The Annals of Mathematical Statistics
1970, Vol. 41, No. 5, 1735–1748

LINEAR SPACES AND UNBIASED ESTIMATION—APPLICATION TO THE MIXED LINEAR MODEL[1]

BY JUSTUS SEELY[2]

Iowa State University

1. Introduction and summary. Exemplification of the theory developed in [9] using a linear space of random variables other than linear combinations of the components of a random vector, and unbiased estimation for the parameters of a mixed linear model using quadratic estimators are the primary reasons for the considerations in this paper. For a random vector Y with expectation $X\beta$ and covariance matrix $\sum_i \nu_i V_i$ ($\nu_1, \cdots, \nu_m$, and β denote the parameters), interest centers upon quadratic estimability for parametric functions of the form $\sum_{i \leqq j} \lambda_{ij}\beta_i\beta_j + \sum_k \lambda_k \nu_k$ and procedures for obtaining quadratic estimators for such parametric functions. Special emphasis is given to parametric functions of the form $\sum_k \lambda_k \nu_k$.

Unbiased estimation of variance components is the main reason for quadratic estimability considerations regarding parametric functions of the form $\sum_k \lambda_k \nu_k$. Concerning variance component models, Airy, in 1861 (Scheffé [6]), appears to have been the first to introduce a model with more than one source of variation. Such a model is also implied (Scheffé [6]) by Chauvenet in 1863. Fisher [1], [2] reintroduced variance component models and discussed, apparently for the first time, unbiased estimation in such models.

Since Fisher's introduction and discussion of unbiased estimation in models with more than one source of variation, there has been considerable literature published on the subject. One of these papers is a description by Henderson [5] which popularized three methods (now known as Henderson's Methods I, II, and III) for obtaining unbiased estimates of variance components. We mention these methods since they seem to be commonly used in the estimation of variance components. For a review as well as a matrix formulation of the methods see Searle [7]. Among the several pieces of work which have dealt with Henderson's methods, only that of Harville [4] seems to have been concerned with consistency of the equations leading to the estimators and to the existence of unbiased (quadratic) estimators under various conditions. Harville, however, only treats a completely random two-way classification model with interaction. One other result which deals with existence of unbiased quadratic estimators in a completely random model is given by Graybill and Hultquist [3].

In Section 2 the form we assume for a mixed linear model is introduced and the pertinent quantities needed for the application of the results in [9] are obtained. Definitions, terminology, and notation are consistent with the usage in [9]. Section 3

Received January 26, 1970.

[1] This work was prepared partly in connection with work on Contract AF33(615) 68C1343, Office of Aerospace Research, United States Air Force, Wright-Patterson Air Force Base, Ohio and NSF Traineeship.

[2] Now at Oregon State University.

considers parametric functions of the form $\sum_{i \leq j} \lambda_{ij}\beta_i\beta_j + \sum_k \lambda_k \nu_k$ and Section 4 concerns parametric functions of the form $\sum_k \lambda_k \nu_k$. One particular method for obtaining unbiased estimators for linear combinations of variance components is given in Section 4 that is computationally simpler than the Henderson Method III procedure which is the most widely used general approach applicable to any mixed linear model. The method described in Section 4 has the added advantage of giving necessary and sufficient conditions for the existence of unbiased quadratic estimators which is not always the case with the Henderson Method III. In the last section an example is given which illustrates the Henderson Method III procedure from the viewpoint of this paper.

2. The mixed linear model—preliminary notions. In the sequel Y denotes an $n \times 1$ random vector with expectation $X\beta$ and covariance matrix

$$\sum_{i=1}^{m} \nu_i V_i, \tag{2.1}$$

where X is a known $n \times p$ matrix; each V_i is a known $n \times n$ symmetric matrix; the unknown parameters are $\beta = (\beta_1, \cdots, \beta_p)'$ and $\nu = (\nu_1, \cdots, \nu_m)'$; and $\Omega = \{\theta\}$ describes the range of the parameter $\theta = (\beta, \nu)$. The parameter space Ω is included for similarity with the general framework in [9] and to allow the possibility of functional relationships among the parameters. Generally the form of Ω would be implicit from assumptions associated with the random vector Y. Note that the usual forms for a mixed linear model may be put into this framework. For example, consider a typical mixed linear model representation $Y = X\beta + \sum_{i=1}^{k} W_i\gamma_i$ where the W_i's are known matrices and the γ_i's are random vectors with zero expectation. Provided that $E[\gamma_i\gamma_j']$ exists for each i and j, the covariance matrix of $e = \sum_{i=1}^{k} W_i\gamma_i$ may be expressed, regardless of the parameters and the structure of $E[\gamma_i\gamma_j']$, in terms of V_i's and ν_i's as in (2.1). To see this observe that one possible representation for the covariance matrix of $e = (e_1, \cdots, e_n)'$ is

$$\sum_i \operatorname{Var}(e_i)\delta_i\delta_i' + \sum_{i<j} \operatorname{Cov}(e_i, e_j)[\delta_i\delta_j' + \delta_j\delta_i'],$$

where δ_i is the ith unit vector in R^n. This representation would of course not generally be used, i.e., under the usual assumptions on such a model the covariance matrix of e is of the form $\sum_i \sigma_i^2 W_i W_i'$ so that the natural choice here is $\nu_i = \sigma_i^2$ and $V_i = W_i W_i'$.

To obtain a convenient inner product representation for the quadratic estimators let $\mathscr{A}$ denote the set of $n \times n$ real symmetric matrices, let $(\cdot, \cdot)$ be defined on $\mathscr{A} \times \mathscr{A}$ by $(A, B) = \operatorname{tr}(AB)$ for all $A, B \in \mathscr{A}$, let $U = YY'$, and let $\bar{\mathscr{A}} = \{(A, U): A \in \mathscr{A}\} = \{Y'AY: A \in \mathscr{A}\}$. With the usual rules of matrix addition and scalar (real) multiplication $\mathscr{A}$, $(\cdot, \cdot)$, U, and $\bar{\mathscr{A}}$ provide an inner product representation for the quadratic estimators. For $\theta = (\beta, \nu) \in \Omega$ observe that

$$A \in \mathscr{A} \Rightarrow E[(A, U) \mid \theta] = (A, X\beta\beta'X' + \sum_{i=1}^{m} \nu_i V_i).$$

Thus μ_θ, the expectation of U at $\theta \in \Omega$, and $\mathscr{E} = \text{sp}\{\mu_\theta : \theta \in \Omega\}$ are given by

$$(2.2) \qquad \text{(a) } \mu_\theta = X\beta\beta' X' + \textstyle\sum_{i=1}^m \nu_i V_i \quad \text{and}$$
$$\text{(b) } \mathscr{E} = \text{sp}\{X\beta\beta' X' + \textstyle\sum_{i=1}^m \nu_i V_i : (\beta, \nu) \in \Omega\}.$$

Note that μ_θ is simply the usual expectation of the matrix YY'.

