ON THE EXTENSION OF GAUSS-MARKOV THEOREM TO COMPLEX MULTIVARIATE LINEAR MODELS*

J. N. SRIVASTAVA

(Received Aug. 29, 1966; Corrected Oct. 13, 1967)

Summary

The purpose of this paper is to develop a theory of linear estimation under various multivariate linear models, which are more general than the usual model to which the standard techniques of multivariate analysis of variance are applicable. In particular, necessary and sufficient conditions under which (unique) best linear unbiased estimates of linear functions of (location) parameters exist are obtained. An extension of the Gauss-Markov theorem to the standard multivariate model was first made by the author in [13]. In this paper, further generalizations of the result to multiresponse designs where the standard technique is inapplicable are considered.

1. Introduction

Multivariate linear models more general than the standard one have been considered in various earlier communications. See, for example, Trawinski [18], Trawinski and Bargmann [19] and Srivastava [14], [15], [17]. However, it will be necessary here to recall them explicitly, and the physical situations where they arise.

We shall assume throughout that (i) there are n experimental units in all, (ii) p responses (or characteristics or variables) $V_1, V_2, \cdots, V_p$ under study, and (iii) any pair of observations arising on *distinct* experimental units are statistically independent. The standard multivariate (SM) model becomes applicable when each response is measured on each unit, resulting in an $(n\times p)$ observation matrix Y (whose rth column $\boldsymbol{y}_r$ $(n\times 1)$ corresponds to observations on response V_r), with means and variances given by

$$\text{(1.1a)} \qquad \mathrm{E}(Y)=A\xi ,$$

* This research was wholly supported by U.S. Air Force under Grant No. AF 33 (615)-3231, monitored by the Aerospace Research Laboratories.

(1.1b) $$\text{Var}(Y)=I_n\otimes\Sigma\ ,$$

where A $(n\times m)$ is the usual design-model matrix, ξ $(m\times p)$ is a matrix of unknown parameters, the rth column $\boldsymbol{\xi}_r$ of which corresponds to V_r, I_n is the $(n\times n)$ identity matrix, and $\Sigma(p\times p)$ is an unknown dispersion matrix. Also, the symbol $\otimes$ denotes Kronecker product, $I_n\otimes\Sigma$ being an $(np\times np)$ matrix partitioned into $(p\times p)$ submatrices, such that each diagonal submatrix is Σ, and each off diagonal one is zero. Let $\boldsymbol{y}_\alpha^{*\prime}\times(p\times 1)$, the αth row of Y, be the vector of the p measurements on the αth unit. Then the above can be re-expressed as

(1.2a) $$\text{E}(\boldsymbol{y}_r)=A\boldsymbol{\xi}_r\ ,$$

(1.2b) $$\text{Var}(\boldsymbol{y}_\alpha^{*\prime})=\Sigma\ ,\qquad \text{Cov}(\boldsymbol{y}_\alpha^{*\prime},\boldsymbol{y}_\beta^{*})=O_{pp}\ ,$$

(1.2c) $$\text{Var}(\boldsymbol{y}_r)=\sigma_{rr}I_n\ ,\quad \text{Cov}(\boldsymbol{y}_r,\boldsymbol{y}_s)=\sigma_{rs}I_n\ ,$$

where O_{pq} is the $(p\times q)$ zero-matrix, and σ_{rs} is the (r,s)th element of Σ.

Coming next to the general incomplete multivariate (GIM) model, let the set S of the n experimental units be divided into u *disjoint* subsets $S_1, S_2,\cdots,S_u$, the set S_i having n_i units. Furthermore, we suppose that there are some units in S on which all the p responses are not measured, the set of responses studied on *each* of the n_i units in S_i being $V_{l_{i1}}, V_{l_{i2}},\cdots,V_{l_{iq_i}}$, where $1\leqq q_i\leqq p$, and l_{ij} are integers between 1 and p such that $l_{ij}<l_{ij'}$, if $j<j'$. (As an example, the set of characteristics observed on each unit may be V_1, V_2 and V_5 for S_1, and V_1, V_5, V_6 and V_7 for S_2, and so on.) Let B_i $(i=1,2,\cdots,u)$ be a $(p\times q_i)$ matrix with unity in the cells (l_{ij}, j), $(j=1,2,\cdots,q_i)$, and zero elsewhere. Notice that the model SM is applicable to each S_i. Let Y_i $(n_i\times q_i)$ be the observation matrix from S_i, and $\boldsymbol{y}_{i\alpha}^{*\prime}$ $(p_i\times 1)$, the αth row of Y_i, be the observations from the αth unit in S_i. Under the model GIM, we then suppose:

(1.3a) $$\text{E}(Y_i)=A_i\xi B_i\ ,$$

(1.3b) $$\text{Var}(\boldsymbol{y}_{i\alpha}^{*\prime})=B_i'\Sigma B_i\ ,\qquad \alpha=1,\cdots,n_i\ ;\quad i=1,\cdots,u\ ,$$

where ξ and Σ are as before, and A_i is the design-model matrix for S_i.

In the above model, let U_r $(r=1,2,\cdots,p)$ denote the subset of S, such that on each unit in U_r, the variable V_r is observed. Notice that if $r\neq s$, we do not necessarily require U_r to be a subset of U_s or vice versa. However, suppose there exists a grading of the p responses, say in the order $(V_{k_1}, V_{k_2},\cdots,V_{k_p})$ where $(k_1,\cdots,k_p)$ is a permutation of $(1,2,\cdots,p)$, and that we require that

(1.4) $$U_{k_1}\supseteqq U_{k_2}\supseteqq\cdots\supseteqq U_{k_p}\ .$$

In this situation, we get a special case of the GIM model, called the hierarchical multivariate (HM) model. Under this model, the response V_{k_r} is measured on any unit in S on which $V_{k_{r+1}}$ is observed, and therefore a kind of hierarchy exists.

Finally, we consider the multiple design multivariate (MDM) linear model. This is not a special case of the above models. The physical situation where this model arises is such that although each response is measured on each experimental unit, the SM model is still inapplicable. Analytically, the MDM model assumes:

$$\text{(1.5a)} \qquad \mathrm{E}(\boldsymbol{y}_r)=A_r\boldsymbol{\xi}_r ,$$

$$\text{(1.5b)} \qquad \mathrm{Var}\,(\boldsymbol{y}_r)=\sigma_{rr}I_n , \qquad \mathrm{Cov}\,(\boldsymbol{y}_r, \boldsymbol{y}_s)=\sigma_{rs}I_n , \qquad r\neq s ;\ r, s=1, 2, \cdots, p ,$$

where A_r $(n\times m)$ is the design-model matrix for the rth response V_r, and the other symbols have the same meaning as in equations (1.2). Clearly, the only difference between the SM and MDM models is that under the latter, the design-model matrices $A_1, \cdots, A_p$ are not necessarily equal. This justifies the name 'multiple-design.'

