^{b} n\ g_{ii} + r(\mathbf{W})\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{24}$$

For the reduction due to $\boldsymbol{\alpha}_1$, the matrix of the quadratic in $\mathbf{W}'\mathbf{y}$ is

$$\begin{pmatrix} (\mathbf{W}_1'\mathbf{W}_1)^- & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Then the expectation of the reduction is

$$\sum_{i=1}^{h} tr(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{Z}_i\mathbf{G}_{ii}\mathbf{Z}_i'\mathbf{W}_1g_{ii}\ +\ \text{rank}\ (\mathbf{W}_1)\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{X}\boldsymbol{\beta}. \tag{25}$$

When all $\mathbf{G}_{ii} = \mathbf{I}$, (25) and when $\mathbf{X}$ is included in $\mathbf{W}_1$ simplifies to

$$\sum_i n\ g_{ii} + \sum_j tr(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{Z}_j\mathbf{Z}_j'\mathbf{W}_1g_{jj}\ + \text{rank}\ (\mathbf{W}_1)\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{26}$$

where i refers to $\mathbf{u}_i$ included in $\boldsymbol{\alpha}_1$, and j refers to $\mathbf{u}_j$ not included in $\boldsymbol{\alpha}_1$. If $\mathbf{Z}_j$ is a linear function of $\mathbf{W}_1$, the coefficient of g_{jj} is n also.

7 Quadratics in $\hat{\mathbf{u}}$ and $\hat{\mathbf{e}}$

MIVQUE computations can be formulated as we shall see in Chapter 11 as quadratics in $\hat{\mathbf{u}}$ and $\hat{\mathbf{e}}$, BLUP of $\mathbf{u}$ and $\mathbf{e}$ when $\mathbf{g} = \tilde{\mathbf{g}}$ and $\mathbf{r} = \tilde{\mathbf{r}}$. The mixed model equations are

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \end{pmatrix}. \tag{27}$$

Let some quadratic in $\hat{\mathbf{u}}$ be $\hat{\mathbf{u}}'\mathbf{Q}\hat{\mathbf{u}}$. The expectation of this is

$$tr\mathbf{Q}\ Var(\hat{\mathbf{u}}). \tag{28}$$

To find $Var(\hat{\mathbf{u}})$, define a g-inverse of the coefficient matrix of (27) as

$$\begin{pmatrix} \mathbf{C}_{00} & \mathbf{C}_{01} \\ \mathbf{C}_{10} & \mathbf{C}_{11}. \end{pmatrix} \equiv \begin{pmatrix} \mathbf{C}_0 \\ \mathbf{C}_1 \end{pmatrix} \equiv \mathbf{C}. \tag{29}$$

$\hat{\mathbf{u}} = \mathbf{C}_1\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{y}$. See (16) for definition of $\mathbf{W}$. Then

$$Var(\hat{\mathbf{u}}) = \mathbf{C}_1\ [Var(\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{y})]\ \mathbf{C}_1', \tag{30}$$

and

$$\begin{aligned} Var(\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{y}) \quad &= \quad \textstyle\sum_{i=1}^{b}\sum_{j=1}^{b}\ \mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{Z}_i\mathbf{G}_{ij}\mathbf{Z}_j'\tilde{\mathbf{R}}^{-1}\mathbf{W}g_{ij} & (31) \\ &\quad +\textstyle\sum_{i=1}^{c}\sum_{j=1}^{c}\ \mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{R}_{ij}^*\tilde{\mathbf{R}}^{-1}\mathbf{W}r_{ij}. & (32) \end{aligned}$$

Let some quadratic in $\hat{\mathbf{e}}$ be $\hat{\mathbf{e}}'\mathbf{Q}\hat{\mathbf{e}}$. The expectation of this is

$$tr\mathbf{Q}\ Var(\hat{\mathbf{e}}). \tag{33}$$

But $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o - \mathbf{Z}\hat{\mathbf{u}} = \mathbf{y} - \mathbf{W}\boldsymbol{\alpha}^o$, where $(\boldsymbol{\alpha}^o)' = [(\boldsymbol{\beta}^o)'\ \ \hat{\mathbf{u}}']$ and $\mathbf{W} = (\mathbf{X}\ \ \mathbf{Z})$, giving

$$\hat{\mathbf{e}} = [\mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1}]\mathbf{y}. \tag{34}$$

Therefore,

$$Var(\hat{\mathbf{e}}) = (\mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1})\ [Var(\mathbf{y})]\ (\mathbf{I} - \mathbf{WCW}'\tilde{\mathbf{R}}^{-1})', \tag{35}$$

and

$$\begin{aligned} Var(\mathbf{y}) \quad &= \quad \textstyle\sum_{i=1}^{b}\sum_{j=1}^{b}\ \mathbf{Z}_i\mathbf{G}_{ij}\mathbf{Z}_j'\mathbf{g}_{ij} & (36) \\ &\quad +\textstyle\sum_{i=1}^{c}\sum_{j=1}^{c}\ \mathbf{R}_{ij}^*r_{ij}. & (37) \end{aligned}$$

When

$$\begin{aligned} \mathbf{G} &= \tilde{\mathbf{G}}, \\ \mathbf{R} &= \tilde{\mathbf{R}}, \\ Var(\hat{\mathbf{u}}) &= \mathbf{G} - \mathbf{C}_{11}, \quad \text{and} & (38) \\ Var(\hat{\mathbf{e}}) &= \mathbf{R} - \mathbf{WCW}'. & (39) \end{aligned}$$

(38) and (39) are used for REML and ML methods to be described in Chapter 12.

8 Henderson's Method 1

We shall now present several methods that have been used extensively for estimation of variances (and in some cases with modifications for covariances). These are modelled after balanced ANOVA methods of estimation. The model for these methods is usually

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^{b} \mathbf{Z}_i\mathbf{u}_i + \mathbf{e}, \tag{40}$$

where $Var(\mathbf{u}_i) = \mathbf{I}\sigma_i^2$, $Cov(\mathbf{u}_i, \mathbf{u}_j^{'}) = \mathbf{0}$ for all $i \neq j$, and $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$. However, it is relatively easy to modify these methods to deal with

$$Var(\mathbf{u}_i) = \mathbf{G}_{ii}\sigma_i^2.$$

For example, $\mathbf{G}_{ii}$ might be $\mathbf{A}$, the numerator relationship matrix.

Method 1, Henderson(1953), requires for unbiased estimation that $\mathbf{X}' = [1...1]$. The model is usually called a random model. The following reductions in sums of squares are computed

$$\mathbf{y}'\mathbf{Z}_i(\mathbf{Z}_i\mathbf{Z}_i)^{-1}\mathbf{Z}_i^{'}\mathbf{y}(i = 1, \ldots, b), \tag{41}$$

$$(\mathbf{1}'\mathbf{y}\mathbf{y}'\mathbf{1})/n, \tag{42}$$

and

$$\mathbf{y}'\mathbf{y}. \tag{43}$$

The first b of these are simply uncorrected sums of squares for the various factors and interactions. The next one is the "correction factor", and the last is the uncorrected sum of squares of the individual observations.

Then these $b+2$ quadratics are equated to their expectations. The quadratics of (41) are easy to compute and their expectations are simple because $\mathbf{Z}_i^{'}\mathbf{Z}_i$ is always diagonal. Advantage should therefore be taken of this fact. Also one should utilize the fact that the coefficient of σ_i^2 is n, as is the coefficient of any σ_j^2 for which $\mathbf{Z}_j$ is linearly dependent upon $\mathbf{Z}_i$. That is $\mathbf{Z}_j = \mathbf{Z}_i\mathbf{K}$. For example the reduction due to sires $\times$ herds has coefficient n for σ_{sh}^2, σ_s^2, σ_h^2 in a model with random sires and herds. The coefficient of σ_e^2 in the expectation is the rank of $\mathbf{Z}_i^{'}\mathbf{Z}_i$, which is the number of elements in $\mathbf{u}_i$.