For quadratic estimability considerations regarding parametric functions of the form $\sum_{i \leqq j} \lambda_{ij}\beta_i\beta_j + \sum_k \lambda_k \nu_k$, some results in [9] are applicable without any additional assumptions. In this paper, however, only the results obtained in Section 4 of [9] are utilized. To use these results we make the following assumption regarding the parameter space Ω:

$$(2.3) \qquad \text{If } \{\lambda_k, \lambda_{ij}\} \text{ is any set of real numbers such that}$$
$$(\beta, \nu) \in \Omega \Rightarrow \textstyle\sum_{i \leqq j} \lambda_{ij}\beta_i\beta_j + \sum_k \lambda_k \nu_k = 0,$$
$$\text{then } \lambda_k = \lambda_{ij} = 0 \qquad \text{for all } i, j, \text{ and } k.$$

Although (2.3) appears quite restrictive, the condition is actually satisfied by most mixed linear models commonly in use. Further, mixed linear model representations for which (2.3) is not satisfied can be reparametrized to satisfy the condition. In the situation that (2.3) is not satisfied and one does not wish to reparametrize, the results in Section 3 of [9] may be employed. Note that using the results in Section 3 of [9], without assuming (2.3), would be similar to the development we give assuming that (2.3) is true, although the final results would be altered according to the statements in Section 3 of [9].

Perhaps the place where (2.3) is most noticeable is in the structure of the subspace $\mathscr{E}$. Assuming (2.3) is true it is easily verified that

$$(2.4) \qquad \mathscr{E} = \text{sp}\{V_1, V_2, \cdots, V_m\} + \{X\Lambda X' : \Lambda = \Lambda'\}.$$

Some characteristics of the subspace $\{X\Lambda X\}$ are given in the following lemmas. For proofs see Seely [8].

LEMMA 1. *Let $x_1, x_2, \cdots, x_p$ denote the columns of the matrix X and define*

$$B_{ii} = x_i x_i' \qquad 1 \leqq i \leqq p$$
$$B_{ij} = x_i x_j' + x_j x_i' \qquad 1 \leqq i < j \leqq p.$$

The following expressions describe the same subspace:

(a) $\text{sp}\{X\beta\beta' X' : \beta \in R^p\}$,
(b) $\text{sp}\{xz' + zx' : x, z \in \underline{R}(X)\}$,
(c) $\text{sp}\{xx' : x \in \underline{R}(X)\}$,
(d) $\text{sp}\{B_{ij} : 1 \leqq i \leqq j \leqq p\}$,
(e) $\{X\Lambda X' : \Lambda = \Lambda'\}$, and
(f) $\{A : A = A', \underline{R}(A) \subset \underline{R}(X)\}$.

LEMMA 2. *If $\underline{r}(X) = r$, then the dimension of the subspace $\{X\Lambda X: \Lambda = \Lambda'\}$ is equal to $(\frac{1}{2})r(r+1)$.*

Let $\mathscr{B}_0 = \{B_{ij}: 1 \leqq i \leqq j \leqq p\}$ where the B_{ij}'s are defined as in Lemma 1, let $\mathscr{B}_1 = \{V_1, \cdots, V_m\}$, and let $\mathscr{B} = \{B_{11}, B_{12}, \cdots, B_{pp}, V_1, \cdots, V_m\}$. It is clear from Lemma 1 and from (2.4) that

$$\mathscr{E} = \operatorname{sp}\mathscr{B} = \operatorname{sp}\mathscr{B}_0 + \operatorname{sp}\mathscr{B}_1. \tag{2.5}$$

Now define ξ_k and ξ_{ij} from Ω into R^1 by the following:

$$\theta = (\beta, v) \in \Omega \Rightarrow \xi_k(\theta) = v_k \qquad 1 \leqq k \leqq m$$

$$\xi_{ij}(\theta) = \beta_i \beta_j \qquad 1 \leqq i \leqq j \leqq p.$$

From these definitions it is clear for $\theta \in \Omega$ that

$$\mu_\theta = \textstyle\sum_{i \leqq j} \xi_{ij}(\theta) B_{ij} + \sum_k \xi_k(\theta) V_k.$$

Thus, the set of elements $\mathscr{B}$ and the parametric functions $\xi_{11}, \cdots, \xi_{pp}, \xi_1, \cdots, \xi_m$ provide a representation as specified in Section 4 of [9]. Moreover, the assumption that (2.3) is true implies that Condition 4.2 in Section 4 of [9] is also true. Some results which may be stated immediately (Theorem 4 and Corollary 4.1 and Corollary 4.2 in [9]) are given in the next theorem and corollaries. In the remainder of this paper we use the terminology estimable to mean $\mathscr{A}$-estimable in the terminology of [9].

THEOREM 1. *The parametric function v_k is estimable if and only if $V_k \notin \operatorname{sp}\mathscr{B}_0 + \operatorname{sp}\{V_i: i \neq k\}$.*

COROLLARY 1.1. *The parametric function v_k is estimable if and only if there does not exist $\{\alpha_i\}$ and $\Lambda = \Lambda'$ such that $V_k = X\Lambda X' + \sum_{i \neq k} \alpha_i V_i$.*

COROLLARY 1.2. *Each v_k $(k = 1, 2, \cdots, m)$ is estimable if and only if $V_1, \cdots, V_m$ are linearly independent and $\operatorname{sp}\mathscr{B}_1$ and $\operatorname{sp}\mathscr{B}_0$ are disjoint subspaces.*

In Corollary 1.2 if $p = 1$ and $X = (1, 1, \cdots, 1)'$, then the result reduces to Theorem 8 of Graybill and Hultquist [3]. Incidentally, the Graybill and Hultquist result at least tacitly assumes a condition like (2.3). Note that the above results may be stated with inclusion of parameter products $\beta_i\beta_j$.

