The Annals of Statistics
1980, Vol. 8, No. 2, 399–406

ESTIMABILITY IN PARTITIONED LINEAR MODELS

BY JUSTUS SEELY AND DAVID BIRKES

Oregon State University

Some estimability facts for partitioned linear models with constraints are presented. For a model $E(Y) = X_1\pi_1 + X_2\pi_2$ with constraints on π_1 and π_2 a reduced model is derived that contains all information regarding the estimability (and also regarding the blues) of parametric functions $b'\pi_2$. For a model $E(Y) = X_0\pi_0 + X_1\pi_1 + X_2\pi_2$ with constraints on π_0, π_1 and π_2, several necessary and sufficient conditions are given for when estimability of $b'\pi_2$ in the original model is equivalent to estimability in the simpler model $E(Y) = X_0\pi_0 + X_2\pi_2$.

1. Introduction and summary. Partitioned linear models occur in several contexts. Perhaps the most familiar place they arise is in the analysis of covariance. They also occur naturally in a situation like a resolution IV design where the main effects are of primary interest and the two-factor interactions are regarded as nuisance parameters. Partitioning can also be used as a mechanism to simplify estimability considerations in classification models. For example, if it can be determined in an additive three-way model $E(Y_{ijkp}) = \mu + \alpha_i + \beta_j + \gamma_k$ that all γ contrasts are estimable, then all α and β contrasts are estimable if and only if they are estimable in the simpler model $E(Y_{ijkp}) = \mu + \alpha_i + \beta_j$. This particular result is known (see Eccleston and Russell [3] or Raghavarao and Federer [4]), but in Theorem 3.11 below it has been strengthened and generalized.

The earliest systematic investigation of partitioned models seems to have been by Rao [5]. More recently Zyskind et al. [11] have considered such models. Also, Searle [6] treats partitioned models but his treatment is applicable only to "true" covariance models and not to arbitrary partitioned models. The above authors do touch on estimability but it is not their main theme. Wynn [9], however, has recently given a result for an unconstrained model that essentially says that if a subset of parameters are estimable, then they can be dropped from the model in further estimability considerations. This result is implied by our results.

Some known facts about constrained linear models are summarized in the next section. In Section 3 a generalization of the usual error analysis of covariance is given, which is then used to obtain the main results. In the last section some examples are given to illustrate the results of Section 3.

The notation $\mathbf{R}(H)$, $\mathbf{N}(H)$, $\mathbf{r}(H)$, and H^- is used for a matrix H to denote the range, null space, rank, and an arbitrary g-inverse respectively. Two facts about

Received March 1977; revised May 1978.

AMS 1970 *subject classifications.* Primary 62J99; secondary 62K99.

Key words and phrases. Partitioned linear model, estimable linear parametric functions, degrees of freedom.

matrices which should be remembered are that $\mathbf{R}(H)^{\perp} = \mathbf{N}(H')$ where $\perp$ denotes orthogonal complement and that $\mathbf{R}(H, G) = \mathbf{R}(H) + \mathbf{R}(G)$, i.e., the range of the partitioned matrix (H, G) is the sum of the individual ranges.

2. Linear model prerequisites. Suppose Y is an $n \times 1$ random vector having the linear model structure

$$E(Y) = X\pi, \qquad \Delta'\pi = 0, \tag{2.1}$$

where X and Δ' are known $n \times p$ and $s \times p$ matrices. Notice that the usual unconstrained model is obtained by setting $\Delta = 0$.

The vector space $\Omega = \{X\pi : \Delta'\pi = 0\}$ is called the *regression space* and $r = \dim \Omega$ is called the *rank* of the model or the *regression degrees of freedom*. A *linear parametric function* of π is a linear functional defined on the parameter space $\mathbf{N}(\Delta')$. As is customary, a linear parametric function is denoted simply as $c'\pi$. A linear parametric function $c'\pi$ is said to be *estimable* if and only if there exists a linear unbiased estimator for $c'\pi$. The vector space of all estimable linear parametric functions is denoted by Γ and $d = \dim \Gamma$. A parametric vector $L'\pi$ is said to be estimable provided that each of its components is estimable.