Examples illustrating the above models will be found in many papers referred to at the end. The standard model, for example, is described in Roy [9] and Anderson [1], and an application in Smith, Gnanadesikan and Hughes [12]. Examples of the GIM will be found in Trawinski [18] and Trawinski and Bargmann [19], where maximum likelihood estimation and likelihood ratio tests based on the assumption of normality are considered. In [15], Srivastava developed a class of designs under GIM, where SM would be applicable after a linear transformation of the data, illustrating the method by an example worked out in [17]. The HM model was introduced by Roy and Srivastava in [10], where a generalization of the step-down procedure was suggested for testing linear hypotheses. In [14], the author considered certain features of the above models and the designs under them, and gave an example where the MDM model arises.

In view of the above papers and for the sake of brevity, further examples will be omitted here. However, it may be appropriate to remark that in general, the GIM model arises whenever it is physically impossible (for example, when measuring a certain response involves destruction of the experimental unit), uneconomic (because of the unequal measuring costs), or inadvisable (because of unequal importance of the different responses, or because of fast changing experimental conditions), to observe each of them on each experimental unit. The special case of the HM model would arise when, because of these considerations, it is possible to decide for each pair of responses as to which should be measured on a larger number of units. Finally, the MDM model can arise, for example, in a block-treatment set-up, where some responses are known, *a priori*, to be insensitive to certain treatments, or to be affected differentially by different treatments.

For an introduction to linear estimation and Gauss-Markov theorem in the univariate case, the reader is referred, for example, to Scheffe [11], Bose [2] and Kempthorne [5]. The generalization of this theorem to the multivariate case under the SM model was considered by the author in [13]. In this paper, the theory is extended to the other models.

2. The hierarchical multivariate model

Although this model is a special case of the GIM model, to be considered in the next section, its discussion first may aid in the clarity of presentation.

Consider equations (1.3) and (1.4) and the definition of the HM model. It clearly implies that $u=p$. Without loss of generality, and for each of presentation, we take $(k_1, \cdots, k_p)$ to be identical with $(1, 2, \cdots, p)$. Then $q_r=r$; Y_r $(n_r \times r)$ is the observation matrix for S_r $(r=1, \cdots, p)$; and $\boldsymbol{y}_{rs}$ $(n_r \times 1)$, the sth column $(s=1, \cdots, r)$ or Y_r, represents the set of observations on the sth response from the n_r units in S_r. Then (1.3) specializes to

$$\text{(2.1a)} \qquad \mathrm{E}(Y_r)=\mathrm{E}(\boldsymbol{y}_{r1}, \boldsymbol{y}_{r2}, \cdots, \boldsymbol{y}_{rr})=A_r(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_r) ,$$

$$\text{(2.1b)} \qquad \mathrm{Var}\,(Y_r)=\sigma_{rr}I_{n_r}\otimes \Sigma_r ,$$

where

$$\text{(2.2)} \qquad \Sigma_r=\begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1r} \\ \cdot & \cdots & \cdot \\ \sigma_{r1} & \cdots & \sigma_{rr} \end{bmatrix} .$$

Consider the problem of estimating a linear function of the unknown parameters ξ

$$\text{(2.3)} \qquad \theta=\boldsymbol{c}_1'\boldsymbol{\xi}_1+\boldsymbol{c}_2'\boldsymbol{\xi}_2+ \cdots +\boldsymbol{c}_p'\boldsymbol{\xi}_p=\sum_{s=1}^{p} \boldsymbol{c}_s'\boldsymbol{\xi}_s ,$$

where $\boldsymbol{c}_r$ $(m \times 1)$, $(r=1, \cdots, p)$ is a given vector of coefficients. Throughout this paper, we shall restrict ourselves to unbiased linear estimates, and call a linear function θ *estimable* if there exists an unbiased linear estimate of θ. Let

$$\text{(2.4)} \qquad l=\sum_{r=1}^{p} \sum_{s=1}^{r} \boldsymbol{b}_{rs}'\boldsymbol{y}_{rs}$$

be any unbiased estimate of θ, where $\boldsymbol{b}_{rs}$ $(n_r \times 1)$, $r=1, \cdots, p$; $s=1, \cdots, r$ are vectors of constants. Since $\mathrm{E}(\boldsymbol{y}_{rs})=A_r\boldsymbol{\xi}_s$, we get

$$\mathrm{E}(l)=\sum_{r=1}^{p}\sum_{s=1}^{r}\boldsymbol{b}'_{rs}A_r\boldsymbol{\xi}_s \tag{2.5}$$
$$=\sum_{s=1}^{p}\sum_{r=s}^{p}\boldsymbol{b}'_{rs}A_r\boldsymbol{\xi}_s=\sum_{s=1}^{p}\boldsymbol{c}'_s\boldsymbol{\xi}_s .$$

To avoid triviality, we assume throughout that the space of the parameters ξ contains a subset of mp linearly independent points. Then (2.5) gives, after equating coefficients of $\boldsymbol{\xi}_s$ on both sides:

$$\boldsymbol{c}_s=\sum_{r=s}^{p}A'_r\boldsymbol{b}_{rs} , \qquad s=1, \cdots, p . \tag{2.6}$$

Define B_s $(m_s \times m)$ by

$$B'_s=[A'_s|A'_{s+1}| \cdots |A'_p] , \qquad s=1, \cdots, p , \tag{2.7}$$

where $m_s=n_s+n_{s+1}+ \cdots +n_p$. Then (2.6) implies

THEOREM 2.1. *A necessary and sufficient condition (n. & s.c.) that θ is estimable is that $\boldsymbol{c}_s$ belongs to the row space of B_s, for all permissible s.*

We next consider best (i.e., minimum variance) unbiased estimates. Define x'_s $(1 \times m_s)$ by

$$\boldsymbol{x}'_s=[\boldsymbol{y}'_{ss}, \boldsymbol{y}'_{s+1,s}, \cdots, \boldsymbol{y}'_{ps}] . \tag{2.8}$$

Then x_s is the vector of the total of m_s observations on V_s from all units in S, and it can be verified that

$$\mathrm{E}(\boldsymbol{x}_s)=B_s\boldsymbol{\xi}_s \tag{2.9a}$$

$$\mathrm{Var}(\boldsymbol{x}_s)=\sigma_{ss}I_{m_s} , \qquad s=1, \cdots, p \tag{2.9b}$$

$$\mathrm{Cov}(\boldsymbol{x}'_r, \boldsymbol{x}_s)=\sigma_{rs}[O_{m_r, m_s-m_r} \mid I_{m_r}] , \tag{2.9c}$$
$$r>s ; \quad r, s=1, \cdots, p .$$

The set-wise and response-wise representations of the HM model by equations (2.1) and (2.9) respectively are clearly equivalent. Consider (2.9) and suppose we ignore all responses except V_s. Then (2.9a, b) gives us the univariate model (M_s say) for V_s. Assume now that for $s=1, \cdots, p$, the (unique) best linear unbiased estimate (BLUE) of $\boldsymbol{c}'_s\boldsymbol{\xi}_s$ exists under M_s, and equals $\boldsymbol{g}'_s\boldsymbol{x}_s$, where $\boldsymbol{g}_s$ $(m_s \times 1)$ is a vector of constants. From Scheffe [11] or Bose [2], we then recall (without proof):