3. Quadratic estimators via a $\mu_\theta = \mathbf{H}\xi_\theta$ representation. Continuing from the previous section we consider a linear transformation $\mathbf{H}$ which will serve for a $\mu_\theta = \mathbf{H}\xi_\theta$ representation. As noted previously, the set $\mathscr{B}$ and the parametric functions $\xi_{11}, \cdots, \xi_{pp}, \xi_1, \cdots, \xi_m$ are such that the results in Section 4 of [9] may be used directly. The set $\mathscr{B}$ has $M = m + (\frac{1}{2})p(p+1)$ elements (we consider elements to be distinct if they have different indices, although two distinct indices might denote the same element) and so $\mathbf{H}$ will be a linear operator from R^M into $\mathscr{A}$. For convenience the elements of a vector $\rho \in R^M$ are indexed in the form

$$\rho = (\rho_{11}, \rho_{12}, \cdots, \rho_{1p}, \rho_{22}, \cdots, \rho_{pp}, \rho_1, \cdots, \rho_m)'.$$

From the theory developed in Section 4 of [9] and using similar notation, it follows for $\rho \in R^M$ and $\theta \in \Omega$ that

(3.1)
(a) $\xi_\theta = (\xi_{11}(\theta), \cdots, \xi_{pp}(\theta), \xi_1(\theta), \cdots, \xi_m(\theta))'$,
(b) $\mathbf{H}\rho = \sum_{i \leqq j} \rho_{ij} B_{ij} + \sum_k \rho_k V_k$,
(c) $\mu_\theta = \mathbf{H}\xi_\theta = \sum_{i \leqq j} \xi_{ij}(\theta) B_{ij} + \sum_k \xi_k(\theta) V_k$, and
(d) $\mathbf{H}^* A = ((B_{11}, A), \cdots, (B_{pp}, A), (V_1, A), \cdots, (V_m, A))'$.

Using these definitions we illustrate some results from [9]. The symbol $\langle \cdot, \cdot \rangle$ denotes the usual inner production on R^M.

Since $\mathbf{H}^*\mathbf{H}$ is a linear operator from R^M into R^M, it is clear that $\mathbf{H}^*\mathbf{H}$ may be thought of as an $M \times M$ matrix and that $\mathbf{H}^*\mathbf{H}\rho$ denotes usual matrix multiplication. In matrix form

$$\mathbf{H}^*\mathbf{H} = \begin{bmatrix} (B_{11}, B_{11}) \cdots (B_{11}, B_{pp}) & (B_{11}, V_1) \cdots (B_{11}, V_m) \\ \vdots \qquad\qquad \vdots & \vdots \qquad\qquad \vdots \\ (B_{pp}, B_{11}) \cdots (B_{pp}, B_{pp}) & (B_{pp}, V_1) \cdots (B_{pp}, V_m) \\ (V_1, B_{11}) \cdots (V_1, B_{pp}) & (V_1, V_1) \cdots (V_1, V_m) \\ \vdots \qquad\qquad \vdots & \vdots \qquad\qquad \vdots \\ (V_m, B_{11}) \cdots (V_m, B_{pp}) & (V_m, V_1) \cdots (V_m, V_m) \end{bmatrix}; \tag{3.2}$$

and the vector $\mathbf{H}^* U$ is given by

$$(\mathbf{H}^* U)' = ((B_{11}, YY'), \cdots, (B_{pp}, YY'), (V_1, YY'), \cdots, (V_m, YY')). \tag{3.3}$$

From Corollaries 2.2, 2.3, and 3.1 in [9] the following results may be stated.

THEOREM 2. *The parametric function* $\langle \lambda, \xi_\theta \rangle = \sum_{i \leqq j} \lambda_{ij} \beta_i \beta_j + \sum_k \lambda_k v_k$ *is estimable if and only if there exists a* ρ *such that* $\mathbf{H}^*\mathbf{H}\rho = \lambda$.

COROLLARY 2.1. *Suppose that* $\langle \lambda, \xi_\theta \rangle$ *is estimable, then the estimator* $(\mathbf{H}\rho, U) = \sum_{i \leqq j} \rho_{ij}(B_{ij}, YY') + \sum_k \rho_k (V_k, YY')$ *is an unbiased estimator for* $\langle \lambda, \xi_\theta \rangle$ *whenever* ρ *is such that* $\mathbf{H}^*\mathbf{H}\rho = \lambda$.

COROLLARY 2.2. *If* $\langle \lambda, \xi_\theta \rangle$ *is estimable and if* $\hat{\xi}$ *is such that* $\mathbf{H}^*\mathbf{H}\hat{\xi} = \mathbf{H}^* U$, *then* $\langle \lambda, \hat{\xi} \rangle = \sum_{i \leqq j} \lambda_{ij} \hat{\xi}_{ij} + \sum_k \lambda_k \hat{\xi}_k$ *is an unbiased estimator for* $\langle \lambda, \xi_\theta \rangle$.

Using the expression for $\mathbf{H}^*\mathbf{H}$ given in (3.2), it is possible from Theorem 2 to determine estimability for a parametric function of the form $\langle \lambda, \xi_\theta \rangle$. Also, unbiased estimators, when they exist, may be obtained from either Corollary 2.1 or Corollary 2.2.

To make the ideas and expressions in this section more tangible, consider the completely random model

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk},$$

where $i = 1, 2, \cdots, a$; $j = 1, 2, \cdots, b$; and $k = 0, 1, \cdots, n_{ij}$. Write the model in matrix notation as

$$Y = \mu X + W_1 \alpha + W_2 \beta + W_3 \gamma + e.$$

Assume, as is generally the case in the completely random model, that the expectation of Y is μX, that the covariance matrix of Y is $\sum_{i=1}^{3} \sigma_i^2 W_i W_i' + \sigma^2 I$, and that the parameters are completely unknown. In the notation of this section the correspondences are

(a) $\Omega = \{\mu : \mu \in R^1\} \times \{(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma^2)' : \sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma^2 \geqq 0\}$

and $\theta \in \Omega$ is of the form (μ, σ) where $\sigma = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma^2)'$.

(b) $p = 1, m = 4$, and $M = 5$.

(3.4) (c) $B_{11} = XX', V_i = W_i W_i'$ for $i = 1, 2, 3$ and $V_4 = I$.