There are several ways of characterizing an estimable linear parametric function (e.g., see [8]). The one used here is that $c'\pi$ is estimable if and only if $c \in \mathbf{R}(X', \Delta)$. Similarly there are various expressions for r (e.g., see [7] page 1728). In particular

$$r = \mathbf{r}(XT) = \mathbf{r}(X', \Delta) - \mathbf{r}(\Delta) = d \tag{2.2}$$

where T can be any matrix satisfying $\mathbf{R}(T) = \mathbf{N}(\Delta')$. Notice, too, for such a T that $\Omega = \mathbf{R}(U)$ where $U = XT$. Finally, note that π is estimable if and only if $p = \mathbf{r}(X', \Delta)$ which by (2.2) is equivalent to $r = p - \mathbf{r}(\Delta)$.

REMARK 2.3. A matrix T satisfying $\mathbf{R}(T) = \mathbf{N}(\Delta')$ has some additional interesting properties. Set $U = XT$. Then $c'\pi$ is estimable if and only if $T'c \in \mathbf{R}(U')$. Also, if $\mathrm{Cov}(Y) = \sigma^2 I$, then $\hat{\pi} = T(U'U)^{-}U'Y$ is such that $c'\hat{\pi}$ is the best linear unbiased estimator (blue) for $c'\pi$ whenever $c'\pi$ is estimable. Computing a $\hat{\pi}$ in this fashion requires basically two inverses—one to get T as indicated in Remark 2.4 below and one to get $(U'U)^{-}$. Oftentimes, however, the amount of work involved will be less than with other methods for computing $\hat{\pi}$ because of the smaller dimensions of the matrices involved.

REMARK 2.4. The conditions required for $\mathbf{R}(T) = \mathbf{N}(\Delta')$ are $\Delta'T = 0$ and $\mathbf{r}(T) = p - \mathbf{r}(\Delta)$. An appropriate T can often be found by inspection. Alternatively, $T = I - \Delta(\Delta'\Delta)^{-}\Delta'$ or, as suggested by a referee, $T = I - (\Delta\Delta^{-})'$ will suffice. Another alternative is the following. Suppose the rows of Δ' are linearly independent. This can always be achieved by eliminating redundant rows. For notational convenience suppose further $\Delta' = (A, B)$ where A is $s \times s$ and nonsingular. Then $T = (-(A^{-1}B)', I_{p-s})'$ satisfies $\mathbf{R}(T) = \mathbf{N}(\Delta')$. This last T has the added feature that its columns are linearly independent which makes $U'U$ in Remark 2.3 of minimal dimensions.

3. The partitioned model. Consider the linear model in (2.1) when $E(Y)$ can be expressed in the partitioned form

$$E(Y) = X_1\pi_1 + X_2\pi_2, \quad \Delta_1'\pi_1 = 0, \quad \Delta_2'\pi_2 = 0. \tag{3.1}$$

That is, $X = (X_1, X_2)$ and $\Delta' = \text{diag}(\Delta_1', \Delta_2')$. The only requirement on (2.1) for such a partition is that the constraints on π_1 be independent of those on π_2.

Our first objective is to characterize the estimable linear parametric functions $b'\pi_2$ which involve only π_2. Let Γ_2 denote the vector space of all such parametric functions and let $d_2 = \dim \Gamma_2$. Our second objective is to use information about Γ_2 to simplify the procedure of checking whether or not a parametric function $a'\pi_1$ is estimable.

We need to consider several linear models that are derived from model (3.1). Two such models are

$$\text{model } Mk : E(Y) = X_k\pi_k, \qquad \Delta_k'\pi_k = 0,$$

for $k = 1, 2$. Let r_k and Ω_k denote the rank and regression space of model Mk. Models $M1$ and $M2$, like all of the derived models we consider, have the form (2.1) so that the ideas and results in Section 2 can be applied to these models. For example, we can apply (2.2) to model $M2$ to obtain $r_2 = \mathbf{r}(X_2', \Delta_2) - \mathbf{r}(\Delta_2)$. In this regard, the important thing is to identify the proper design matrix (X) and constraint matrix (Δ'). Hereafter, the term estimable will always be taken to mean estimable with respect to (w.r.t.) model (3.1) whereas estimability w.r.t. any of the derived models will be explicitly indicated.

For a linear estimator $t'Y$ to have an expectation $b'\pi_2$ for some vector b, it is necessary and sufficient that $t'X_1\pi_1 = 0$ for all $\pi_1 \in \mathbf{N}(\Delta_1')$ or equivalently that $t \in \Omega_1^\perp$. Set $Z = W'Y$ where W is any matrix satisfying $\mathbf{R}(W) = \Omega_1^\perp$. The model for Z is

$$E(Z) = W'X_2\pi_2, \qquad \Delta_2'\pi_2 = 0. \tag{3.2}$$

Since the class of estimators $t'Y$, $t \in \Omega_1^\perp$, is the same as the class of $h'Z$, $h \in R^q$ (where q is the number of columns of W), we can state

THEOREM 3.3. *A parametric function $b'\pi_2$ is estimable if and only if it is estimable* w.r.t. *model* (3.2).