THEOREM 2.2. *A necessary and sufficient condition that $\boldsymbol{c}'_s\boldsymbol{\xi}_s$ is estimable under M_s is:*

(2.10) $$\mathrm{Rank}\,(B_s') = \mathrm{Rank}\,(B_s' \mid \boldsymbol{c}_s)\ .$$

Furthermore, if $\boldsymbol{g}_s'\boldsymbol{x}_s$ *is the BLUE for* $\boldsymbol{c}_s'\boldsymbol{c}_s$, *then* $\boldsymbol{g}_s$ *belongs to the estimation space; in other words,*

$$\mathrm{Rank}\,(B_s) = \mathrm{Rank}\,(B_s \mid \boldsymbol{g}_s)\ .$$

Let

(2.11) $$l_0 = \sum_{s=1}^{p} \boldsymbol{g}_s'\boldsymbol{x}_s\ ,$$

be the sum of the individual estimates. We proceed to compare the variances of l_0 and l. Rewrite l in the form

(2.12) $$l = \sum_{s=1}^{p} \boldsymbol{d}_s'\boldsymbol{x}_s\ ,$$

where

(2.13) $$\boldsymbol{d}_s' = [\boldsymbol{b}_{ss}', \boldsymbol{b}_{s+1,s}', \cdots, \boldsymbol{b}_{ps}']\ .$$

Then

(2.14) $$\begin{aligned}\mathrm{Var}\,(l) &= \mathrm{Var}\left(\sum_{s=1}^{p} \boldsymbol{d}_s'\boldsymbol{x}_s\right)\\ &= \sum_{s=1}^{p} \boldsymbol{d}_s'[\mathrm{Var}\,(\boldsymbol{x}_s)]\,\boldsymbol{d}_s + 2\sum_{r>s} \boldsymbol{d}_r'\ \mathrm{Cov}\,(\boldsymbol{x}_r', \boldsymbol{x}_s)\,\boldsymbol{d}_s\\ &= \sum_{s=1}^{p} (\boldsymbol{d}_s'\boldsymbol{d}_s)\sigma_{ss} + 2\sum_{r>s} \{\boldsymbol{d}_r'[O_{m_r, m_s - m_r} \mid I_{m_r}]\,\boldsymbol{d}_s\}\sigma_{rs}\ .\end{aligned}$$

A similar expression holds for $\mathrm{Var}\,(l_0)$. Also

$$\theta = \mathrm{E}(l) = \sum_{s=1}^{p} \boldsymbol{d}_s' B_s \boldsymbol{\xi}_s = \mathrm{E}(l_0) = \sum_{s=1}^{p} \boldsymbol{g}_s' B_s \boldsymbol{\xi}_s\ .$$

Hence

(2.15) $$(\boldsymbol{d}_s' - \boldsymbol{g}_s')B_s = O_{1m}\ , \qquad s = 1, 2, \cdots, p\ .$$

Let W_s be the column space of B_s, and $\overline{W}_s$ the space orthogonal to W_s. Then (3.15) implies that $(\boldsymbol{d}_s - \boldsymbol{g}_s)$ belongs to $\overline{W}_s$. Let

(2.16) $$\mathrm{Rank}\,(B_s) = \rho_s\ .$$

Hence

(2.17) $$\mathrm{Dim}\,(W_s) = \rho_s\ , \qquad \mathrm{Dim}\,(\overline{W}_s) = m_s - \rho_s = \mu_s\ , \qquad \text{say.}$$

Let $\boldsymbol{\theta}_{s1}, \boldsymbol{\theta}_{s2}, \cdots, \boldsymbol{\theta}_{s\mu_s}$ be an orthogonal basis of $\overline{W}_s$. Then there exist constants $h_{s1}, \cdots, h_{s\mu_s}$ such that

$$d_s - g_s = h_{s1}\theta_{s1} + \cdots + h_{s\mu_s}\theta_{s\mu_s}\,, \qquad s=1,\cdots,p\,. \tag{2.18}$$

From theorem 2.2, $g_s \in W_s$, and hence g_s is orthogonal to each θ_{si}. Thus we get, for $r>s$,

$$(d_s'd_s) = (g_s'g_s) + h_{s1}^2 + h_{s2}^2 + \cdots + h_{s\mu_s}^2\,, \tag{2.19}$$

and

$$\begin{aligned} d_r'[O_{m_r,m_s-m_r} \mid I_{m_r}]\, d_s &= [g_r' + h_{r1}\theta_{r1}' + \cdots + h_{r\mu_r}\theta_{r\mu_r}'][O_{m_r,m_s-m_r} \mid I_{m_r}] \\ &\quad \times [g_s + h_{s1}\theta_{s1} + \cdots + h_{s\mu_s}\theta_{s\mu_s}] \\ &= [O_{1,m_s-m_r} \mid g_r' + h_{r1}\theta_{r1} + \cdots + h_{r\mu_r}\theta_{r\mu_r}'] \\ &\quad \times [g_s + h_{s1}\theta_{s1} + \cdots + h_{s\mu_s}\theta_{s\mu_s}] \\ &= [g_r' + h_{r1}\theta_{r1}' + \cdots + h_{r\mu_r}\theta_{r\mu_r}'][g_s^{(r)} + h_{s1}\theta_{s1}^{(r)} + \cdots + h_{s\mu_s}\theta_{s\mu_s}^{(r)}]\,, \end{aligned} \tag{2.20}$$

where (r) in the *superscripts* in the vectors in the second bracket denotes that the first part of the corresponding original vector, involving $(m_j - m_i)$ coordinates, has been omitted. But

$$\mathrm{Var}\,(l) - \mathrm{Var}\,(l_0) = \sum_{s=1}^{p} \omega_{ss}\sigma_{ss} + 2\sum_{r>s} \omega_{rs}\sigma_{rs}\,, \tag{2.21}$$

where, using (2.19) and (2.20)

$$\omega_{ss} = d_s'd_s - g_s'g_s = \sum_{j=1}^{\mu_s} h_{sj}^2 \tag{2.22}$$

$$\begin{aligned} \omega_{rs} &= [d_r'(O_{m_r,m_s-m_r} \mid I_{m_r})\, d_s - g_r'(O_{m_r,m_s-m_r} \mid I_{m_r}) g_s] \\ &= g_r'\left(\sum_{i=1}^{\mu_s} h_{si}\theta_{si}^{(r)}\right) + g_s^{(r)\prime}\left(\sum_{j=1}^{\mu_r} h_{rj}\theta_{rj}\right) \\ &\qquad + \left(\sum_{j=1}^{\mu_r} h_{rj}\theta_{rj}'\right)\left(\sum_{i=1}^{\mu_s} h_{si}\theta_{si}^{(r)}\right). \end{aligned} \tag{2.23}$$

Let Ω be the $(p\times p)$ matrix (ω_{rs}). Then the Schur's product $(\Omega^*\Sigma)$ of Ω and $\Sigma=(\sigma_{rs})$ is the $(p\times p)$ matrix whose (r, s) cell has the element $\omega_{rs}\sigma_{rs}$. Let J_{pq} be the $(p\times q)$ matrix having unity everywhere. Then the above results can be noted in

LEMMA 2.1. *We have*

$$\mathrm{Var}\,(l) - \mathrm{Var}\,(l_0) = J_{1p}[\Sigma^*\Omega]J_{p1} = u\,, \tag{2.24}$$

where the elements of Ω are given by (2.22) *and* (2.23).