(d) $\xi_{11}(\theta) = \mu^2, \xi_i(\theta) = \sigma_i^2$ for $i = 1, 2, 3$, and $\xi_4(\theta) = \sigma^2$.

(e) $\mu_\theta = \mu^2 XX' + \sum_{i=1}^{3} \sigma_i^2 W_i W_i' + \sigma^2 I$.

To obtain $\mathbf{H}^*\mathbf{H}$ use (3.2). Substituting the symbols used in this example the form of $\mathbf{H}^*\mathbf{H}$ is

$$H^*H = \begin{bmatrix} (XX', XX') & (XX', W_1W_1') & (XX', W_2W_2') & (XX', W_3W_3') & (XX', I) \\ & (W_1W_1', W_1W_1') & (W_1W_1', W_2W_2') & (W_1W_1', W_3W_3') & (W_1W_1', I) \\ & & (W_2W_2', W_2W_2') & (W_2W_2', W_3W_3') & (W_2W_2', I) \\ & & & (W_3W_3', W_3W_3') & (W_3W_3', I) \\ & & & & (I, I) \end{bmatrix}$$

$$= \begin{bmatrix} n_{..}^2 & \sum_i n_{i.}^2 & \sum_j n_{.j}^2 & \sum_{ij} n_{ij}^2 & n_{..} \\ & \sum_i n_{i.}^2 & \sum_{ij} n_{ij}^2 & \sum_{ij} n_{ij}^2 & n_{..} \\ & & \sum_j n_{.j}^2 & \sum_{ij} n_{ij}^2 & n_{..} \\ & & & \sum_{ij} n_{ij}^2 & n_{..} \\ & & & & n_{..} \end{bmatrix}.$$

Since $\mathbf{H}^*\mathbf{H}$ is symmetric only the upper half of the matrix is indicated and $n_{i.} = \sum_{j=1}^{b} n_{ij}$; $n_{.j} = \sum_{i=1}^{a} n_{ij}$; $n_{..} = \sum_{ij} n_{ij}$. A necessary and sufficient condition for $\langle \lambda, \xi_\theta \rangle = \lambda_{11}\mu^2 + \sum_{i=1}^{3} \lambda_i \sigma_i^2 + \lambda_4 \sigma^2$ to be estimable is that λ' be a member of the row space of the matrix $\mathbf{H}^*\mathbf{H}$. Further, if ρ is such that $\mathbf{H}^*\mathbf{H}\rho = \lambda$, then $(\mathbf{H}\rho, U) = \rho_{11} Y_{..}^2 + \rho_1 \sum_i Y_{i.}^2 + \rho_2 \sum_j Y_{.j}^2 + \rho_3 \sum_{ij} Y_{ij.}^2 + \rho_4 \sum_{ijk} Y_{ijk}^2$ is an unbiased estimator for $\langle \lambda, \xi_\theta \rangle$. To see this note that

$$\mathbf{H}^*U = \begin{bmatrix} (XX', YY') \\ (W_1 W_1', YY') \\ (W_2 W_2', YY') \\ (W_3 W_3', YY') \\ (I, YY') \end{bmatrix} = \begin{bmatrix} Y_{..}^2 \\ \sum_i Y_{i.}^2 \\ \sum_j Y_{.j}^2 \\ \sum_{ij} Y_{ij.}^2 \\ \sum_{ijk} Y_{ijk}^2 \end{bmatrix}.$$

A dot indicates summation over the missing subscript. Also note that if $\lambda_{11}\mu^2 + \sum_{i=1}^{3} \lambda_i \sigma_i^2 + \lambda_4 \sigma^2$ is estimable, then $\langle \lambda, \hat{\xi} \rangle$ is an unbiased estimator for $\langle \lambda, \xi_\theta \rangle$ provided that $\hat{\xi}$ satisfies $\mathbf{H}^*\mathbf{H}\hat{\xi} = \mathbf{H}^* U$.

We do not go into any more detail with this example. However, it may be noted that Harville [4] has considered this same example in detail from the viewpoint of the Henderson Methods I and III. As a final point it is clear that the general approach of this section will work with any mixed model; however, from the dimensions of $\mathbf{H}^*\mathbf{H}$ it is equally clear that the approach is computationally more suitable to a completely random model when interest is in parametric functions of the form $\sum_k \lambda_k \nu_k$.

4. Quadratic estimators via Theorem 5 in [9]. Necessary and sufficient conditions for estimability of parametric functions of the form $\sum_{i \leq j} \lambda_{ij} \beta_i \beta_j + \sum_k \lambda_k \nu_k$ and ways of obtaining unbiased estimators for such functions are considered in Section 3. Usually, however, interest is not centered upon the estimability of an arbitrary $\langle \lambda, \xi_\theta \rangle$, but only on parametric functions of the form $\sum_k \lambda_k \nu_k$. The methods of the last section, although applicable to parametric functions of this form, require that one consider an $M \times M$ matrix to determine estimability. For a fixed m it is clear that the size of the matrix $\mathbf{H}^*\mathbf{H}$ increases in relation to $(\frac{1}{2})p(p+1)$, and thus can become extremely big and unmanageable if p gets very large. Thus, this section is devoted to ways of reducing the size of the matrix involved when emphasis is on parametric functions of the form $\sum_k \lambda_k \nu_k$.

Considering only parametric functions of the form $\sum_k \lambda_k \nu_k$ is similar to focusing attention on a subset of the parameters in the fixed part of the random vector Y. Such considerations regarding only a subset of the fixed parameters, as discussed in Zyskind *et al.* [10], can be obtained by application of Theorem 5 in [9]. For example, consider the fixed part $X\beta$ of the random vector Y. Suppose that $X\beta$ is in the partitioned form $X\beta = X_1\beta_1 + X_2\beta_2$ and that interest is in estimability of parametric functions of the form $\lambda'\beta_2$ within the class $\{a'Y: a \in R^n\}$. For any matrix W such that

$$\underline{R}(W) + \underline{N}(X') = \underline{N}(X_1'), \tag{4.1}$$

it follows from Theorem 5 in [9] that a necessary and sufficient condition for $\lambda'\beta_2$ to be estimable within the class $\{a'Y\}$ is the existence of a ρ such that $X_2'W\rho = \lambda$. In [10] the matrix used for W is $(I - X_1(X_1'X_1)^- X_1')X_2$ where $(X_1'X_1)^-$ denotes any g-inverse for the matrix $X_1'X_1$. This matrix yields a direct sum in Equation 4.1 so that Corollary 5.1 in [9] is also applicable. This same technique, i.e., choosing $W = (I - X_1(X_1'X_1)^- X_1')X_2$ for linear estimability considerations regarding parametric functions $\lambda'\beta_2$, may be employed to obtain an $m \times m$ matrix for estimability considerations regarding parametric functions of the form $\sum_k \lambda_k \nu_k$. Although such a technique could be used, we utilize Theorem 5 of [9] to investigate another way to obtain a more manageable matrix for estimability considerations for parametric functions of the form $\sum_k \lambda_k \nu_k$.