The importance of this theorem is that the results in Section 2 can be applied. In particular, it follows that $b'\pi_2$ is estimable if and only if $b \in \mathbf{R}(X_2'W, \Delta_2)$ and that d_2 is the rank of model (3.2).

REMARK 3.4. The condition $\{t'Y : t \in \Omega_1^\perp\} = \{h'Z : h \in R^q\}$ also trivially implies that blues for $b'\pi_2$ under models (3.1) and (3.2) are identical. Also, an error contrast $f'Y$ satisfies $f \in \Omega_1^\perp$ so that the residual sum of squares is the same under both models. For computations it must, of course, be remembered that $\text{Cov}(Z) = W'\text{Cov}(Y)W$. From these comments it can be seen that (3.2) can be regarded as the model which, in the analysis of covariance, leads to the usual error analysis.

REMARK 3.5. Note that if $\mathbf{R}(T_1) = \mathbf{N}(\Delta_1')$, then the condition on W is $\mathbf{R}(W) = \mathbf{N}(U_1')$ where $U_1 = X_1 T_1$. One could use Remark 2.4 to compute T_1 from Δ_1 and then to compute W from U_1. In connection with Remark 3.4 the choice $W = I - U_1(U_1'U_1)^- U_1'$ has an advantage when $\mathrm{Cov}(Y) = \sigma^2 I$ because then $\mathrm{Cov}(Z) = \sigma^2 W$ and $Wh = h$ for all h in the regression space of model (3.2). By Theorem 2.8 in Zyskind [10] this means that blues and least squares estimators coincide under model (3.2).

If $b'\pi_2$ is estimable, then it is estimable w.r.t. model $M2$. It sometimes happens that the reverse implication is also true, which is convenient because estimability in model $M2$ is generally easier to check. By Theorem 3.3 this occurs when $\mathbf{R}(X_2'W, \Delta_2) = \mathbf{R}(X_2', \Delta_2)$. Applying (2.2) to models (3.2) and $M2$ one can write $\mathbf{r}(X_2'W, \Delta_2) = d_2 + \mathbf{r}(\Delta_2)$ and $\mathbf{r}(X_2', \Delta_2) = r_2 + \mathbf{r}(\Delta_2)$. This shows the equivalence of (a) and (b) in Theorem 3.6 below. The rest of the proof is omitted since Theorem 3.11 is a more general version.

THEOREM 3.6. *The following statements are equivalent*:
(a) $b'\pi_2$ *is estimable* $\Leftrightarrow$ $b'\pi_2$ *is estimable* w.r.t. *model* $M2$.
(b) $r_2 = d_2$.
(c) $r = r_1 + r_2$.
(d) $\Omega_1 \cap \Omega_2 = \{0\}$.
(e) $a'\pi_1$ *is estimable* $\Leftrightarrow$ $a'\pi_1$ *is estimable* w.r.t. *model* $M1$.

It is clear that a sufficient condition for statement (3.6a) to be true is that π_2 is estimable. For an unconstrained model, Dahan and Styan [2] recently gave the conditions that $\Omega_1 \cap \Omega_2 = \{0\}$ and X_2' has full row rank as necessary and sufficient for π_2 to be estimable. The generalization to model (3.1) is that π_2 is estimable if and only if π_2 is estimable w.r.t. model $M2$ and any one of the conditions in Theorem 3.6 is true. Another condition given as (b) in the following proposition can be obtained by applying the paragraph containing (2.2) to model (3.2).

PROPOSITION 3.7.
(a) *If* π_2 *is estimable, then the statements in Theorem* 3.6 *are true.*
(b) *Suppose* π_2 *is* $p_2 \times 1$. *Then* π_2 *is estimable if and only if* $d_2 = p_2 - \mathbf{r}(\Delta_2)$.

One can compute d_2 by applying (2.2) to model (3.2). This is probably unsatisfactory unless the W matrix is desired for additional purposes such as indicated in Remark 3.4. There are, however, other possibilities for determining d_2. Some illustrations of this are given in the examples of the next section. Another formula for d_2 is given in Proposition 3.9 below.