Next, we investigate the conditions under which the difference u, between the variances of l and l_0 remains non-negative under the variation of Σ, subject to the condition that Σ be positive definite. The last requirement is obviously necessary, for non-triviality. We prove

LEMMA 3.2. *A n. & s.c. that* $u \geqq 0$, *for all values of* Σ *(subject to* Σ *being positive definite (p.d.)) is that* Ω *be positive semi-definite (p.s.d.).*

PROOF. (i) *Sufficiency.* The Schur's product $(\Omega^* \Sigma)$ is a principal $(p \times p)$ submatrix of the Kronecker product $(\Omega \otimes \Sigma)$. Since the Kronecker product of a p.d. matrix with a p.s.d. matrix is a p.s.d. matrix, and since a principal submatrix of a p.s.d. matrix is also p.s.d., it follows that $(\Omega^* \Sigma)$ is p.s.d. Since u is a quadratic form in $(\Omega^* \Sigma)$, we have therefore $u \geqq 0$.
(ii) *Necessity.* Suppose Ω is not p.s.d. Let $\boldsymbol{v}$ be a vector (with ith element v_i) such that $\boldsymbol{v}' \Omega \boldsymbol{v} < 0$. Let $\Sigma = (\boldsymbol{v}\boldsymbol{v}' + \varepsilon I_p)$. Then clearly, Σ is p.d. for all values of $\varepsilon > 0$, howsoever small. But then,

$$(\Omega^* \Sigma) = ((\pi_{rs})) ,$$

where

$$\pi_{rs} = \begin{cases} v_r v_s \omega_{rs} , & \text{if } r \neq s \\ v_r^2 \omega_{rr} + \varepsilon \omega_{rr} , & \text{if } r = s . \end{cases}$$

Hence

$$\begin{aligned} u &= \sum_{r,s=1}^{p} \pi_{rs} \\ &= \sum_{r,s=1}^{p} \omega_{rs} v_r v_s + \sum_{r=1}^{p} \varepsilon \omega_{rr} \\ &= \varepsilon \left(\sum_{r=1}^{p} \omega_{rr} \right) + (\boldsymbol{v}' \Omega \boldsymbol{v}) . \end{aligned}$$

Since $(\boldsymbol{v}' \Omega \boldsymbol{v}) < 0$, we find that $u < 0$, for sufficiently small values of ε. This completes the proof.

If l_1 and l_2 are two estimates of θ, we shall say, in accordance with the usual convention, that l_1 is as good as l_2 if $\operatorname{Var}(l_2) - \operatorname{Var}(l_1) \geqq 0$, for all values of Σ, subject to the condition of being p.d. If in addition to this, strict inequality holds for some value of Σ, then l_1 is said to be better than l_2.

Comparing l_0 and l, the last lemma shows that l_0 is better than l, provided Ω is p.s.d. If Ω is indefinite, then l_0 will have less or more variance compared to l, depending upon the value of Σ. However, since Σ is unknown, one cannot decide whether l_0 or l should be used. Hence in order that we can choose between l_0 and l, Ω should be a definite matrix. Now, Ω can't be negative definite, since the diagonal elements of Ω (see (2.22)) are all non-negative. Thus in order to choose among l_0 and l, Ω must be p.s.d., in which case the choice will fall on l_0. On the other hand, l itself depends upon the chosen value of the constants

h_{jr} (see (2.17) and (2.18)), and becomes identical with l_0 when all h_{jr} equal zero. For any given Σ, optimal values of h_{si} (say $\hat{h}_{si}$) can be found (by differentiating $\mathrm{Var}(l)$ and equating to zero), such that $\mathrm{Var}(l)$ is minimized. However, it can be easily checked that $\hat{h}_{si}$ depends upon Σ (which is unknown), and therefore cannot be used. The above facts, together with the last two lemmas, lead us to

THEOREM 2.3. *A n. & s.c. that a best linear unbiased estimate (BLUE) of θ may exist (given Σ is unknown) is that Ω possesses the property P_h, namely that Ω be p.s.d., for all values of the constants h_{si} $(i=1, \cdots, \mu_s; s=1, 2, \cdots, p)$. Furthermore, if this is the case, then a BLUE is given by l_0.*

We next prove

LEMMA 2.3. *Suppose, for nontriviality, that Σ is p.d. Also, let l and hence the corresponding h_{sj} be such that Ω is p.s.d. and* $\mathrm{Var}(l)=\mathrm{Var}(l_0)$. *Then Ω has to be a zero matrix, and the h_{sj} must all be zero.*

PROOF. Let $\gamma=\mathrm{Rank}(\Omega)$, and let $d_1, \cdots, d_\gamma$ be the latent roots of Ω. Then there exists an orthogonal matrix P such that $P'\Omega P=D$, where $D=\mathrm{diag}(d_1, d_2, \cdots, d_\gamma; 0, 0, \cdots, 0)$. Let D_i $(i=1, \cdots, \gamma)$ be a diagonal matrix which contains d_i in the cell (i, i) and zero elsewhere. Then $D=\sum_{i=1}^{\gamma} D_i$. Let $\Omega_i=PD_iP'$. Then

$$\Omega=PDP'=\sum_{i=1}^{\gamma}\Omega_i . \tag{2.25}$$

Now Ω_i is clearly of rank 1 since D_i is so. Hence there exist non-zero $(p\times 1)$ vectors $\boldsymbol{f}_i$ $(i=1, \cdots, \gamma)$ such that $\Omega_i=\boldsymbol{f}_i\boldsymbol{f}_i'$. Let $(\Sigma^*\Omega)=\Pi$, and $(\Sigma^*\Omega_i)=\Pi_i$, where $*$ denotes Schur's product. Clearly, we have from (2.25), $\Pi=\sum_{i=1}^{\gamma}\Pi_i$. Let π_{irs} be the element in the cell (r, s) of Π_i. Then $\pi_{irs}=f_{ir}f_{is}\sigma_{rs}$ where f_{ir} is the rth element in $\boldsymbol{f}_i$. From (2.24),

$$u=\mathrm{Var}(l)-\mathrm{Var}(l_0)=J_{1p}\Pi J_{p1}=\sum_{i=1}^{\gamma}[J_{1p}\Pi_iJ_{p1}]$$
$$=\sum_{i=1}^{\gamma}\left[\sum_{r,s=1}^{p} f_{ir}f_{is}\sigma_{rs}\right]=\sum_{i=1}^{\gamma}[\boldsymbol{f}_i'\Sigma\boldsymbol{f}_i] .$$

Since Σ is p.d., $\boldsymbol{f}_i'\Sigma\boldsymbol{f}_i\geqq 0$ for all i. Hence $u=0$ implies that for all i, $\boldsymbol{f}_i$ is the zero vector and hence Ω_i is the zero matrix. This completes the proof.