Because Method 1 is so easy, it is often tempting to use it on a model in which $\mathbf{X}' \neq (1...1)$, but to pretend that one or more fixed factors is random. This leads to biased estimators, but the bias can be evaluated in terms of unknown $\boldsymbol{\beta}\boldsymbol{\beta}'$. In balanced designs no bias results from using this method.

We illustrate Method 1 with a treatment $\times$ sire design in which treatments are regarded as random. The data are arranged as follows.

Number of Observations					
	Sires				
Treatment	1	2	3	4	Sums
1	8	3	2	5	18
2	7	4	1	0	12
3	6	2	0	1	9
Sums	21	9	3	6	39

Sums of Observations					
	Sires				
Treatment	1	2	3	4	Sums
1	54	21	13	25	113
2	55	33	8	0	96
3	44	17	0	9	70
Sums	153	71	21	34	279

$\mathbf{y}'\mathbf{y} = 2049$.

The ordinary least squares equations for these data are useful for envisioning Method 1 as well as some others. The coefficient matrix is in (44). The right hand side vector is (279, 113, 96, 70, 153, 71, 21, 34, 54, 21, 13, 25, 55, 33, 8, 44, 17, 9)′.

$$\left(\begin{array}{cccccccccccccccccc}
39 & 18 & 12 & 9 & 21 & 9 & 3 & 6 & 8 & 3 & 2 & 5 & 7 & 4 & 1 & 6 & 2 & 1 \\
 & 18 & 0 & 0 & 8 & 3 & 2 & 5 & 8 & 3 & 2 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & 12 & 0 & 7 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 7 & 4 & 1 & 0 & 0 & 0 \\
 & & & 9 & 6 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 2 & 1 \\
 & & & & 21 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 7 & 0 & 0 & 6 & 0 & 0 \\
 & & & & & 9 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 4 & 0 & 0 & 2 & 0 \\
 & & & & & & 3 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 & & & & & & & 6 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 1 \\
 & & & & & & & & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & 7 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & 4 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & 1 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & & 6 & 0 & 0 \\
 & & & & & & & & & & & & & & & & 2 & 0 \\
 & & & & & & & & & & & & & & & & & 1
\end{array}\right) \quad (44)$$

$$\begin{aligned}
\text{Red }(\mathbf{ts}) &= \frac{54^2}{8} + \frac{21^2}{3} \ldots + \frac{9^2}{1} = 2037.56. \\
\text{Red }(\mathbf{t}) &= \frac{113^2}{18} + \frac{96^2}{12} + \frac{70^2}{9} = 2021.83. \\
\text{Red }(\mathbf{s}) &= \frac{153^2}{21} + ... + \frac{34^2}{6} = 2014.49. \\
\text{C.F.} &= 279^2/39 = 1995.92. \\
E[\text{Red }(\mathbf{ts})] &= 10\sigma_s^2 + 39(\sigma_s^2 + \sigma_t^2 + \sigma_{ts}^2) + 39\ \mu^2.
\end{aligned}$$

For the expectations of other reductions as well as for the expectations of quadratics used in other methods including MIVQUE we need certain elements of $\mathbf{W}'\mathbf{Z}_1\mathbf{Z}_1'\mathbf{W}$, $\mathbf{W}'\mathbf{Z}_2\mathbf{Z}_2'\mathbf{W}$, and $\mathbf{W}'\mathbf{Z}_3\mathbf{Z}_3'\mathbf{W}$, where $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{Z}_3$ refer to incidence matrices for $\mathbf{t}$, $\mathbf{s}$, and $\mathbf{ts}$, respectively, and $\mathbf{W} = [\mathbf{1} \;\; \mathbf{Z}]$. The coefficients of $\mathbf{W}'\mathbf{Z}_1\mathbf{Z}_1'\mathbf{W}$ are in (45), (46), and (47).

Upper left 9×9

$$\begin{pmatrix} 549 & 324 & 144 & 81 & 282 & 120 & 48 & 99 & 144 \\ & 324 & 0 & 0 & 144 & 54 & 36 & 90 & 144 \\ & & 144 & 0 & 84 & 48 & 12 & 0 & 0 \\ & & & 81 & 54 & 18 & 0 & 9 & 0 \\ & & & & 149 & 64 & 23 & 46 & 64 \\ & & & & & 29 & 10 & 17 & 24 \\ & & & & & & 5 & 10 & 16 \\ & & & & & & & 26 & 40 \\ & & & & & & & & 64 \end{pmatrix} \tag{45}$$

Upper right 9×9 and (lower left 9×9)'

$$\begin{pmatrix} 54 & 36 & 90 & 84 & 48 & 12 & 54 & 18 & 9 \\ 54 & 36 & 90 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 84 & 48 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 54 & 18 & 9 \\ 24 & 16 & 40 & 49 & 28 & 7 & 36 & 12 & 6 \\ 9 & 6 & 15 & 28 & 16 & 4 & 12 & 4 & 2 \\ 6 & 4 & 10 & 7 & 4 & 1 & 0 & 0 & 0 \\ 15 & 10 & 25 & 0 & 0 & 0 & 6 & 2 & 1 \\ 24 & 16 & 40 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{46}$$

Lower right 9×9

$$\begin{pmatrix} 9 & 6 & 15 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 4 & 10 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 25 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 49 & 28 & 7 & 0 & 0 & 0 \\ & & & & 16 & 4 & 0 & 0 & 0 \\ & & & & & 1 & 0 & 0 & 0 \\ & & & & & & 36 & 12 & 6 \\ & & & & & & & 4 & 2 \\ & & & & & & & & 1 \end{pmatrix} \tag{47}$$

The coefficients of $\mathbf{W}'\mathbf{Z}_3\mathbf{Z}_3'\mathbf{W}$ are in (48), (49), and (50).

Upper left 9×9

$$\begin{pmatrix}
209 & 102 & 66 & 41 & 149 & 29 & 5 & 26 & 64 \\
 & 102 & 0 & 0 & 64 & 9 & 4 & 25 & 64 \\
 & & 66 & 0 & 49 & 16 & 1 & 0 & 0 \\
 & & & 41 & 36 & 4 & 0 & 1 & 0 \\
 & & & & 149 & 0 & 0 & 0 & 64 \\
 & & & & & 29 & 0 & 0 & 0 \\
 & & & & & & 5 & 0 & 0 \\
 & & & & & & & 26 & 0 \\
 & & & & & & & & 64
\end{pmatrix} \tag{48}$$

Upper right 9×9 and (lower left 9×9)'

$$\begin{pmatrix}
9 & 4 & 25 & 49 & 16 & 1 & 36 & 4 & 1 \\
9 & 4 & 25 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 49 & 16 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 36 & 4 & 1 \\
0 & 0 & 0 & 49 & 0 & 0 & 36 & 0 & 0 \\
9 & 0 & 0 & 0 & 16 & 0 & 0 & 4 & 0 \\
0 & 4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 25 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \tag{49}$$

Lower right 9×9

$$\text{dg } (9, 4, 25, 49, 16, 1, 36, 4, 1) \tag{50}$$

The coefficients of $\mathbf{W}'\mathbf{Z}_2\mathbf{Z}_2'\mathbf{W}$ are in (51), (52), and (53).