LEMMA 3.8. *Suppose* $G = (G_1, G_2)$ *is a partitioned matrix. If* H *is any matrix satisfying* $\mathbf{R}(H) = \mathbf{N}(G_1')$, *then* $\mathbf{r}(G) = \mathbf{r}(G_1) + \mathbf{r}(H'G_2)$.

PROOF. This is obtained from equation (2.2) by setting $\Delta = G_1$, $X' = G_2$, and $T = H$. ∎

PROPOSITION 3.9. $r = r_1 + d_2$.

PROOF. Let $U = (U_1, U_2)$ where $U_i = X_i T_i$ and $\mathbf{R}(T_i) = \mathbf{N}(\Delta_i')$. Note that $U = XT$ where $T = \text{diag}(T_1, T_2)$ and that $\mathbf{R}(T) = \mathbf{N}(\Delta')$. By (2.2) we get $r = \mathbf{r}(U)$, $r_1 = \mathbf{r}(U_1)$, and $d_2 = \mathbf{r}(W'U_2)$. Since $\mathbf{N}(U_1') = \Omega_1^{\perp}$, the result follows from Lemma 3.8. □

There are several conceivable ways the above results could be utilized. We comment on three of them. If interest is only in π_2, then consideration should be given to using model (3.2). This is particularly so in light of Remark 3.4. If interest is in π_1 and/or π_2, one needs to compute r to determine the residual degrees of freedom. By calculating r in a judicious manner, one should be able to get r_1 and r_2 with little additional work. If $r = r_1 + r_2$, then Theorem 3.6 is applicable. Even if $r \neq r_1 + r_2$, one can still use Proposition 3.9 to determine $d_2 = r - r_1$ (and $d_1 = r - r_2$), which is the numerator degrees of freedom for testing the null hypothesis that all estimable $b'\pi_2$ are zero. Also, d_2 would tell one how many linearly independent estimable $b'\pi_2$ to look for in further estimability investigations. Alternatively, one might first consider determining d_2 via model (3.2) or directly from model (3.1) via the estimability definition of d_2 and then determining r_2. This is particularly appealing when an upper bound is known a priori on r_2 so that if d_2 attains the upper bound one can conclude $d_2 = r_2$ without calculating r_2. Then if $d_2 = r_2$, statement (3.6e) would be useful and, depending on how d_2 was determined, statement (3.6a) might also be useful. Even if $d_2 \neq r_2$, one can still calculate $r = r_1 + d_2$ (and $d_1 = r - r_2$) by only calculating r_1. A natural case where this last approach is taken is in the analysis of covariance. That is, if π_2 denotes the covariate coefficients, then d_2 is the rank of the error normal equations. If $d_2 = r_2$, then estimability considerations for the classification part of the model can be ascertained by disregarding the covariates. Actually, in this situation Proposition 3.7a is usually applicable since, typically, the error normal equations are nonsingular implying that π_2 is estimable.

When Theorem 3.6 is not applicable, a generalized version can sometimes be used. For this, additional notation is needed. Suppose $E(Y)$ has the partitioned form

$$E(Y) = X_0\pi_0 + X_1\pi_1 + X_2\pi_2, \quad \Delta_i'\pi_i = 0, \qquad i = 0, 1, 2. \tag{3.10}$$

Consider the two reduced models

$$\text{model } M01\text{: } E(Y) = X_0\pi_0 + X_1\pi_1, \quad \Delta_0'\pi_0 = 0, \quad \Delta_1'\pi_1 = 0,$$
$$\text{model } M02\text{: } E(Y) = X_0\pi_0 + X_2\pi_2, \quad \Delta_0'\pi_0 = 0, \quad \Delta_2'\pi_2 = 0.$$

Let r_{01} and Ω_{01} denote the rank and regression space of model $M01$. Let r_{02}, Ω_{02}, r_0 and Ω_0 be similarly defined w.r.t. models $M02$ and $E(Y) = X_0\pi_0$, $\Delta_0'\pi_0 = 0$, respectively. Also, let d_2^* denote the dimension of the vector space of parametric functions $b'\pi_2$ that are estimable w.r.t. model $M02$. Let d_2 be similarly defined w.r.t. the full model (3.10).

THEOREM 3.11. *The following statements are equivalent*:
(a) $b'\pi_2$ *is estimable* $\Leftrightarrow$ $b'\pi_2$ *is estimable* w.r.t. *model* $M02$.
(b) $d_2 = d_2^*$.
(c) $r = r_{01} + r_{02} - r_0$.
(d) $\Omega_{01} \cap \Omega_{02} = \Omega_0$.
(e) $a'\pi_1$ *is estimable* $\Leftrightarrow$ $a'\pi_1$ *is estimable* w.r.t. *model* $M01$.