To utilize Theorem 5 in [9] we need a linear operator $\mathbf{W}$ such that $\underline{R}(\mathbf{W})+\underline{H}(\mathbf{H}^*)=\mathscr{B}_0{}^{\perp}$, then the pertinent condition for $\sum_k\lambda_k\nu_k$ to be estimable is the existence of a ρ such that $\mathbf{H}^*\mathbf{W}\rho=\lambda(\lambda_{ij}=0$ for $1\leqq i\leqq j\leqq p)$. In order that the linear operator $\mathbf{H}^*\mathbf{W}$ be conveniently represented in the form of a matrix, $\mathbf{W}$ is selected as a linear operator from R^m into $\mathscr{A}$; and to show that $\mathbf{W}$ satisfies the required condition the following two lemmas are given. For proofs of the lemmas see [8].

LEMMA 3. *If $\mathbf{T}$ is any symmetric and nonnegative linear operator on $\mathscr{A}$ and $\mathbf{B}$ is a linear operator into $\mathscr{A}$, then $\underline{R}(\mathbf{TB})\cap\underline{N}(\mathbf{B}^*)=\{0\}$. Further, if $\underline{N}(\mathbf{B}^*)\subset\underline{R}(\mathbf{T})$ then $\underline{R}(\mathbf{TB})\oplus\underline{N}(\mathbf{B}^*)=\underline{R}(\mathbf{T})$.*

LEMMA 4. *Let P denote the symmetric idempotent matrix such that $\underline{R}(P)=\underline{R}(X)$ and let $N=I-P$. Define $\mathbf{T}$ from $\mathscr{A}$ into $\mathscr{A}$ by $A\in\mathscr{A}\Rightarrow\mathbf{T}A=NA+AN$, then $\mathbf{T}$ is a symmetric and nonnegative linear operator on $\mathscr{A}$ such that $\underline{R}(\mathbf{T})=\mathscr{B}_0{}^{\perp}$.*

Let N and $\mathbf{T}$ be defined as in Lemma 4. Observe that $\underline{N}(\mathbf{H}^*)\subset\mathscr{B}_0{}^{\perp}$ so that Lemma 3 and Lemma 4 imply $\underline{R}(\mathbf{TH})\oplus\underline{N}(\mathbf{H}^*)=\mathscr{B}_0{}^{\perp}$. Let $\mathbf{Q}$ denote the linear operator from R^m into $\mathscr{A}$ defined by $\mathbf{Q}\rho=\sum_{i=1}^m\rho_iV_i$ for all $\rho\in R^m$ and define $\mathbf{W}=(\frac{1}{2})\mathbf{TQ}$, i.e., for $\rho\in R^m$

$$\mathbf{W}\rho=(\tfrac{1}{2})\textstyle\sum_{i=1}^m\rho_i(NV_i+V_iN).$$

Note that $\underline{R}(\mathbf{TQ})=\underline{R}(\mathbf{TH})$; and so $\mathbf{W}$ satisfies the relationship $\underline{R}(\mathbf{W})\oplus\underline{N}(\mathbf{H}^*)=\mathscr{B}_0{}^{\perp}$.

From the preceding paragraph it is clear that $\mathbf{W}$ is a suitable linear operator to utilize Theorem 5 of [9]. Let $\mathbf{0}$ denote a $(\frac{1}{2})p(p+1)$ vector of zeros and note that $(B_{ij},\mathbf{W}\rho)=0$ for all i,j, and ρ. For $\lambda'=(\lambda_1,\cdots,\lambda_m)$ it is clear for $\rho\in R^m$ that

$$\mathbf{H}^*\mathbf{W}\rho=\begin{bmatrix}\mathbf{0}\\ \mathbf{Q}^*\mathbf{W}\rho\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ \lambda\end{bmatrix}\Leftrightarrow\mathbf{Q}^*\mathbf{W}\rho=\lambda.$$

Hence, estimability considerations regarding parametric functions $\sum_k\lambda_k\nu_k$ may be obtained through the linear operator $\mathbf{Q}^*\mathbf{W}$. The matrix form for $\mathbf{Q}^*\mathbf{W}$ and the vector form for $\mathbf{W}^*U$ are

$$\mathbf{Q}^*\mathbf{W}=\begin{bmatrix}\operatorname{tr}(V_1NV_1)\cdots\operatorname{tr}(V_1NV_m)\\ \vdots \qquad\qquad \vdots\\ \operatorname{tr}(V_mNV_1)\cdots\operatorname{tr}(V_mNV_m)\end{bmatrix}\tag{4.2}$$

and

$$\mathbf{W}^*U=\begin{bmatrix}Y'NV_1Y\\ \vdots\\ Y'NV_mY\end{bmatrix}.$$

Using these expressions the following statements, which are immediate consequences of Theorem 5 and comments in Section 4 of [9], may be made.

THEOREM 3. *A necessary and sufficient condition for $\sum_k\lambda_k\nu_k$ to be estimable is the existence of a ρ such that $\mathbf{Q}^*\mathbf{W}\rho=\lambda$.*

COROLLARY 3.1. *Suppose* ρ *and* λ *are such that* $\mathbf{Q}^*\mathbf{W}\rho = \lambda$, *then the estimator* $(\mathbf{W}\rho, U) = \langle \rho, \mathbf{W}^*U\rangle = \sum_{i=1}^{m} \rho_i Y'NV_iY$ *is an unbiased estimator for* $\sum_k \lambda_k v_k$.

COROLLARY 3.2. *Suppose* $\hat{v} = (\hat{v}_1, \cdots, \hat{v}_m)'$ *is such that* $\mathbf{W}^*\mathbf{Q}\hat{v} = \mathbf{W}^*U$, *then* $\sum_k \lambda_k \hat{v}_k$ *is an unbiased estimator for any estimable parametric function of the form* $\sum_k \lambda_k v_k$.