Upper left 9×9

$$\begin{pmatrix}
567 & 231 & 186 & 150 & 441 & 81 & 9 & 36 & 168 \\
 & 102 & 70 & 59 & 168 & 27 & 6 & 30 & 64 \\
 & & 66 & 50 & 147 & 36 & 3 & 0 & 56 \\
 & & & 41 & 126 & 18 & 0 & 6 & 48 \\
 & & & & 441 & 0 & 0 & 0 & 168 \\
 & & & & & 81 & 0 & 0 & 0 \\
 & & & & & & 9 & 0 & 0 \\
 & & & & & & & 36 & 0 \\
 & & & & & & & & 64
\end{pmatrix} \tag{51}$$

Upper right 9×9 and (lower left 9×9)'

$$\begin{pmatrix} 27 & 6 & 30 & 147 & 36 & 3 & 126 & 18 & 6 \\ 9 & 4 & 25 & 56 & 12 & 2 & 48 & 6 & 5 \\ 12 & 2 & 0 & 49 & 16 & 1 & 42 & 8 & 0 \\ 6 & 0 & 5 & 42 & 8 & 0 & 36 & 4 & 1 \\ 0 & 0 & 0 & 147 & 0 & 0 & 126 & 0 & 0 \\ 27 & 0 & 0 & 0 & 36 & 0 & 0 & 18 & 0 \\ 0 & 6 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 30 & 0 & 0 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 56 & 0 & 0 & 48 & 0 & 0 \end{pmatrix} \quad (52)$$

Lower right 9×9

$$\begin{pmatrix} 9 & 0 & 0 & 0 & 12 & 0 & 0 & 6 & 0 \\ & 4 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ & & 25 & 0 & 0 & 0 & 0 & 0 & 5 \\ & & & 49 & 0 & 0 & 42 & 0 & 0 \\ & & & & 16 & 0 & 0 & 8 & 0 \\ & & & & & 1 & 0 & 0 & 0 \\ & & & & & & 36 & 0 & 0 \\ & & & & & & & 4 & 0 \\ & & & & & & & & 1 \end{pmatrix} \quad (53)$$

$$\begin{aligned} E[\text{Red } (\mathbf{t})] &= 3\sigma_e^2 + 39\sigma_t^2 + k_1(\sigma_s^2 + \sigma_{ts}^2) + 39\ \mu^2. \\ k_1 &= \frac{102}{18} + \frac{66}{12} + \frac{41}{9} \ = \ 15.7222. \end{aligned}$$

The numerators above are the 2nd, 3rd, and 4th diagonals of (48) and (51). The denominators are the corresponding diagonals of the least squares coefficient matrix of (44). Also note that

$$\begin{aligned} 102 &= \sum_j n_{1j}^2 = 8^2 + 3^2 + 2^2 + 5^2, \\ 66 &= 7^2 + 4^2 + 1^2, \\ 41 &= 6^2 + 2^2 + 1^2. \\ E[\text{Red } (\mathbf{s})] &= 4\sigma_e^2 + 39\sigma_s^2 + k_2(\sigma_t^2 + \sigma_{ts}^2) + 39\ \mu^2. \\ k_2 &= \frac{149}{21} + \frac{29}{9} + \frac{5}{3} + \frac{26}{6} \ = \ 16.3175. \\ E(\text{ C.F.}) &= \sigma_e^2 + k_3\sigma_{ts}^2 + k_4\sigma_t^2 + k_5\sigma_s^2 + 39\ \mu^2. \\ k_3 &= \frac{209}{39} = 5.3590,\ k_4 = \frac{549}{39} = 14.0769,\ k_5 = \frac{567}{39} = 14.5385. \end{aligned}$$

It turns out that

$$\begin{aligned}\hat{\sigma}_e^2 &= [\mathbf{y}'\mathbf{y} - \text{Red } (\mathbf{ts})]/(39-10) \\ &= (2049 - 2037.56)/29 = .3945.\end{aligned}$$

$$E\begin{pmatrix}\hat{\sigma}_e^2 \\ \text{R}(\mathbf{ts}) \\ \text{R}(\mathbf{s}) \\ \text{R}(\mathbf{t}) \\ \text{CF}\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 10 & 39 & 39 & 39 & 1 \\ 4 & 16.3175 & 39 & 16.3175 & 1 \\ 3 & 15.7222 & 15.7222 & 39 & 1 \\ 1 & 5.3590 & 14.5385 & 14.0769 & 1 \end{pmatrix} \begin{pmatrix}\sigma_e^2 \\ \sigma_{ts}^2 \\ \sigma_s^2 \\ \sigma_t^2 \\ 39\mu^2\end{pmatrix}.$$

$$\begin{pmatrix}\hat{\sigma}_e^2 \\ \hat{\sigma}_{ts}^2 \\ \hat{\sigma}_s^2 \\ \hat{\sigma}_t^2 \\ 39\hat{\mu}^2\end{pmatrix} =$$

$$\begin{pmatrix} 1. & 0 & 0 & 0 & 0 \\ -.31433 & .07302 & -.06979 & -.06675 & .06352 \\ .01361 & -.03006 & .06979 & .02379 & -.06352 \\ .04981 & -.02894 & .02571 & .06675 & -.06352 \\ -.21453 & .45306 & -1.00251 & -.92775 & 2.47720 \end{pmatrix} \begin{pmatrix} .3945 \\ 2037.56 \\ 2014.49 \\ 2021.83 \\ 1995.92 \end{pmatrix}$$

= [.3945, -.1088, .6660, 1.0216, 1972.05]'.

The 5×5 matrix just above is the inverse of the expectation matrix.

What if $\mathbf{t}$ is fixed but we estimate by Method 1 nevertheless? We can evaluate the bias in $\hat{\sigma}_{ts}^2$ and $\hat{\sigma}_s^2$ by noting that

$$\hat{\sigma}_{ts}^2 = \mathbf{y}'\mathbf{W}\mathbf{Q}_1\mathbf{W}'\mathbf{y} - .31433\ \hat{\sigma}_e^2$$

where $\mathbf{Q}_1$ is a matrix formed from these elements of the inverse just above, (.07302, -.06979, -.06675, 06352) and the matrices of quadratics in right hand sides representing Red (ts), Red (s), Red (t), C.F.

$\mathbf{Q}_1$ is dg [.0016, -.0037, -.0056, -.0074, -.0033, -0.0078, -.0233, -.0116, .0091, .0243, .0365, .0146, .0104, .0183, .0730, .0122, .0365, .0730]. dg refers to the diagonal elements of a matrix. Then the contribution of $\mathbf{tt}'$ to the expectation of $\hat{\sigma}_{ts}^2$ is

$$tr(\mathbf{Z}_1'\mathbf{W}\mathbf{Q}_1\mathbf{W}'\mathbf{Z}_1)\ (\mathbf{tt}')$$

where $\mathbf{Z}_1$ is the incidence matrix for $\mathbf{t}$ and $\mathbf{W} = (\mathbf{1}\ \ \mathbf{Z})$.

This turns out to be

$$tr \begin{pmatrix} -.0257 & .0261 & -.0004 \\ & -.0004 & -.0257 \\ & & .0261 \end{pmatrix} \mathbf{tt}',$$

that is, $-.0257\ t_1^2 + 2(.0261)\ t_1\ t_2\ - 2(.0004)\ t_1 t_3\ - .0004\ t_2^2\ - 2(.0257)\ t_2 t_3\ +\ .0261\ t_2^2$. This is the bias due to regarding $\mathbf{t}$ as random. Similarly the quadratic in right hand sides for estimation of σ_s^2 is

dg [-.0016, .0013, .0020, .0026, .0033, .0078, .0233, .0116, -.0038, -.0100, -.0150, -.0060, -.0043, -.0075, -.0301, -.0050, -.0150, -.0301].