THEOREM 2.4. *If l is such that Ω is p.s.d., then* $\mathrm{Var}(l)>\mathrm{Var}(l_0)$. *Furthermore, if Ω has the property P_h, then l_0 is the unique BLUE for θ.*

The proof follows directly from the last lemma, since $\mathrm{Var}(l)=\mathrm{Var}(l_0)$ implies Ω is a zero matrix, which means all the h_{si} are zero so that $l=l_0$.

The development in this section shows that the whole matter of the existence and uniqueness of a BLUE estimate hinge on Ω having the property P_h. We now proceed to examine the conditions under which this happens.

THEOREM 2.5. *A n. & s.c. that Ω has the property P_h is that for $i=1, \cdots, \mu_s$; $j=1, \cdots, \mu_r$, and all $r>s$, with $r, s=1, \cdots, p$; we have*

(2.26) (a) $\boldsymbol{g}_r$ *is orthogonal to* $\boldsymbol{\theta}_{si}^{(r)}$,

(2.27) (b) $\boldsymbol{g}_s^{(r)}$ *is orthogonal to* $\boldsymbol{\theta}_{rj}$.

PROOF. (i) *Necessity.* Consider first the case $p=2$. We have

$$\omega_{11}=\sum_{j=1}^{\mu_1} h_{1j}^2=\boldsymbol{\delta}_1'\boldsymbol{\delta}_1, \qquad \omega_{22}=\sum_{i=1}^{\mu_2} h_{2i}^2=\boldsymbol{\delta}_2'\boldsymbol{\delta}_2 \tag{2.28}$$

$$\omega_{21}=\boldsymbol{g}_2'\boldsymbol{\delta}_{12}+\boldsymbol{g}_1^{(2)\prime}\boldsymbol{\delta}_2+(\boldsymbol{\delta}_2'\boldsymbol{\delta}_{12}), \tag{2.29}$$

where

$$\boldsymbol{\delta}_2=\sum_{i=1}^{\mu_2} h_{2i}\boldsymbol{\theta}_{2i}, \qquad \boldsymbol{\delta}_{12}=\sum_{j=1}^{\mu_1} h_{1j}\boldsymbol{\theta}_{1j}^{(2)}, \qquad \boldsymbol{\delta}_1=\sum_{j=1}^{\mu_1} h_{1j}\boldsymbol{\theta}_{1j}. \tag{2.30}$$

If either of the conditions (a) and (b) is violated, there will exist a value of the constant h_{si} (say $h_{si}=h_{si}^*$) such that $z\equiv\boldsymbol{g}_2'\boldsymbol{\delta}_{12}+\boldsymbol{\delta}_1^{(2)\prime}\boldsymbol{\delta}_{12}\neq 0$. Also then, let the values of ω_{11}, etc. be: $\omega_{11}=x$, $\omega_{22}=y$, $\omega_{21}=z+z_1$. Now consider taking $h_{si}=\varepsilon h_{si}^*$, where $\varepsilon>0$. Then the new values of ω_{rs} will be: $\omega_{11}=\varepsilon^2 x$, $\omega_{22}=\varepsilon^2 y$, $\omega_{21}=\varepsilon z+\varepsilon^2 z_1$, so that

$$\begin{aligned}|\Omega|&=\omega_{11}\omega_{22}-\omega_{21}^2=\varepsilon^4 xy-\varepsilon^2(z+\varepsilon z_1)^2\\&=\varepsilon^2[xy\varepsilon^2-(z+\varepsilon z_1)^2].\end{aligned}$$

Since z is fixed and non-zero, the last expression will clearly become negative for sufficiently small values of ε.

This proves for $p=2$, Ω has the property P_h only if conditions (a) and (b) hold. The case for general p follows from this by considering the top left-hand (2×2) submatrix of Ω, and recalling that all principal submatrices of a p.s.d. matrix must be p.s.d.

(ii) *Sufficiency.* We consider here only the case $p=2$. This is to build up a background for the proof of theorem 3.1 (to be proved later), of which this is a special case.

For $p=2$ we have under conditions (a) and (b),

$$|\Omega|=(\boldsymbol{\delta}_1'\boldsymbol{\delta}_1)(\boldsymbol{\delta}_2'\boldsymbol{\delta}_2)-(\boldsymbol{\delta}_2'\boldsymbol{\delta}_{12})^2$$
$$=[(\boldsymbol{\delta}_1'\boldsymbol{\delta}_1)(\boldsymbol{\delta}_{12}'\boldsymbol{\delta}_{12})-(\boldsymbol{\delta}_2'\boldsymbol{\delta}_{12})^2]+[(\boldsymbol{\delta}_1'\boldsymbol{\delta}_1)\{(\boldsymbol{\delta}_2'\boldsymbol{\delta}_2)-(\boldsymbol{\delta}_{12}'\boldsymbol{\delta}_{12})\}] .$$

That the first bracket is always non-negative follows from Cauchy-Schwartz inequality. The same property for the second bracket is implied by the fact that $\boldsymbol{\delta}_1'\boldsymbol{\delta}_1 \geqq 0$ always, and that $\boldsymbol{\delta}_{12}$ is obtained from $\boldsymbol{\delta}_2$ by deleting a certain portion. This completes the proof for $p=2$.

We shall return to the hierarchical case after completing the theory under the GIM model.

3. General incomplete multivariate model

Consider the sets $S_1, S_2, \cdots, S_u$. Out of these, let $S_{t_{r1}}, S_{t_{r2}}, \cdots, S_{t_{ru_r}}$ be the sets belonging to the class U_r of sets on which V_r is measured. Here we assume that the integers t_{rk} satisfy $1 \leqq t_{r1} < t_{r2} < \cdots < t_{ru_r} \leqq u$. Thus U_r contains u_r sets. Let $\boldsymbol{x}_{il_{ir}}$ $(n_i \times 1)$, $(r=1, \cdots, q_i)$ be the vector of observations on the set S_i corresponding to the response $V_{l_{ir}}$, such that

$$(3.1) \qquad Y_i=[\boldsymbol{x}_{il_{i1}}, \boldsymbol{x}_{il_{i2}}, \cdots, \boldsymbol{x}_{il_{ip_i}}] , \qquad i=1, \cdots, u .$$

Let $\boldsymbol{y}_r(m_r \times 1)$ be the vector of all observations on V_r, such that

$$(3.2) \qquad \boldsymbol{y}_r' = [\boldsymbol{x}_{t_{r1}r}' \mid \boldsymbol{x}_{t_{r2}r}' \mid \cdots \mid \boldsymbol{x}_{t_{ru_r}r}']$$

$$(3.3) \qquad m_r = n_{t_{r1}} + n_{t_{r2}} + \cdots + n_{t_{ru_r}} , \qquad r=1, \cdots, p .$$

Define $B_r(m_r \times m)$ by

$$(3.4) \qquad B_r' = [A_{t_{r1}}' \mid A_{t_{r2}}' \mid \cdots \mid A_{t_{ru_r}}'] .$$