As in Section 3 we illustrate the preceding results with a specific example. Consider the mixed linear model

$$Y_{ijk} = \mu + \beta_i + \gamma_j + e_{ijk},$$

where $i = i, 2, \cdots, b$; $j = 1, 2, \cdots, c$; $k = 0, 1, \cdots, n_{ij}$; μ, β_i are fixed effects; and γ_j, e_{ijk} are random. Write the model in matrix notation as

$$Y = X\beta + G\gamma + e \qquad (\beta = (\mu, \beta_1, \cdots, \beta_b)');$$

then making the usual assumptions on such a model the notational correspondences are

(4.3)

(a) $\Omega = \{\beta : \mu, \beta_i \in R^1\} \times \{(\sigma_1^2, \sigma^2)' : \sigma_1^2, \sigma^2 \geqq 0\}$ and $\theta \in \Omega$ is of the form (β, σ) where $\sigma = (\sigma_1^2, \sigma^2)'$.

(b) $p = b+1$, $m = 2$, and $M = 2 + (\frac{1}{2})(b+1)(b+2)$.

(c) $V_1 = GG'$ and $V_2 = I$.

(d) $X = (x_0, x_1, \cdots, x_b)$ and the B_{ij}'s are formed as previously defined.

(e) $\mu_\theta = X\beta\beta'X' + \sigma_1^2 GG' + \sigma^2 I$.

(f) $\mathscr{B}_0 = \{B_{ij} : 0 \leqq i \leqq j \leqq b\}$ and $\mathscr{B}_1 = \{GG', I\}$.

From (4.2) the following description for $\mathbf{Q}^*\mathbf{W}$ may be obtained:

$$\mathbf{Q}^*\mathbf{W} = \begin{bmatrix} \operatorname{tr}[(G'G)^2] - \operatorname{tr}[(G'G)(G'PG)] & \operatorname{tr}(G'G) - \operatorname{tr}(G'PG) \\ \operatorname{tr}(G'G) - \operatorname{tr}(G'PG) & \operatorname{tr}(I-P) \end{bmatrix} \tag{4.4}$$

$$= \begin{bmatrix} \sum_j n_{.j}^2 - \sum_i n_{i.}^{-1} \sum_j n_{.j} n_{ij}^2 & n_{..} - \sum_i n_{i.}^{-1} \sum_j n_{ij}^2 \\ n_{..} - \sum_i n_{i.}^{-1} \sum_j n_{ij}^2 & n_{..} - b \end{bmatrix}.$$

In the last expression we have assumed that $n_{i.} \neq 0$ for $i = 1, 2, \cdots, b$. A dot in place of a subscript indicates that summation has been carried out over that subscript. It is now straightforward to apply the results stated in this section. For example, a necessary and sufficient condition for $\lambda_1\sigma_1^2 + \lambda\sigma^2$ to be estimable is that (λ_1, λ) be a member of the row space of the 2×2 matrix $\mathbf{Q}^*\mathbf{W}$.

5. The Henderson Method III procedure—an example. In this section the Henderson Method III procedure for estimating variance components is illustrated, via an example, from the viewpoint of this paper. In addition to obtaining the pertinent estimating equations, their status with regard to the following two conditions is also indicated:

(5.1) (a) $\underline{R}(\mathbf{W})+\underline{N}(\mathbf{H}^*) = \mathscr{B}_0^{\perp}$ and

(b) $\underline{R}(\mathbf{W}^*\mathbf{Q}) = \underline{R}(\mathbf{W}^*)$.

That is, does the Henderson procedure provide necessary and sufficient conditions for estimability of linear functions of the variance components (5.1.a) and are the estimating equations consistent (5.1.b)? Although these two questions, i.e., estimability and consistency, are not generally considered in the literature on Henderson's procedure, they would seem to be relevant questions. An exception to this statement is the treatment of a completely random two-way classification model with interaction given by Harville [4]. Harville considers Method I and Method III of Henderson and answers both the questions of estimability and of consistency.

For the example consider a mixed linear model $Y = X\beta + e$ with covariance structure $\sigma_1^2 V + \sigma^2 I$. As is usually the case assume that $\Omega = \{\beta : \beta \in R^p\} \times \{(\sigma_1^2, \sigma^2)' : \sigma_1^2, \sigma^2 \geqq 0\}$. The parameter $\theta \in \Omega$ is taken to be of the form $\theta = (\beta, \sigma)$ where $\sigma = (\sigma_1^2, \sigma^2)'$. Let P denote the matrix in $\mathscr{A}$ which is the orthogonal projection on $\underline{R}(X)$, i.e., $P = P' = P^2$ and $\underline{R}(P) = \underline{R}(X)$, and let L denote the matrix in $\mathscr{A}$ which is the orthogonal projection on $\underline{R}(X)+\underline{R}(V)$.

Let $\mathbf{W}$ be the linear operator from R^2 into $\mathscr{A}$ defined by $\mathbf{W}\rho = \rho_1 P_1 + \rho_2 P_2$ for all $\rho \in R^2$ where $P_1 = L - P$ and $P_2 = I - L$. Using the definitions previously given for $\mathbf{H}$ and $\mathbf{Q}$, observe for $\alpha \in R^M$ and $\rho = (\alpha_1, \alpha_2)'$, i.e., ρ denotes the last two components of α, that

$$\begin{aligned} \mathbf{W}^*\mathbf{H}\alpha &= \textstyle\sum_{i \leqq j} \alpha_{ij} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \alpha_1 \begin{bmatrix} (P_1, V) \\ (P_2, V) \end{bmatrix} + \alpha_2 \begin{bmatrix} (P_1, I) \\ (P_2, I) \end{bmatrix} \\ &= \begin{bmatrix} (P_1, V) & (P_1, I) \\ 0 & (P_2, I) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \qquad (5.2) \\ &= \mathbf{W}^*\mathbf{Q}\rho. \end{aligned}$$

It is clear that $\underline{R}(\mathbf{W}^*\mathbf{H}) = \underline{R}(\mathbf{W}^*\mathbf{Q})$ and that we may consider $\mathbf{W}^*\mathbf{Q}$ as the 2×2 matrix given in (5.2). The following is a statement of the Henderson Method III procedure for our example.