The bias in $\hat{\sigma}_s^2$ is

$$tr \begin{pmatrix} .0257 & -.0261 & .0004 \\ & .0004 & .0257 \\ & & -.0261 \end{pmatrix} \mathbf{tt}'.$$

This is the negative of the bias in $\hat{\sigma}_{ts}^2$.

9 Henderson's Method 3

Method 3 of Henderson(1953) can be applied to any general mixed model for variance components. Usually the model assumed is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}\ +\ \sum\ \mathbf{Z}_i \mathbf{u}_i + \mathbf{e}. \tag{54}$$

$Var(\mathbf{u}_i) = \mathbf{I}\sigma_i^2$, $Cov(\mathbf{u}_i \mathbf{u}_j^{'}) = \mathbf{0}$, $Var(\mathbf{e}) = \mathbf{I}\sigma_i^2$. In this method $b+1$ different quadratics of the following form are computed.

$$\text{Red } (\boldsymbol{\beta} \text{ with from 0 to } b \text{ included } \mathbf{u}_i) \tag{55}$$

Then σ_e^2 is estimated usually by

$$\hat{\sigma}_e^2 = [\mathbf{y}'\mathbf{y} - \text{Red } (\boldsymbol{\beta}, \mathbf{u}_1, ..., \mathbf{u}_b)]/[n - \text{rank}(\mathbf{W})] \tag{56}$$

where $\mathbf{W} = (\mathbf{X}\ \ \mathbf{Z})$, and the solution to $\boldsymbol{\beta}^o$, $\mathbf{u}^o$ is OLS.

In some cases it is easier to compute σ_e^2 by expanding the model to include all possible interactions. Then if there is no covariate, $\hat{\sigma}_e^2$ is the within "smallest subclass" mean square. Then $\hat{\sigma}_e^2$ and the $b+1$ reductions are equated to their expectations. Method 3 has the unfortunate property that there are often more than $b+1$ reductions like (55) possible. Consequently more than one Method 3 estimator exists, and in unbalanced designs the estimates will not be invariant to the choice. One would like to select the set that will give smallest sampling variance, but this is unknown. Consequently it is tempting to select the easiest subset. This usually is

Red $(\boldsymbol{\beta}, \mathbf{u}_1)$, Red $(\boldsymbol{\beta}, \mathbf{u}_2)$, ..., Red $(\boldsymbol{\beta}, \mathbf{u}_b)$, Red $(\boldsymbol{\beta})$. For example: Red $(\boldsymbol{\beta}, \mathbf{u}_2)$ is computed as follows. Solve

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_2 \\ \mathbf{Z}_2'\mathbf{X} & \mathbf{Z}_2'\mathbf{Z}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}_2^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}_2{}'\mathbf{y} \end{pmatrix}.$$

Then reduction $= (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{y} + (\mathbf{u}_2^o)'\mathbf{Z}_2'\mathbf{y}$. To find the expectation of a reduction let a g-inverse of the coefficient matrix of the i^{th} reduction, $(\mathbf{W}_i'\mathbf{W}_i)$, be $\mathbf{C}_i$. Then

$$E(i^{th} \text{ reduction}) = \text{rank } (\mathbf{C}_i)\sigma_e^2 + \sum_{j=1}^{s} tr\mathbf{C}_i\mathbf{W}_i'\mathbf{Z}_j\mathbf{Z}_j'\mathbf{W}_i\sigma_j^2 \; + \; \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{57}$$

$\mathbf{W}_i = [\mathbf{X} \;\; \mathbf{Z}_i$ for any included $\mathbf{u}_i]$, and $\mathbf{C}_i$ is the g-inverse. For example, in Red $(\boldsymbol{\beta},\ \mathbf{u}_1,\ \mathbf{u}_3)$, $\mathbf{W} = [\mathbf{X} \;\; \mathbf{Z}_1 \;\; \mathbf{Z}_3]$.

Certain of the coefficients in (57) are n. These are all σ_j^2 included in the reduction and also any σ_k^2 for which

$$\mathbf{Z}_k \; = \; \mathbf{W}_j\mathbf{L}.$$

A serious computational problem with Method 3 is that it may be impossible with existing computers to find a g-inverse of some of the $\mathbf{W}_i'\mathbf{W}_i$. Partitioned matrix methods can sometimes be used to advantage. Partition

$$\mathbf{W}_i'\mathbf{W}_i = \begin{pmatrix} \mathbf{W}_1'\mathbf{W}_1 & \mathbf{W}_1'\mathbf{W}_2 \\ \mathbf{W}_2'\mathbf{W}_1 & \mathbf{W}_2'\mathbf{W}_2 \end{pmatrix},$$

and

$$\mathbf{W}_i'\mathbf{y} = \begin{pmatrix} \mathbf{W}_1'\mathbf{y} \\ \mathbf{W}_2'\mathbf{y} \end{pmatrix}.$$

It is advantageous to have $\mathbf{W}_1'\mathbf{W}_1$ be diagonal or at least of some form that is easy to invert. Define $\boldsymbol{\beta}$ and included $\mathbf{u}_i$ as $\boldsymbol{\alpha}$ and partition as $\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix}$. Then the equations to solve are

$$\begin{pmatrix} \mathbf{W}_1'\mathbf{W}_1 & \mathbf{W}_1'\mathbf{W}_2 \\ \mathbf{W}_2'\mathbf{W}_1 & \mathbf{W}_2'\mathbf{W}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{W}_1'\mathbf{y} \\ \mathbf{W}_2'\mathbf{y} \end{pmatrix}.$$

Absorb $\boldsymbol{\alpha}_1$ by writing equations

$$\mathbf{W}_2'\mathbf{P}\mathbf{W}_2\boldsymbol{\alpha}_2 = \mathbf{W}_2'\mathbf{P}\mathbf{y} \tag{58}$$

where $\mathbf{P} = \mathbf{I} - \mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'$. Solve for $\boldsymbol{\alpha}_2$ in (58). Then

$$\text{reduction} \; = \; \mathbf{y}'\mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{y} + \boldsymbol{\alpha}_2'\mathbf{W}_2'\mathbf{P}\mathbf{y}. \tag{59}$$

To find the coefficient of σ_j^2 in the expectation of this reduction, define

$$(\mathbf{W}_2'\mathbf{P}\mathbf{W}_2)^- \ = \ \mathbf{C}.$$

The coefficient of σ_j^2 is

$$tr(\mathbf{W}_1'\mathbf{W}_1)^-\mathbf{W}_1'\mathbf{Z}_j\mathbf{Z}_j'\mathbf{W}_1 \ + \ tr\mathbf{C}\mathbf{W}_2'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{W}_2. \tag{60}$$

Of course if $\mathbf{u}_j$ is included in the reduction, the coefficient is n.