Then

$$(3.5a) \qquad \mathrm{E}(\boldsymbol{y}_r) = B_r \boldsymbol{\xi}_r ,$$

$$(3.5b) \qquad \mathrm{Var}(\boldsymbol{y}_r) = \sigma_{rr} I_{m_r} , \qquad r=1, \cdots, p .$$

Also,

$$(3.5c) \qquad \mathrm{Var}(\boldsymbol{y}_r, \boldsymbol{y}_s') = \mathrm{Var}\left\{ \begin{bmatrix} \boldsymbol{x}_{t_{r1}r} \\ \vdots \\ \boldsymbol{x}_{t_{ru_r}r} \end{bmatrix} , \; [\boldsymbol{x}_{t_{s1}s}', \cdots, \boldsymbol{x}_{t_{su_s}s}'] \right\}$$
$$= \sigma_{rs} \begin{bmatrix} \Sigma_{11}^{*(rs)} & \cdots & \Sigma_{1u_s}^{*(rs)} \\ \vdots & \cdots & \vdots \\ \Sigma_{u_r1}^{*(rs)} & \cdots & \Sigma_{u_ru_s}^{*(rs)} \end{bmatrix} ,$$

where

$$(3.6)\qquad \Sigma^{*(rs)}_{\alpha\beta}=\begin{cases} O_{n_{t_{r\alpha}},n_{t_{s\beta}}} & \text{if } t_{r\alpha}\neq t_{s\beta}\ , \\ I_{n_{t_{r\alpha}}} & \text{if } t_{r\alpha}=t_{s\beta}\ , \end{cases}$$

$$\alpha=1,\cdots,u_r\ ,\quad \beta=1,\cdots,u_s\ ,\quad r\neq s,\ r,s=1,2,\cdots,p\ .$$

Notice that the model (3.5) is a generalization of the one in (2.9). As in the hierarchical case, we consider the problem of estimating $\theta=\sum_{r=1}^{p} c_r'\xi_r$. As before, let $g_r'y_r$ be the BLUE of $c_r'\xi_r$ from the model (3.5) when all responses except V_r are ignored. Let $l_0=\sum_{r=1}^{p} g_r'y_r$, and $l=\sum_{r=1}^{p} d_r'y_r$ be another unbiased estimate. Write

$$(3.7)\qquad g_r'=[g_{rt_{r1}}'\,|\,g_{rt_{r2}}'\,|\cdots|\,g_{rt_{ru_r}}']\ ,$$

and similarly for d_r. Then, as before,

$$(3.8)\qquad u=\mathrm{Var}\,(l)-\mathrm{Var}\,(l_0)=J_{1p}(\Omega^*\Sigma)J_{p1}\ ,$$

where $\Omega=(\omega_{rs})$ and

$$(3.9)\qquad \omega_{rr}=\sum_{\alpha=1}^{\mu_r}[d_{rt_{r\alpha}}'d_{rt_{r\alpha}}-g_{rt_{r\alpha}}'g_{rt_{r\alpha}}]\ ,$$

$$(3.10)\qquad \omega_{rs}=\sum_{\alpha=1}^{u_r}\sum_{\beta=1}^{u_s}[(d_{rt_{r\alpha}}'\Sigma^{*(rs)}_{t_{r\alpha}t_{s\beta}}d_{st_{s\beta}})-(g_{rt_{r\alpha}}'\Sigma^{*(rs)}_{t_{r\alpha}t_{s\beta}}g_{st_{s\beta}})]\ .$$

Let U_{rs} be the intersection of U_r and U_s, so that if $S_i\in U_{rs}$, then both responses V_r and V_s are measured on each unit in S_i. Let the sets in U_{rs} be $S_{t_{rs1}},\cdots,S_{t_{rsu_{rs}}}$, where

$$(3.11)\qquad 1\leqq t_{rs1}<t_{rs2}<\cdots<t_{rsu_{rs}}\leqq u\ .$$

Analogous to the HM model, and with the same notation, we introduce the spaces W_r and $\overline{W}_r$, and an orthogonal basis θ_{rj} ($j=1,\cdots,\mu_r$) for $\overline{W}_r$; and there exist constants h_{rj} such that

$$(3.12)\qquad \omega_{rr}=\sum_{j=1}^{\mu_r}h_{rj}^2\ ,$$

$$(3.13)\qquad \omega_{rs}=\sum_{\nu=1}^{\mu_{rs}}[d_{rt_{rs\nu}}'d_{st_{rs\nu}}-g_{rt_{rs\nu}}'g_{st_{rs\nu}}]\ .$$

Consider the $(m_r\times 1)$ vector θ_{rj} in $\overline{W}_r$. In view of (3.4), any vector in the column space of B_r can be partitioned into u_r components belonging respectively to the column spaces of $A_{t_{r1}},\cdots,A_{t_{ru_r}}$. Let the component of θ_{rj} corresponding to $A_{t_{rs\nu}}$ be denoted by $\theta_{rjt_{rs\nu}}$. Then

$$(3.14)\qquad \boldsymbol{d}_{rt_{rs\nu}}=\boldsymbol{g}_{rt_{rs\nu}}+\sum_{j=1}^{\mu_r} h_{rj}\boldsymbol{\theta}_{rjt_{rs\nu}}=\boldsymbol{g}_{rt_{rs\nu}}+\boldsymbol{g}^*_{rt_{rs\nu}}\,, \qquad \text{say.}$$

Hence

$$(3.15)\qquad \omega_{rs}=\sum_{\nu=1}^{\mu_{rs}}[\boldsymbol{g}'_{st_{rs\nu}}\boldsymbol{g}^*_{rt_{rs\nu}}+\boldsymbol{g}'_{rt_{rs\nu}}\boldsymbol{\xi}_{st_{rs\nu}}+\boldsymbol{g}^{*\prime}_{rt_{rs\nu}}\boldsymbol{g}^*_{st_{rs\nu}}]\,.$$

As in the case of HM model, one can establish that a necessary and sufficient condition for the existence of a (unique) BLUE for θ is that Ω has the property P_h. We now characterize this property in

THEOREM 3.1. *Under the GIM model, a n. & s.c. that Ω has the property P_h is the condition C^*: $\boldsymbol{g}_s^{(r)}$ be orthogonal to $\boldsymbol{\theta}_{rj}^{(s)}$, for $j=1,\cdots,\mu_r$; $r\neq s$; $r, s=1,\cdots,p$. Here*

$$(3.16)\qquad \boldsymbol{g}_s^{(r)\prime}=[\boldsymbol{g}'_{st_{rs1}},\cdots,\boldsymbol{g}'_{st_{rsu_{rs}}}]$$

$$(3.17)\qquad \boldsymbol{\theta}_{rj}^{(s)}=[\boldsymbol{\theta}'_{rjt_{rs1}},\cdots,\boldsymbol{\theta}_{rjt_{rsu_{rs}}}]\,,$$

i.e., $\boldsymbol{g}_s^{(r)}$ and $\boldsymbol{\theta}_{rj}^{(s)}$ are obtained from $\boldsymbol{g}_s$ and $\boldsymbol{\theta}_{rj}$ respectively by retaining that part which corresponds to both variables V_r and V_s, and discarding the rest.