Henderson Method III. Suppose that $\hat{\sigma} = (\hat{\sigma}_1^2, \hat{\sigma}^2)'$ satisfies the following:

$$\mathbf{W}^*\mathbf{Q}\hat{\sigma} = \begin{bmatrix} (P_1, V) & (P_1, I) \\ 0 & (P_2, I) \end{bmatrix} \begin{bmatrix} \hat{\sigma}_1^2 \\ \hat{\sigma}^2 \end{bmatrix} = \begin{bmatrix} Y'P_1 Y \\ Y'P_2 Y \end{bmatrix} = W^*U, \qquad (5.3)$$

then $\lambda_1 \hat{\sigma}_1^2 + \lambda \hat{\sigma}^2$ is an unbiased estimator for $\lambda_1 \sigma_1^2 + \lambda \sigma^2$ provided that (λ_1, λ) is a member of the row space of the matrix $\mathbf{W}^*\mathbf{Q}$.

The statement referred to as the Henderson Method III will follow immediately from Theorem 3 in [9] after we demonstrate the consistency of the equations in (5.3). In most references to the Henderson procedure it is apparently assumed, and is usually the case, that the 2×2 matrix $\mathbf{W}^*\mathbf{Q}$ is nonsingular so that no complications arise with regard to estimability or with regard to consistency. With regard to estimability, however, it is shown later than even if σ_1^2 and σ^2 are estimable, it does not follow that $\mathbf{W}^*\mathbf{Q}$ is nonsingular. The way the equations in (5.3) are usually obtained is by taking expectations; that is,

$$E[Y'P_1 Y | \theta] = \sigma_1^2(P_1, V) + \sigma^2(P_1, I)$$

$$E[Y'P_2 Y | \theta] = \sigma^2(P_2, I).$$

The procedure is then usually stated as follows: Set the expectations equal to their respective functions and solve to obtain the estimates. This type of argument, however, does not demonstrate that the equations are consistent.

One way to demonstrate that the equations $\mathbf{W}^*\mathbf{Q}\hat{\sigma} = \mathbf{W}^*U$ are consistent is to show that $\underline{R}(\mathbf{W}^*\mathbf{Q}) = \underline{R}(\mathbf{W}^*)$ or equivalently that $\underline{R}(\mathbf{W}) \cap \underline{N}(\mathbf{Q}^*) = \{0\}$. Suppose that $\mathbf{W}\rho \in \underline{N}(\mathbf{Q}^*)$ for some ρ, then it must follow that

$$\rho_1(V, P_1) + \rho_2(V, P_2) = 0$$

$$\rho_1(I, P_1) + \rho_2(I, P_2) = 0.$$

These conditions may be shown to imply that $\rho_1 P_1 + \rho_2 P_2 = \mathbf{W}\rho = 0$ so that $\underline{N}(\mathbf{Q}^*) \cap \underline{R}(\mathbf{W}) = \{0\}$. Thus, the equations are consistent.

In determining whether Condition 5.1.a holds for the example we use a dimension argument. Since it is already established that $\underline{R}(\mathbf{W}^*\mathbf{Q}) = \underline{R}(\mathbf{W}^*\mathbf{H}) = \underline{R}(\mathbf{W}^*)$, we need only determine if $\underline{r}(\mathbf{W}) = \dim \mathscr{B}_0^\perp - \underline{n}(\mathbf{H}^*)$. Note that $\underline{R}(\mathbf{H}) = \operatorname{sp} \mathscr{B}_0 + \operatorname{sp} \mathscr{B}_1$ so that by appropriate substitution the condition on $\underline{r}(\mathbf{W})$ reduces to

$$\underline{r}(W) = \dim [\operatorname{sp} \mathscr{B}_1] - \dim [\operatorname{sp} \mathscr{B}_1 \cap \operatorname{sp} \mathscr{B}_0]. \tag{5.4}$$

In words (5.4) says (Corollary 5.2 in [9]) that the rank of $\mathbf{W}$ must equal the number of linearly independent estimable functions of the form $\lambda_1 \sigma_1^2 + \lambda \sigma^2$ in order for Condition 5.1.a to be satisfied.

Table 1 summarizes the various relationships between $\mathscr{B}_1$ and $\mathscr{B}_0$ and gives

TABLE 1

The possible relationships between $\mathscr{B}_1$ and $\mathscr{B}_0$.

Line	Basis for $\operatorname{sp} \mathscr{B}_1$	Basis for $\operatorname{sp} \mathscr{B}_0 \cap \operatorname{sp} \mathscr{B}_1$	estimable functions	$\dim [\operatorname{sp} \mathscr{B}_1] - \dim [\operatorname{sp} \mathscr{B}_1 \cap \operatorname{sp} \mathscr{B}_0]$
1	I	I	—	0
2	$I(V = kI)$	0	$k\sigma_1^2 + \sigma^2$	1
3	V, I	V, I	—	0
4	V, I	V	σ^2	1
5	V, I	I	impossible	impossible
6	V, I	$kI + k_1 V (k, k_1 \neq 0)$	$\sigma^2 - (k/k_1)\sigma_1^2$	1
7	V, I	0	σ_1^2, σ^2	2

$\dim[\text{sp}\,\mathscr{B}_1] - \dim[\text{sp}\,\mathscr{B}_1 \cap \text{sp}\,\mathscr{B}_0]$. The results in this table may be easily verified. For example, in Line 6 the matrices V and I are independent and $kI + k_1 V$ is a basis for $\text{sp}\,\mathscr{B}_0 \cap \text{sp}\,\mathscr{B}_1$. Thus, $\dim[\text{sp}\,\mathscr{B}_1] - \dim[\text{sp}\,\mathscr{B}_0 \cap \text{sp}\,\mathscr{B}_1] = 1$ and there is only one independent estimable function of the form $\lambda\sigma^2 + \lambda_1 \sigma_1^2$. To obtain the form of the estimable function in Line 6 use Theorem 4 of [9]. That is, define a linear functional F on $\mathscr{E}$ as follows:

$$F(B_{ij}) = 0 \qquad \text{for} \quad 1 \leqq i \leqq j \leqq p$$
$$F(V) = \lambda_1$$
$$F(I) = \lambda,$$

and extend linearly. Since $kI + k_1 V \in \text{sp}\,\mathscr{B}_0$, it follows that $0 = F(kI + k_1 V) = kF(I) + k_1 F(V) = k\lambda + k_1 \lambda_1$; and so $k\lambda + k_1 \lambda_1$ must be zero. Therefore, λ_1 must be of the form $-(k/k_1)\lambda$.