Let us illustrate Method 3 by the same example used in Method 1 except now we regard $\mathbf{t}$ as fixed. Consequently the σ_i^2 are σ_{ts}^2, σ_s^2, and we need 3 reductions, each including μ, $\mathbf{t}$. The only possible reductions are Red (μ,$\mathbf{t}$,$\mathbf{ts}$), Red (μ,$\mathbf{t}$,$\mathbf{s}$), and Red (μ,$\mathbf{t}$). Consequently in this special case Method 3 is unique. To find the first of these reductions we can simply take the last 10 rows and columns of the least squares equations. That is, dg [8, 3, 2, 5, 7, 4, 1, 6, 2, 1] $\hat{\mathbf{st}}$ = [54, 21, 13, 25, 55, 33, 8, 44, 17, 9]$'$. The resulting reduction is 2037.56 with expectation,

$$10\sigma_e^2 \ + \ 39(\sigma_{ts}^2 + \sigma_s^2) \ + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}.$$

For the reduction due to (μ, t, s) we can take the subset of OLS equations represented by rows (and columns) 2-7 inclusive. This gives equations to solve as follows.

$$\begin{pmatrix} 18 & 0 & 0 & 8 & 3 & 2 \\ & 12 & 0 & 7 & 4 & 1 \\ & & 9 & 6 & 2 & 0 \\ & & & 21 & 0 & 0 \\ & & & & 9 & 0 \\ & & & & & 3 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} 113 \\ 96 \\ 70 \\ 153 \\ 71 \\ 21 \end{pmatrix} \tag{61}$$

We can delete μ and s_4 because the above is a full rank subset of the coefficient matrix that includes μ and s_4. The inverse of the above matrix is

$$\begin{pmatrix} .1717 & .1602 & .1417 & -.1593 & -.1599 & -.1678 \\ & .3074 & .1989 & -.2203 & -.2342 & -.2093 \\ & & .2913 & -.2035 & -.2004 & -.1608 \\ & & & .2399 & .1963 & .1796 \\ & & & & .3131 & .1847 \\ & & & & & .5150 \end{pmatrix}, \tag{62}$$

and this gives a solution vector [5.448, 6.802, 6.760, 1.011, 1.547, 1.100]. The reduction is 2029.57. The coefficient of σ_s^2 in the expectation is 39 since $\mathbf{s}$ is included. To find the coefficient of σ_{ts}^2 define as $\mathbf{T}$ the submatrix of (51) formed by taking columns and rows (2-7). Then the coefficient of σ_{ts}^2 = trace [matrix (62)] $\mathbf{T}$ = 26.7638. The coefficient of σ_e^2 is 6. The reduction due to $\mathbf{t}$ and its expectation has already been done for Method 1.

Another way of formulating a reduction and corresponding expectations is to compute i^{th} reduction as follows. Solve

$$\begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_i'\mathbf{W}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \gamma_1^o \\ \gamma_2^o \\ \gamma_3^o \end{pmatrix} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix} = \mathbf{r}. \tag{63}$$

$$\begin{aligned} \mathbf{r} &= \mathbf{W}'\mathbf{y}, \text{where } \mathbf{W} = (\mathbf{X} \ \ \mathbf{Z}) \\ \text{Red} &= \mathbf{r}'\mathbf{Q}_i\mathbf{r}, \end{aligned}$$

where $\mathbf{Q}_i$ is some g-inverse of the coefficient matrix, (63). Then the coefficient of σ_e^2 in the expectation is

$$\text{rank } (\mathbf{Q}_i) = \text{rank } (\mathbf{W}_i'\mathbf{W}_i). \tag{64}$$

Coefficient of σ_j^2 is

$$tr \ \mathbf{Q}_i\mathbf{W}'\mathbf{Z}_j\mathbf{Z}_j'\mathbf{W}. \tag{65}$$

Let the entire vector of expectations be

$$E \begin{pmatrix} \hat{\sigma}_e^2 \\ \text{Red } (1) \\ \vdots \\ \text{Red } (b+1) \end{pmatrix} = \mathbf{P} \begin{pmatrix} \sigma_e^2 \\ \sigma_1^2 \\ \vdots \\ \sigma_b^2 \\ \beta'\mathbf{X}'\mathbf{X}\beta \end{pmatrix}.$$

Then the unbiased estimators are

$$\begin{pmatrix} \hat{\sigma}_e^2 \\ \hat{\sigma}_1^2 \\ \vdots \\ \hat{\sigma}_b^2 \\ \widehat{\beta'\mathbf{X}'\mathbf{X}\beta} \end{pmatrix} = \mathbf{P}^{-1} \begin{pmatrix} \hat{\sigma}_e^2 \\ \text{Red } (1) \\ \vdots \\ \text{Red } (\text{b}+1) \end{pmatrix} \tag{66}$$

provided $\mathbf{P}^{-1}$ exists. If it does not, Method 3 estimators, at least with the chosen $b+1$ reductions, do not exist. In our example

$$E \begin{pmatrix} \hat{\sigma}_e^2 \\ \text{Red } \ (ts) \\ \text{Red } \ (ts) \\ \text{Red } \ (t) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 10 & 39 & 39 & 1 \\ 6 & 26.7638 & 39 & 1 \\ 3 & 15.7222 & 15.7222 & 1 \end{pmatrix} \begin{pmatrix} \sigma_e^2 \\ \sigma_{ts}^2 \\ \sigma_s^2 \\ \beta'\mathbf{X}'\mathbf{X}\beta \end{pmatrix}.$$

$$\begin{pmatrix} \hat{\sigma}_e^2 \\ \hat{\sigma}_{ts}^2 \\ \hat{\sigma}_s^2 \\ \widehat{\beta'\mathbf{X}'\mathbf{X}\beta} \end{pmatrix} = \begin{pmatrix} .3945 \\ .5240 \\ .0331 \\ 2011.89 \end{pmatrix},$$

$$= \begin{pmatrix} 1. & 0 & 0 & 0 \\ -.32690 & .08172 & -.08172 & 0 \\ .02618 & -.03877 & .08172 & -.04296 \\ 1.72791 & -.67542 & 0 & 1.67542 \end{pmatrix} \begin{pmatrix} .3945 \\ 2037.56 \\ 2029.57 \\ 2021.83 \end{pmatrix}.$$

These are different from the Method 1 estimates.

10 A Simple Method for General $\mathbf{X}\beta$

We now present a very simple method for the general $\mathbf{X}\boldsymbol{\beta}$ model provided an easy g-inverse of $\mathbf{X}'\mathbf{X}$ can be obtained. Write the following equations.

$$\begin{pmatrix} \mathbf{Z}_1'\mathbf{P}\mathbf{Z}_1 & \mathbf{Z}_1'\mathbf{P}\mathbf{Z}_2 & \cdots & \mathbf{Z}_1'\mathbf{P}\mathbf{Z}_b \\ \mathbf{Z}_2'\mathbf{P}\mathbf{Z}_1 & \mathbf{Z}_2'\mathbf{P}\mathbf{Z}_2 & \cdots & \mathbf{Z}_2'\mathbf{P}\mathbf{Z}_b \\ \vdots & & & \\ \mathbf{Z}_b'\mathbf{P}\mathbf{Z}_1 & \mathbf{Z}_b'\mathbf{P}\mathbf{Z}_2 & \cdots & \mathbf{Z}_b'\mathbf{P}\mathbf{Z}_b \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_b \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1'\mathbf{P}\mathbf{y} \\ \mathbf{Z}_2'\mathbf{P}\mathbf{y} \\ \vdots \\ \mathbf{Z}_b'\mathbf{P}\mathbf{y} \end{pmatrix} \quad (67)$$

$\mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$. $\boldsymbol{\beta}^o$ is absorbed from the least squares equations to obtain (67). We could then compute b reductions from (67) and this would be Method 3. An easier method, however, is described next.

Let $\mathbf{D}_i$ be a diagonal matrix formed from the diagonals of $\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_i$. Then compute the following b quadratics,

$$\mathbf{y}'\mathbf{P}\mathbf{Z}_i\mathbf{D}_i^{-1}\mathbf{Z}_i'\mathbf{P}\mathbf{y}. \quad (68)$$

This computation is simple because $\mathbf{D}_i^{-1}$ is diagonal. It is simply the sum of squares of elements of $\mathbf{Z}_i'\mathbf{P}\mathbf{y}$ divided by the corresponding element of $\mathbf{D}_i$. The expectation is also easy. It is

$$q_i\sigma_e^2 \; + \; \textstyle\sum_{j=1}^s \; tr \; \mathbf{D}_i^{-1}\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\sigma_j^2. \quad (69)$$

Because $\mathbf{D}_i^{-1}$ is diagonal we need to compute only the diagonals of $\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i$ to find the last term of (69). Then as in Methods 1 and 3 we find some estimate of σ_e^2 and equate $\hat{\sigma}_e^2$ and the s quadratics of (68) to their expectations.