PROOF. Let $\mathcal{W}$ (resp., $\mathcal{W}^*$) denote the set of vectors b such that $b'\pi_2$ is estimable (resp., estimable w.r.t. model $M02$). Similar to the paragraph following Remark 3.5, write $\dim \mathcal{W} = d_2 + \mathbf{r}(\Delta_2)$ and $\dim \mathcal{W}^* = d_2^* + \mathbf{r}(\Delta_2)$. Since $\mathcal{W} \subset \mathcal{W}^*$, we get (a)$\Leftrightarrow$(b). From Proposition 3.9 we can write $r = r_{01} + d_2$ and $r_{02} = r_0 + d_2^*$ which gives (b)$\Leftrightarrow$(c). Using $\Omega = \Omega_{01} + \Omega_{02}$ and $\Omega_0 \subset \Omega_{01} \cap \Omega_{02}$, one can get (c)$\Leftrightarrow$(d). The symmetry of (d) and (a)$\Leftrightarrow$(d) then gives (d)$\Leftrightarrow$(e). □

Notice that the above theorem reduces to Theorem 3.6 when $X_0 = 0$. Also, notice that the comments following Proposition 3.9 are applicable (with minor adaptations) to the above theorem. Finally, statement (3.11d) sometimes provides insight as how to partition $E(Y)$ before attempting to utilize the theorem; e.g., in classification models it is often obvious that one reduced model regression space is in the intersection of two other reduced model regression spaces.

4. Examples. To illustrate some of the results in Section 3, consider first the unconstrained model

$$E(Y_{ijkp}) = \mu + \alpha_i + \beta_j + \gamma_k \tag{4.1}$$

where $i = 1, \cdots, a, j = 1, \cdots, b, k = 1, \cdots, c$, and $p = 1, \cdots, n_{ijk}$. Here n_{ijk} is the number (possibly 0) of replications at levels (i, j, k) of the factors.

Theorem 3.6 will not be useful for any of the natural partitions because (3.6d) cannot be true. If it is known a priori that the model is of maximal rank $a + b + c - 2$ (e.g., all $n_{ijk} \neq 0$ or the n_{ijk} have a Latin square pattern), then estimability facts are known so there is no need for Theorem 3.11. When the rank of the model is unknown, however, Theorem 3.11 (and Proposition 3.9) can sometimes be useful. Clearly Theorem 3.11 cannot be applied without some preliminary work. The work required, however, is work which must eventually be done to completely analyse the model. Thus, it is suggested that an analysis of the model begin with a check of one of the conditions in the theorem. If the theorem is found applicable, then later work is simplified; and if the theorem is not applicable, nothing has been lost. For example, consider the reduced models

$$\text{model } M01\text{: } E(Y_{ijkp}) = \mu + \alpha_i + \beta_j$$

$$\text{model } M02\text{: } E(Y_{ijkp}) = \mu + \gamma_k.$$

Let $d_\gamma = d_2$ and $r_{\alpha\beta} = r_{01}$. A spanning set for Γ_γ can be obtained as in Example 10.1 of Birkes et al. [1]. So d_γ can be determined by extracting a basis from the spanning set. If $d_\gamma = c - 1$, then Theorem 3.11 implies (since $d_2 \leqslant d_2^* \leqslant c - 1$) that estimability of α and β contrasts can be determined via model $M01$. In any event, Proposition 3.9 implies $r = r_{\alpha\beta} + d_\gamma$, $d_\alpha = r - r_{\beta\gamma}$, and $d_\beta = r - r_{\alpha\gamma}$ can be determined by calculating $r_{\alpha\beta}$, $r_{\alpha\gamma}$ and $r_{\beta\gamma}$, which are straightforward to obtain; e.g., see Section 4 of [1].