To determine if Condition 5.1.a is satisfied we need to obtain the rank of $\mathbf{W}$ under the various relationships between $\mathscr{B}_0$ and $\mathscr{B}_1$. This is done in Table 2. Note that $\underline{r}(\mathbf{W}) = \dim[\text{sp}\{P_1, P_2\}]$ is equal to the number of non-zero elements in $\{P_1, P_2\}$. The results in Table 2, as those in Table 1, are easily verified. For example,

TABLE 2
The rank of the operator **W**

Line	Basis for sp $\mathscr{B}_1$	Basis for sp $\mathscr{B}_1 \cap$ sp $\mathscr{B}_0$	P_1, P_2	$\underline{r}(\mathbf{W})$
1	I	I	$P_1 = P_2 = 0$	0
2	$I(V = kI)$	0	$P_1 \neq 0, P_2 = 0$	1
3	V, I	V, I	$P_1 = P_2 = 0$	0
4	V, I	V	$P_1 = 0, P_2 \neq 0$	1
5	V, I	I	impossible	impossible
6	V, I	$kI + k_1 V (k_1, k \neq 0)$	$P_1 \neq 0, P_2 = 0$	1
7	V, I	0	$P_1 \neq 0, P_2 = ?$	$\underline{r}(\mathbf{W}) \geqq 1$

suppose (Line 7) that V and I are independent and that $\text{sp}\,\mathscr{B}_0 \cap \text{sp}\,\mathscr{B}_1 = \{0\}$. If $P_1 = 0$ then V must be in $\text{sp}\,\mathscr{B}_0$; but by assumption this is not true so that $P_1 \neq 0$. Since $P_2 = 0$ if and only if $\underline{R}(X) + \underline{R}(V) = R^n$ the only conclusions which may be drawn regarding P_2 are

(5.5) (a) $\underline{R}(X) + \underline{R}(V) = R^n \Rightarrow \underline{r}(\mathbf{W}) = 1$ $(P_2 = 0)$ and

(b) $\underline{R}(X) + \underline{R}(V) \neq R^n \Rightarrow \underline{r}(\mathbf{W}) = 2$ $(P_2 \neq 0)$.

Thus we may conclude that Condition 5.1.a is true provided that the following conditions are not simultaneously satisfied:

$$(5.6) \qquad \begin{aligned} &\text{(a) } V \text{ and } I \text{ independent,} \\ &\text{(b) } \mathrm{sp}\,\mathscr{B}_0 \cap \mathrm{sp}\,\mathscr{B}_1 = \{0\}, \qquad \text{and} \\ &\text{(c) } \underline{R}(X) + \underline{R}(V) = R^n. \end{aligned}$$

Note that Corollary 1.2 implies that both ${\sigma_1}^2$ and σ^2 are estimable when (5.6.a) and (5.6.b) are satisfied. Hence, if Conditions 5.6 are satisfied the Henderson Method III will not provide estimators for both σ^2 and ${\sigma_1}^2$ even though such estimators exist. In the second of the following two specific examples we illustrate that such a situation can exist.

Consider the linear model $Y = \mu X + G\alpha + e$ arising from the one-way classification model $Y_{ij} = \mu + \alpha_i + e_{ij}$, where α_i, e_{ij} are random and μ is fixed. Make the usual assumptions on the covariance matrix, i.e., ${\sigma_1}^2 GG' + \sigma^2 I$. It is easily verified that $\underline{R}(X) + \underline{R}(GG') = R^n$ if and only if $G = I$. Thus, the conditions in (5.6) cannot all be satisfied so that the Henderson procedure satisfies Condition 5.1.a and Condition 5.1.b.

Now consider the example in Section 4. That is, the mixed linear model $Y = X\beta + G\gamma + e$, which arises from a two-way classification with one factor random and the other fixed. For a more complete description see the expressions in (4.3). The following observation arrangement

μ	β_1	β_2	γ_1	γ_2	γ_3
1	1	0	1	0	0
1	1	0	0	1	0
1	1	0	0	0	1
1	0	1	1	0	0

can be shown to satisfy Conditions 5.6.a, 5.6.b, and 5.6.c. Thus, there can exist situations such that the conditions in (5.6) are all satisfied. As a side point note that the 2×2 matrix in (4.4) is

$$\begin{bmatrix} 8/3 & 2 \\ 2 & 2 \end{bmatrix},$$

and its determinant is $\frac{4}{3}$. Thus, we have another verification that both ${\sigma_1}^2$ and σ^2 are estimable. Harville [4], using a somewhat different approach and considering a completely random two factor model with interaction, found an analogous situation with regard to the Henderson Method III, although he did find that the Henderson Method I procedure satisfies both Condition 5.1.a and Condition 5.1.b.

Acknowledgment. I wish to express my appreciation to Professor George Zyskind who served as my Major Professor while I was at Iowa State University. This work is a part of my Ph.D. thesis which was done under the guidance of Professor Zyskind.

REFERENCES

[1] Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinburgh* **52** 399–433.

[2] Fisher, R. A. (1925). *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh.

[3] Graybill, F. A. and Hultquist, R. A. (1961). Theorems concerning Eisenhart's Model II. *Ann. Math. Statist.* **32** 261–269.

[4] Harville, D. A. (1967). Estimability of variance components for the two-way classification with interaction. *Ann. Math. Statist.* **38** 1508–1519.

[5] Henderson, C. R. (1953). Estimation of variance and covariance components. *Biometrics* **9** 226–252.

[6] Scheffé, H. (1956). Alternative models for the analysis of variance. *Ann. Math. Statist.* **27** 251–271.

[7] Searle, S. R. (1969). Another look at Henderson's methods of estimating variance components. *Biometrics* **24** 749–778.

[8] Seely, J. (1969). Estimation in finite-dimensional vector spaces with application to the mixed linear model. Ph.D. Dissertation, Iowa State Univ.

[9] Seely, J. (1970). Linear spaces and unbiased estimation. *Ann. Math. Statist.* **41** 1725–1734.

[10] Zyskind, G., Kempthorne, O., White, R. F., Dayhoff, E. E. and Doerfler, T. E. (1964). Research on analysis of variance and related topics. Aerospace Research Laboratories, Tech. Report 64–193, Wright-Patterson Air Force Base.