Let us illustrate the method with our same example, regarding $\mathbf{t}$ as fixed.

$$\mathbf{X}'\mathbf{X} \; = \; \begin{pmatrix} 39 & 18 & 12 & 9 \\ & 18 & 0 & 0 \\ & & 12 & 0 \\ & & & 9 \end{pmatrix},$$

and a g-inverse is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ & 18^{-1} & 0 & 0 \\ & & 12^{-1} & 0 \\ & & & 9^{-1} \end{pmatrix}.$$

The coefficient matrix of equations like (67) is in (70), (71) and (72) and the right hand side is (.1111, 4.6111, .4444, -5.1667, 3.7778, 2.1667, .4444, -6.3889, -1, 1, 0, -2.6667, 1.4444, 1.2222)′.

Upper left 7×7

$$\begin{pmatrix} 9.3611 & -5.0000 & -1.4722 & -2.8889 & 4.4444 & -1.3333 & -.8889 \\ & 6.7222 & -.6667 & -1.0556 & -1.3333 & 2.5 & -.3333 \\ & & 2.6944 & -.5556 & -.8889 & -.3333 & 1.7778 \\ & & & 4.5 & -2.2222 & -.8333 & -.5556 \\ & & & & 4.4444 & -1.3333 & -.8889 \\ & & & & & 2.5 & -.3333 \\ & & & & & & 1.7778 \end{pmatrix} \quad (70)$$

Upper right 7×7 and (lower left 7×7)'

$$\begin{pmatrix} -2.2222 & 2.9167 & -2.3333 & -.5833 & 2.0 & -1.3333 & -.6667 \\ -.8333 & -2.3333 & 2.6667 & -.3333 & -1.3333 & 1.5556 & -.2222 \\ -.5556 & -.5833 & -.3333 & .9167 & 0 & 0 & 0 \\ 3.6111 & 0 & 0 & 0 & -.6667 & -.2222 & .8889 \\ -2.2222 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.8333 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.5556 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (71)$$

Lower right 7×7

$$\begin{pmatrix} 3.6111 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 2.9167 & -2.3333 & -.5833 & 0 & 0 & 0 \\ & & 2.6667 & -.3333 & 0 & 0 & 0 \\ & & & .9167 & 0 & 0 & 0 \\ & & & & 2.0 & -1.3333 & -.6667 \\ & & & & & 1.5556 & -.2222 \\ & & & & & & .8889 \end{pmatrix} \quad (72)$$

The diagonals of the variance of the reduced right hand sides are needed in this method and other elements are needed for approximate MIVQUE in Chapter 11. The coefficients of σ_e^2 in this variance are in (70), ..., (72). The coefficients of σ_{ts}^2 are in (73), (74) and (75). These are computed by (Cols. 5-14 of 10.70) (same)$'$.

Upper left 7×7

$$\begin{pmatrix} 47.77 & -24.54 & -5.31 & -17.93 & 27.26 & -7.11 & -3.85 \\ & 25.75 & .39 & -1.60 & -7.11 & 8.83 & .22 \\ & & 5.66 & -.74 & -3.85 & .22 & 4.37 \\ & & & 20.27 & -16.30 & -1.94 & -.74 \\ & & & & 27.26 & -7.11 & -3.85 \\ & & & & & 8.83 & .22 \\ & & & & & & 4.37 \end{pmatrix} \quad (73)$$

Upper right 7×7 and (lower left 7×7)'

$$\begin{pmatrix} -16.30 & 14.29 & -12.83 & -1.46 & 6.22 & -4.59 & -1.63 \\ -1.94 & -12.83 & 12.67 & .17 & -4.59 & 4.25 & .35 \\ -.74 & -1.46 & .17 & 1.29 & 0 & 0 & 0 \\ 18.98 & 0 & 0 & 0 & -1.63 & .35 & 1.28 \\ -16.30 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1.94 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.74 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (74)$$

Lower right 7×7

$$\begin{pmatrix} 18.98 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 14.29 & -12.83 & -1.46 & 0 & 0 & 0 \\ & & 12.67 & .17 & 0 & 0 & 0 \\ & & & 1.29 & 0 & 0 & 0 \\ & & & & 6.22 & -4.59 & -1.63 \\ & & & & & 4.25 & .35 \\ & & & & & & 1.28 \end{pmatrix} \quad (75)$$

The coefficients of σ_s^2 are in (76), (77), and (78). These are computed by (Cols 1-4 of 10.70) (same)$'$.

Upper left 7×7

$$\begin{pmatrix} 123.14 & -76.39 & -12.81 & -33.95 & 56.00 & -22.08 & -7.67 \\ & 71.75 & 1.67 & 2.97 & -28.25 & 24.57 & 1.60 \\ & & 10.18 & .96 & -6.81 & -.14 & 6.63 \\ & & & 30.02 & -20.94 & -2.35 & -.57 \\ & & & & 27.26 & -7.11 & -3.85 \\ & & & & & 8.83 & .22 \\ & & & & & & 4.37 \end{pmatrix} \quad (76)$$

Upper right 7×7 and (lower left 7×7)'

$$\begin{pmatrix} -26.25 & 39.83 & -34.69 & -5.14 & 27.31 & -19.62 & -7.70 \\ 2.07 & -29.88 & 29.81 & .06 & -18.26 & 17.36 & .90 \\ .32 & -4.31 & .76 & 3.55 & -1.69 & 1.05 & .64 \\ 23.86 & -5.64 & 4.11 & 1.53 & -7.37 & 1.21 & 6.16 \\ -16.30 & 16.59 & -13.63 & -2.96 & 12.15 & -7.51 & -4.64 \\ -1.94 & -9.53 & 9.89 & -.36 & -5.44 & 5.85 & -.41 \\ -.74 & -2.85 & .59 & 2.26 & -.96 & .79 & .17 \end{pmatrix} \quad (77)$$

Lower right 7×7

$$\begin{pmatrix} 18.98 & -4.21 & 3.15 & 1.06 & -5.74 & .86 & 4.88 \\ & 14.29 & -12.83 & -1.46 & 8.94 & -7.52 & -1.43 \\ & & 12.67 & .17 & -8.22 & 7.26 & .96 \\ & & & 1.29 & -.72 & .26 & .46 \\ & & & & 6.22 & -4.59 & -1.63 \\ & & & & & 4.25 & .35 \\ & & & & & & 1.28 \end{pmatrix} \tag{78}$$

The reduction for **ts** is

$$\frac{3.778^2}{4.444} + \cdots + \frac{1.222^2}{.889} = 23.799.$$

The expectation is $10\ \sigma_e^2 + 35.7262\ (\sigma_{ts}^2 + \sigma_s^2)$, where 10 is the number of elements in the **ts** vector and

$$35.7262 = \frac{27.259}{4.444} + \cdots + \frac{1.284}{.889}.$$

The reduction for **s** is

$$\frac{.111^2}{9.361} + \cdots + \frac{(-5.167)^2}{4.5} = 9.170.$$

The expectation is $4\ \sigma_e^2 + 15.5383\ \sigma_{ts}^2 + 34.2770\ \sigma_s^2$, where

$$15.5383 = \frac{47.773}{9.361} + \cdots + \frac{20.265}{4.5}.$$
$$34.2770 = \frac{123.144}{9.361} + \cdots + \frac{30.019}{4.5}.$$

Thus

$$E \begin{pmatrix} \hat{\sigma}_e^2 \\ \text{Red}(\mathbf{ts}) \\ \text{Red}(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 10 & 35.7262 & 35.7262 \\ 4 & 15.5383 & 34.2770 \end{pmatrix} \begin{pmatrix} \sigma_e^2 \\ \sigma_{ts}^2 \\ \sigma_s^2 \end{pmatrix}.$$

Then

$$\begin{pmatrix} \hat{\sigma}_e^2 \\ \hat{\sigma}_{ts}^2 \\ \hat{\sigma}_s^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -.29855 & .05120 & -.05337 \\ .01864 & -.02321 & .05337 \end{pmatrix} \begin{pmatrix} .3945 \\ 23.799 \\ 9.170 \end{pmatrix} = \begin{pmatrix} .3945 \\ .6114 \\ -.0557 \end{pmatrix}.$$

11 Henderson's Method 2

Henderson's Method 2 (1953) is probably of interest from an historical viewpoint only. It has the disadvantage that random by fixed interactions and random within fixed nesting are not permitted. It is a relatively easy method, but usually no easier than the method described in Sect. 10.10, absorption of $\boldsymbol{\beta}$, and little if any easier than an approximate MIVQUE procedure described in Chapter 11.