As a second example, consider (4.1) with an $\alpha\beta$ interaction term, that is,

$$E(Y_{ijkp}) = \mu + \alpha_i + \beta_j + \theta_{ij} + \gamma_k \tag{4.2}$$

where i, j, k, p are as before. Consider the reduced models

$$\text{model } M01\text{: } E(Y_{ijkp}) = \mu + \theta_{ij} + \gamma_k$$

$$\text{model } M02\text{: } E(Y_{ijkp}) = \mu + \theta_{ij} + \alpha_i + \beta_j.$$

It is immediate to see $r = r_{01}$ and $r_{02} = r_0$. Thus, Theorem 3.11 implies estimability for the γ_k can be ascertained from model $M01$. Because model $M01$ is an $ab \times c$ additive two-way model, the estimable functions involving the γ_k and the dimension d_γ of the vector space of such estimable functions are straightforward to determine; e.g., see Section 4 of [1]. Then $r = r_{02} + d_\gamma$ can be determined since r_{02} is the number of nonzero $n_{ij} = \Sigma_k n_{ijk}$. Next consider the reduced models

$$\text{model } M01\text{: } E(Y_{ijkp}) = \mu + \alpha_i + \beta_j + \theta_{ij}$$

$$\text{model } M02\text{: } E(Y_{ijkp}) = \mu + \gamma_k.$$

Now $d_2 = d_\gamma$ can be determined as above so that $d_2 = d_2^*$ can be checked to see if estimability for the α, β and θ contrasts can be deduced from model $M01$.

As a final example consider model $(4.2)_c$ which is the same as model (4.2) except that the parameters are constrained by $\alpha. = \beta. = \gamma. = 0$ and $\theta_{i.} = \theta_{.j} = 0$ for all i, j. Let d_γ be defined in the obvious manner. Because of the constraints, it is possible that Theorem 3.6 is applicable. Consider the partition leading to model $M2$: $E(Y_{ijkp}) = \gamma_k$, $\gamma. = 0$. Proposition 3.7b implies $\gamma = (\gamma_1, \cdots, \gamma_c)'$ is estimable if and only if $d_\gamma = c - 1$, in which case Proposition 3.7a would imply that Theorem 3.6 could be invoked. To check this d_γ must be determined. One possibility is to compute d_γ as in model (4.2) since d_γ is the same for models (4.2) and $(4.2)_c$. (This is also true for r and d_θ but not for d_α and d_β.) Other partitions are also possible. For example, consider model $M2$: $E(Y_{ijkp}) = \theta_{ij}$, $\theta_{i.} = \theta_{.j} = 0$. Proposition 3.7b implies $\theta = (\theta_{11}, \cdots, \theta_{ab})'$ is estimable if and only if $d_\theta = ab - (a + b - 1) = (a - 1)(b - 1)$, in which case Theorem 3.6 can be used. A possibility for computing d_θ is to compute the ranks for models (4.2) and (4.1) as previously described and then take the difference of these ranks.

Acknowledgment. The authors wish to thank the referees for their helpful comments.

REFERENCES

[1] BIRKES, D., DODGE, Y. and SEELY, J. (1976). Spanning sets for estimable contrasts in classification models. *Ann. Statist.* **4** 86–107.

[2] DAHAN, S. and STYAN, G. P. H. (1977). Estimability of a proper subset of the regression coefficients in the general linear model with less than full rank. (abstract) *IMS Bulletin* **6** 48.

[3] ECCLESTON, J. and RUSSELL, K. (1975). Connectedness and orthogonality in multi-factor designs. *Biometrika* **62** 341–345.

[4] RAGHAVARAO, D. and FEDERER, W. T. (1975). On connectedness in two-way elimination of heterogeneity designs. *Ann. Statist.* **3** 730–735.

[5] RAO, C. R. (1946). On the linear combination of observations and the general theory of least squares. *Sankhyā* **7** 237–256.

[6] SEARLE, S. R. (1971). *Linear Models*. Wiley.

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

[8] SEELY, J. (1977). Estimability and linear hypotheses. *Amer. Statist.* **31** 121–123.

[9] WYNN, H. P. (1976). A note on estimability in multi-way layouts. Unpublished manuscript.

[10] ZYSKIND, G. (1967). On canonical forms, nonnegative covariance matrices and best and simple least squares linear estimators in linear models. *Ann. Math. Statist.* **38** 1092–1109.

[11] ZYSKIND, G., KEMPTHORNE, O., WHITE, R. F., DAYHOFF, E. E. and DOERFLER, T. R. (1964). Research on analysis of variance and related topics. Aerospace Research Laboratories, Technical Report 64–193, Wright-Patterson Air Force Base, Ohio.

[12] ZYSKIND, G. and MARTIN, F. (1969). On best linear estimation and a general Gauss-Markov theorem in linear models with arbitrary nonnegative covariance structure. *SIAM J. Appl. Math.* **17** 1190–1202.

DEPARTMENT OF STATISTICS
OREGON STATE UNIVERSITY
CORVALLIS, OREGON 97331