PROOF. (i) *Necessity.* This part can be proved on the lines of theorem 2.5 and will be omitted.

(ii) *Sufficiency.* Let $\boldsymbol{T}_r=\sum_{i=1}^{\mu_r} h_{ri}\boldsymbol{\theta}_{ri}$. Let $\boldsymbol{T}_{s_0,s_1,\dots,s_k}$ ($1\leqq s_0\leqq p$, $1\leqq s_1<s_2<\cdots<s_k\leqq p$, the integers $s_0, s_1, \cdots, s_k$ are all distinct) be that part of the vector $\boldsymbol{T}_{s_0}$ which corresponds to the sets in which the responses V_{s_0}, V_{s_1}, $\cdots$, V_{s_k} are measured on each experimental unit, while responses other than these may or may not be measured. Similarly, let $\boldsymbol{t}_{s_0,s_1,\dots,s_k}$ be the part of the vector $\boldsymbol{T}_{s_0}$ which corresponds to the sets on which the responses $V_{s_0}, V_{s_1},\cdots, V_{s_k}$, *and only these responses*, are measured on each unit.

Let $\boldsymbol{\lambda}'=(\lambda_1,\cdots,\lambda_p)$ be any real vector, and consider the quadratic form $Q=\boldsymbol{\lambda}'\Omega\boldsymbol{\lambda}$. In order to prove that Ω is p.s.d., it is sufficient to show that $Q\geqq 0$, whatever $\boldsymbol{\lambda}$ may be. Under the condition C^*, we have from (3.12) and (3.15),

$$(3.18)\qquad \omega_{rr}=\left[\sum_{j=1}^{\mu_r} h_{rj}\boldsymbol{\theta}'_{rj}\right]\left[\sum_{j=1}^{\mu_r} h_{rj}\boldsymbol{\theta}_{rj}\right]=\boldsymbol{T}'_r\boldsymbol{T}_r\,,$$

$$(3.19)\qquad \omega_{rs}=\sum_{\nu=1}^{u_{rs}}[\boldsymbol{g}^{*\prime}_{rt_{rs\nu}}\boldsymbol{g}^*_{st_{rs\nu}}]=\boldsymbol{T}'_{r,s}\boldsymbol{T}_{s,r}\,.$$

Hence

$$Q=\sum_{r=1}^{p}\lambda_r^2(T_r'T_r)+\sum_{r\neq s}\lambda_r\lambda_s(T_{r,s}'T_{s,r}) . \tag{3.20}$$

Next, we consider the equality:

$$Q=\sum_{r=1}^{p}\lambda_r^2(t_r't_r)+\sum_{r=1}^{p-1}\sum_{k=1}^{p-r}\Sigma^*\Phi_{r,s_1,\cdots,s_k} , \tag{3.21}$$

where $*$ denotes summation over all positive integral values of $s_1, \cdots, s_k$ such that $r<s_1<\cdots<s_k\leqq p$, and where

$$\begin{aligned}\Phi_{r,s_1,\cdots,s_k}=&\lambda_r^2[t_{r,s_1,s_2,\cdots,s_k}'t_{r,s_1,\cdots,s_k}]\\&+2\lambda_r[t_{r,s_1,\cdots,s_k}'t_{r,s_1,\cdots,s_k}^*]+[t_{r,s_1,\cdots,s_k}^{*\prime}t_{r,s_1,\cdots,s_k}^*] ,\end{aligned} \tag{3.22}$$

$$t_{r,s_1,\cdots,s_k}^*=\lambda_{s_1}t_{s_1,r,s_2,\cdots,s_k}+\lambda_{s_2}t_{s_2,r,s_1,s_2,\cdots,s_k}+\cdots+\lambda_{s_k}t_{s_k,r,s_1,\cdots s_{k-1}} . \tag{3.23}$$

Before proving (3.21) we show that its truth implies the statement of the theorem. For this purpose, one can easily show (in a manner analogous to the one in theorem 2.5) that $\Phi_{r,s_1,\cdots,s_k}\geqq 0$. Since $t_r't_r\geqq 0$, (3.21) therefore implies $Q\geqq 0$.

To complete the proof of the theorem, we therefore show that the coefficient of λ_α^2 and $\lambda_\alpha\lambda_\beta$ ($\alpha\neq\beta$; $\alpha, \beta=1, \cdots, p$) respectively on the r.h.s. of (3.21) are the same as on the r.h.s. of (3.20). Consider λ_1^2. Its coefficient in (3.21) is

$$(t_1't_1)+\sum_{k=1}^{p-1}\Sigma^*[t_{1,s_1,\cdots,s_k}'t_{1,s_1,\cdots,s_k}] .$$

It is clear that this expression equals the total sum of squares of the elements from all the vectors t, which have the suffix 1 in them (with or without further suffixes). It therefore equals $(T_1'T_1)$, the coefficient of λ_1^2 in (3.20).

Next take λ_α^2, $\alpha>1$. In view of (3.23), its coefficient in (3.21) is

$$\begin{aligned}&[t_\alpha't_\alpha]+\sum_{k=1}^{p-\alpha}\Sigma^*[t_{\alpha,s_1,\cdots,s_k}'t_{\alpha,s_1,\cdots,s_k}]\\&\qquad+\sum_{s_0=1}^{\alpha-1}\sum_{k=1}^{p-s_0}\Sigma^{**}[t_{\alpha,s_0,s_1,\cdots,s_k}'t_{\alpha,s_0,s_1,\cdots,s_k}] ,\end{aligned} \tag{3.24}$$

where $**$ indicates summation over all s_j's such that $s_0<s_1<\cdots<s_k$, and all s_j's are distinct from α. A scrutiny of (3.24) shows that it involves all the vectors t which have α as one of the suffixes. The result regarding λ_α^2 is therefore proved.

Consider now the coefficient of $(2\lambda_\alpha\lambda_\beta)$ ($\alpha<\beta$, say) in (3.21), which equals

$$(3.25)\qquad \sum_{k=1}^{p-\alpha} \Sigma^*[t'_{\alpha,s_1,\cdots,s_k} t_{\beta,\alpha,s_1,*,s_k}] + \sum_{s_0=1}^{\alpha-1}\sum_{k=1}^{p-s_0} \Sigma^{**}[t'_{\beta,s_0,*,s_k} t_{\alpha,s_0,**,s_k}] ,$$

where (i) in the first term $\alpha < s_1 < \cdots < s_k$, and the set $(s_1, \cdots, s_k)$ includes the integer β, (ii) $*$ in the suffix of t indicates the set of s_j's *excluding* β, (iii) $**$ means the set $(s_0, \cdots, s_k)$ excluding α, and (iv) Σ^{**} indicates summation over $1 \leqq s_0 < s_1 < \cdots < s_k \leqq p$, where the set $(s_0, \cdots, s_k)$ includes both the integers α and β.