Method 2 involves correction of the data by a least squares solution to $\boldsymbol{\beta}$ excluding μ. Then a Method 1 analysis is carried out under the assumption of a model

$$\mathbf{y} \;=\; \mathbf{1}\alpha \;+\; \sum_i \mathbf{Z}_i\mathbf{u}_i + \mathbf{e}.$$

If the solution to $\boldsymbol{\beta}^o$ is done as described below, the expectations of the Method 1 reductions are identical to those for a truly random model except for an increase in the coefficients of σ^2_e . Partition

$$\mathbf{Z} \;=\; [\mathbf{Z}_a \;\; \mathbf{Z}_b]$$

such that rank

$$(\mathbf{Z}_a) \;=\; \text{rank } (\mathbf{Z}).$$

Then partition

$$\mathbf{X} \;=\; (\mathbf{X}_a \;\; \mathbf{X}_b)$$

such that

$$\text{rank } (\mathbf{X}_a \;\; \mathbf{Z}_a) \;=\; \text{rank } (\mathbf{X} \;\; \mathbf{Z}).$$

See Henderson, Searle, and Schaeffer (1974). Solve equations (79) for $\boldsymbol{\beta}_a$.

$$\begin{pmatrix} \mathbf{X}_a'\mathbf{X}_a & \mathbf{X}_a'\mathbf{Z}_a \\ \mathbf{Z}_a'\mathbf{X}_a & \mathbf{Z}_a'\mathbf{Z}_a \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_a \\ \mathbf{u}_a \end{pmatrix} = \begin{pmatrix} \mathbf{X}_a'\mathbf{y} \\ \mathbf{Z}_a'\mathbf{y} \end{pmatrix} \tag{79}$$

Let the upper submatrix (pertaining to $\boldsymbol{\beta}_a$) of the inverse of the matrix of (79) be denoted by $\mathbf{P}$. This can be computed as

$$\mathbf{P} = [\mathbf{X}_a'\mathbf{X}_a - \mathbf{X}_a'\mathbf{Z}_a \; (\mathbf{Z}_a'\mathbf{Z}_a)^{-1}\mathbf{Z}_a'\mathbf{X}_a]^{-1}. \tag{80}$$

Now compute

$$\mathbf{1}'\mathbf{y}^* \;=\; \mathbf{1}'\mathbf{y} - \mathbf{1}'\mathrm{X}_a\boldsymbol{\beta}_a. \tag{81}$$

$$\mathbf{Z}_i'\mathbf{y}^* \;=\; \mathbf{Z}_i'\mathbf{y} - \mathbf{Z}_i'\mathrm{X}_a\boldsymbol{\beta}_a \; i = 1, \ldots, b. \tag{82}$$

Then compute the following quadratics

$$(\mathbf{1}'\mathbf{y}^*)^2/n, \text{ and} \tag{83}$$

$$(\mathbf{Z}_i'\mathbf{y}^*)'(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}(\mathbf{Z}_i'\mathbf{y}^*) \text{ for } i = 1, \ldots, b. \tag{84}$$

The expectations of these quadratics are identical to those with $\mathbf{y}$ in place of $\mathbf{y}^*$ except for an increase in the coefficient of σ_e^2 computed as follows. Increase in coefficient of σ_e^2 in expectation of (83) by

$$tr\mathbf{P}(\mathbf{X}_a'\mathbf{1}\mathbf{1}'\mathbf{X}_a)/n. \tag{85}$$

Increase in the coefficient of σ_e^2 in expectation of (84) is

$$tr\mathbf{P}(\mathbf{X}_a'\mathbf{Z}_i(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}\mathbf{Z}_i'\mathbf{X}_a). \tag{86}$$

Note that $\mathbf{X}_a'\mathbf{Z}_i(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}\mathbf{Z}_i'\mathbf{X}_a$ is the quantity that would be subtracted from $\mathbf{X}_a'\mathbf{X}_a$ if we were to "absorb" $\mathbf{u}_i$. σ_e^2 can be estimated in a number of ways but usually by the conventional residual

$$[\mathbf{y}'\mathbf{y} - (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{y} - (\mathbf{u}^o)'\mathbf{Z}'\mathbf{y}]/[n - \text{rank }(\mathbf{X}\ \ \mathbf{Z})].$$

Sampling variances for Method 2 can be computed by the same procedure as for Method l except that the variance of adjusted right hand sides of μ and $\mathbf{u}$ equations is increased by $\begin{pmatrix} \mathbf{1}'\mathbf{X}_a \\ \mathbf{Z}'\mathbf{X}_a \end{pmatrix} \mathbf{P}(\mathbf{X}_a'\mathbf{1}\quad \mathbf{X}_a'\mathbf{Z})\ \sigma_e^2$ over the unadjusted. As is true for other quadratic estimators, quadratics in the adjusted right hand sides are uncorrelated with σ_e^2, the OLS residual mean square.

We illustrate Method 2 with our same data, but now we assume that σ_{ts}^2 does not exist. This 2 way mixed model could be done just as easily by Method 3 as by Method 2, but it suffices to illustrate the latter. Delete μ and t_3 and include all 4 levels of $\mathbf{s}$. First solve for β_a in these equations.

$$\begin{pmatrix} 18 & 0 & 8 & 3 & 2 & 5 \\ & 12 & 7 & 4 & 1 & 0 \\ & & 21 & 0 & 0 & 0 \\ & & & 9 & 0 & 0 \\ & & & & 3 & 0 \\ & & & & & 6 \end{pmatrix} \begin{pmatrix} \beta_a \\ u_a \end{pmatrix} = \begin{pmatrix} 113 \\ 96 \\ 153 \\ 71 \\ 21 \\ 34 \end{pmatrix}$$

The solution is $\boldsymbol{\beta}_a$ = [-1.31154, .04287]$'$, $\mathbf{u}_a$ = (7.7106, 8.30702, 7.86007, 6.75962)$'$. The adjusted right hand sides are

$$\begin{pmatrix} 279 \\ 153 \\ 71 \\ 21 \\ 34 \end{pmatrix} - \begin{pmatrix} 18 & 12 \\ 8 & 7 \\ 3 & 4 \\ 2 & 1 \\ 5 & 0 \end{pmatrix} \begin{pmatrix} -1.31154 \\ .04287 \end{pmatrix} = \begin{pmatrix} 302.093 \\ 163.192 \\ 74.763 \\ 23.580 \\ 40.558 \end{pmatrix}.$$

Then the sum of squares of adjusted right hand sides for sires is

$$\frac{(163.192)^2}{12} + \cdots + \frac{(40.558)^2}{6} = 2348.732.$$

The adjusted C.F. is $(302.093)^2/39 = 2340.0095$. $\mathbf{P}$ is the upper 2x2 of the inverse of the coefficient matrix (79) is

$$\mathbf{P} = \begin{pmatrix} .179532 & .110888 \\ & .200842 \end{pmatrix}.$$

$$\begin{aligned}\mathbf{X}_a'\mathbf{Z}_i(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}\mathbf{Z}_i'\mathbf{X}_a &= \begin{pmatrix} 8 & 3 & 2 & 5 \\ 7 & 4 & 1 & 0 \end{pmatrix} \begin{pmatrix} 21 & 0 & 0 & 0 \\ & 9 & 0 & 0 \\ & & 3 & 0 \\ & & & 6 \end{pmatrix}^{-1} \begin{pmatrix} 8 & 7 \\ 3 & 4 \\ 2 & 1 \\ 5 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 9.547619 & 4.666667 \\ & 4.444444 \end{pmatrix}.\end{aligned}$$

The trace of this (86) is 3.642 to be added to the coefficient of σ_e^2 in E (sires S.S). The trace of $\mathbf{P}$ times the following matrix

$$\begin{pmatrix} 18 \\ 12 \end{pmatrix} \begin{pmatrix} 39 \end{pmatrix}^{-1} (18 \;\; 12) = \begin{pmatrix} 8.307692 & 5.538462 \\ & 3.692308 \end{pmatrix}.$$

gives 3.461 to be added to the coefficient of σ_e^2 in E(CF). Then

$$\begin{aligned} E(Sire\ SS) &= 7.642\ \sigma_e^2 \ + \ 39\ \sigma_s^2 \ + \ \text{a quadratic.} \\ E(C.F.) &= 4.461\ \sigma_e^2 \ + \ 14.538\ \sigma_s^2 \ + \ \text{the same quadratic.} \end{aligned}$$

Then taking some estimate of σ_e^2 one equates these expectations to the computed sums of squares.

12 An Unweighted Means ANOVA

A simple method for testing hypotheses approximately is the unweighted means analysis described in Yates (1934). This method is appropriate for the mixed model described in Section 4 provided that every subclass is filled and there are no covariates. The "smallest" subclass means are taken as the observations as in Section 6 in Chapter 1. Then a conventional analysis of variance for equal subclass numbers (in this case 1) is performed. The expectations of these mean squares, except for the coefficients of σ_e^2 are exactly the same as they would be had there actually been only one observation per subclass. An algorithm for finding such expectations is given in Henderson (1959).

The coefficient of σ_e^2 is the same in every mean square. To compute this let s = the number of "smallest" subclasses, and let n_i be the number of observations in the i^{th} subclass. Then the coefficient of σ_e^2 is

$$\textstyle\sum_{i=1}^{s} \quad n_i^{-1}/s. \tag{87}$$

Estimate σ_e^2 by

$$[y'y - \sum_{i=1}^{s} \ y_i^2/n_i]/(n-s), \tag{88}$$

where y_i is the sum of observations in the i^{th} subclass. Henderson (1978a) described a simple algorithm for computing sampling variances for the unweighted means method.

We illustrate estimation by a two way mixed model,

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

b_j is fixed.

$Var(\mathbf{a}) \ = \ \mathbf{I}\sigma_a^2, \ Var(\boldsymbol{\gamma}) \ = \ \mathbf{I}\sigma_\gamma^2, \ Var(\mathbf{e}) \ = \ \mathbf{I}\sigma_e^2.$

Let the data be

	n_{ij} B			$\bar{y}_{ij.}$ B		
A	1	2	3	1	2	3
1	5	4	1	8	10	5
2	2	10	5	7	8	4
3	1	4	2	6	9	3
4	2	1	5	10	12	8

The mean squares and their expectation in the unweighted means analysis are

	df	MS	E(ms)
A	3	9.8889	$.475 \ \sigma_e^2 \ + \ \sigma_\gamma^2 \ + \ 3\sigma_a^2$
B	2	22.75	$.475 \ \sigma_e^2 \ + \ \sigma_\gamma^2 \ + \ Q(b)$
AB	6	.3056	$.475 \ \sigma_e^2 \ + \ \sigma_\gamma^2$

Suppose $\hat{\sigma}_e^2$ estimated as described above is .2132. Then

$$\hat{\sigma}_\gamma^2 \ = \ .3056 - .475(.2132) = .2043,$$

and

$$\hat{\sigma}_a^2 \ = \ (9.8889 - .3056)/3 = 3.1944.$$

The coefficient of σ_e^2 is $(5^{-1} + 4^{-1} + ... + 5^{-1})/12 = .475$.

13 Mean Squares For Testing $\mathbf{K}'\mathbf{u}$

Section 2.c in Chapter 4 described a general method for testing the hypothesis, $\mathbf{K}'\boldsymbol{\beta} = \mathbf{0}$ against the unrestricted hypothesis. The mean square for this test is

$$(\boldsymbol{\beta}^o)'\mathbf{K}(\mathbf{K}'\mathbf{C}\mathbf{K})^{-1}\mathbf{K}'\boldsymbol{\beta}^o/\text{f}.$$

$\mathbf{C}$ is a symmetric g-inverse of the GLS equations or is the corresponding partition of a g-inverse of the mixed model equations and f is the number of rows in $\mathbf{K}'$ chosen to have full row rank. Now as in other ANOVA based methods of estimation of variances we can compute as though $\mathbf{u}$ is fixed and then take expectations of the resulting mean squares to estimate variances. The following precaution must be observed. $\mathbf{K}'\mathbf{u}$ must be estimable under a fixed $\mathbf{u}$ model. Then we compute

$$(\mathbf{u}^o)'\mathbf{K}(\mathbf{K}'\mathbf{C}\mathbf{K})^{-1}\mathbf{K}'\mathbf{u}^o/\text{f}, \tag{89}$$

where $\mathbf{u}^o$ is some solution to (90) and f = number of rows in $\mathbf{K}'$.

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}. \tag{90}$$

The assumption is that $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$. $\mathbf{C}$ is the lower $q \times q$ submatrix of a g-inverse of the coefficient matrix in (90). Then the expectation of (89) is

$$f^{-1}\ tr\ \mathbf{K}(\mathbf{K}'\mathbf{C}\mathbf{K})^{-1}\mathbf{K}'\ Var(\mathbf{u}) + \sigma_e^2. \tag{91}$$

This method seems particularly appropriate in the filled subclass case for then with interactions it is relatively easy to find estimable functions of $\mathbf{u}$. To illustrate, consider the two way mixed model of Section 11. Functions for estimating σ_a^2 are

$$\begin{pmatrix} a_1 + \bar{\gamma}_{1.} - a_4 - \bar{\gamma}_{4.} \\ a_2 + \bar{\gamma}_{2.} - a_4 - \bar{\gamma}_{4.} \\ a_3 + \bar{\gamma}_{3.} - a_4 - \bar{\gamma}_{4.} \end{pmatrix} /3.$$

Functions for estimating σ_γ^2 are

$$[\gamma_{ij} \ - \ \gamma_{i3} \ - \ \gamma_{4j} \ + \ \gamma_{34}]/6; \quad i = 1,2,3; \quad j = 1,2.$$

This is an example of a weighted square of means analysis.

The easiest solution to the OLS equations for the 2 way case is $\mathbf{a}^o$, $\mathbf{b}^o$ = null and $\gamma_{ij}^o = \bar{y}_{ij}$. Then the first set of functions can be estimated as $\bar{\gamma}_{i.}^o - \bar{\gamma}_{4.}^o$ $(i = 1,2,3)$. Reduce $\mathbf{K}'$ to this same dimension and take $\mathbf{C}$ as a 12×12 diagonal matrix with diagonal elements $= n_{ij}^{-1}$.