Consider the term inside the first bracket in (3.25). This is a scalar product of two vectors t, whose suffixes (which include *both* α and β) are the same except for a permutation, and which are such that α occurs as the first suffix in one of them, and β as the first suffix in the other. The same situation prevails with the term in the second square bracket. Also, an examination of the ranges of the summation sign shows that all such products are included, and therefore (3.25) equals $T'_{\alpha,\beta} T_{\beta,\alpha}$. This completes the proof.

The above theorem implies

THEOREM 3.2. *A n. & s.c. for the existence of a (unique) BLU estimate for θ is "condition C^*," that $g_s^{(r)}$ be orthogonal to $\theta_{rj}^{(s)}$ for all permissible j, r and s.*

Further reduction of the condition C^* in terms of the matrices A_i and the coefficient vectors c_j will be considered later.

4. Multiple-design multivariate model

A development analogous to the earlier models exists here. Define θ, l and l_0 as under the GIM model, by considering one response at a time, and adding. Then $\mathrm{Var}(l) - \mathrm{Var}(l_0)$ equals $J_{1p}(\Sigma^*\Omega)J_{p1}$, where as before $\omega_{rs} = (d'_r d_s) - (g'_r g_s)$, for $r, s = 1, \cdots, p$. Let W_r be the column space of A_r (occurring in (2.5)), $\overline{W}_r$ the space orthogonal to W_r, and θ_{rj} ($j = 1, \cdots, \mu_r$; $\mu_r = \dim(\overline{W}_r)$) an orthogonal basis for $\overline{W}_r$. As before, we then have $d_r - g_r = T_r$, where

$$(4.1)\qquad T_r = \sum_{j=1}^{\mu_r} h_{rj}\theta_{rj} , \qquad r = 1, \cdots, p .$$

Thus

$$(4.2)\qquad \omega_{rr} = T'_r T_r , \qquad \omega_{rs} = g'_r T_s + g'_s T_r + T'_r T_s .$$

We then have

THEOREM 4.1. *A n. & s.c. that a (unique) BLUE for θ exists under* *the MDM model is condition C^{**}: $\boldsymbol{g}_r$ be orthogonal to $\boldsymbol{\theta}_{sj}$, for all permissible r, s and j.*

PROOF. Necessity follows on the lines of theorem 2.5. To prove sufficiency, we observe that under the condition C^{**}, we get $\omega_{rs}=\boldsymbol{T}_r'\boldsymbol{T}_s$ $(r, s=1, 2, \cdots, p)$. Thus, $\Omega=T'T$ where $T=[\boldsymbol{T}_1, \boldsymbol{T}_2, \cdots, \boldsymbol{T}_p]$, and hence $\text{Rank}(\Omega)\geqq 1$ and Ω is at least p.s.d. Hence, as in the proof of lemma 2.3, it follows that $(\Sigma^*\Omega)$ is p.d. and hence $J_{1p}(\Sigma^*\Omega)J_{p1}>0$. This completes the proof.

Note that condition C^{**} is equivalent to "$\boldsymbol{g}_r \in W_s$, for all permissible r and s".

5. Further discussion of existence conditions

(a) MDM model. We discuss this model first because of its simplicity. The BLU estimate (if C^{**} holds) of $\theta=\Sigma\boldsymbol{c}_s'\boldsymbol{\xi}_s$ is $l_0=\Sigma\boldsymbol{g}_s'\boldsymbol{y}_s$, where $\boldsymbol{g}_s\in W_s$. Under C^{**}, we require $\boldsymbol{g}_s\in W_r$ for all r, s. This can happen only if the intersection $W=W_1\cap W_2\cap\cdots\cap W_p$ is non-empty, and $\boldsymbol{g}_s\in W$ for all s. Thus for the existence of the BLUE for a given θ, it is not necessary that the W_s be all identical.

Consider any fixed multivariate linear model, and let θ be called "piece-wise estimable" under that model if each component $\boldsymbol{c}_s'\boldsymbol{\xi}_s$ of θ has a BLUE $\boldsymbol{g}_s'\boldsymbol{y}_s$ under the univariate model obtained by considering the response V_s alone, and ignoring the rest. Let Θ be the set of all piece-wise estimable functions θ. Then a given design is called an orthogonal multiresponse design (OMD), if it is such that there exists a BLUE of θ, for all $\theta\in\Theta$. We prove:

THEOREM 5.1. *Under the MDM model, a n. & s.c. that a design is OMD is that $A_1,\cdots, A_p$ be such that their column spaces $W_1,\cdots, W_p$ are identical.*

PROOF. Suppose $W_1\neq W$. Then there exists a vector $\boldsymbol{g}_1^*\in W_1$, such that $\boldsymbol{g}_1^*\notin W_r$, for some r. Let $\mathrm{E}(\boldsymbol{g}_1^{*\prime}\boldsymbol{y}_1)=\boldsymbol{c}_1^{*\prime}\boldsymbol{\xi}_1$. Consider the estimation of $\theta=\boldsymbol{c}_1^{*\prime}\boldsymbol{\xi}_1+\boldsymbol{c}_2'\boldsymbol{\xi}_2+\cdots+\boldsymbol{c}_p'\boldsymbol{\xi}_p$, where $\boldsymbol{c}_j$ $(j>1)$ are chosen so that the BLUE for $\boldsymbol{c}_j'\boldsymbol{\xi}_j$ exists. Then clearly $\theta\in\Theta$. But since $\boldsymbol{g}_1^*\notin W_r$, condition C^{**} is not satisfied, and a BLUE for θ does not exist. This completes the proof of 'necessity'. Sufficiency (proved earlier in [14]) follows in an obvious manner.

The above points out the important fact that even if we are interested in all $\theta\in\Theta$, the need for being identical is for the column spaces of A_s, not A_s themselves. It seems to the author that by choosing the design carefully, one could achieve this condition in a majority of cases.

Indeed, the condition is satisfied in the example of MDM model given in [14].

(b) GIM and HM models. For clarity's sake, we first consider the case $p=2$. For full generality, we then take $u=3$. Let V_1 be measured on S_1, V_2 on S_2, and both V_1 and V_2 on S_3. We have the GIM model if each on n_1, n_2 and n_3 is non-zero, the HM if either n_1 or n_2 is zero, but $n_3 \neq 0$, and the trivial case when $n_3=0$. Also then,

$$B_1'=[A_1' \vdots A_3'] , \qquad B_2'=[A_2' \vdots A_3'] \tag{5.1}$$

$$m_1=n_1+n_3 , \qquad m_2=n_2+n_3 \tag{5.2}$$

$$\mu_1=m_1-\rho_1 , \qquad \mu_2=m_2-\rho_2 . \tag{5.3}$$

Define the set Θ as in subsection (a). We prove:

LEMMA 5.1. *A n. & s.c. that a design be OMD under the given GIM model is that there exist matrices* $K_{11}(n_1 \times \mu_1)$, $K_{12}(n_3 \times \mu_1)$, $K_{21}(n_2 \times \mu_2)$, $K_{22}(n_3 \times \mu_2)$ *such that*

(5.4) (a) $A_1'K_{11}$, $A_3'K_{12}$, $A_2'K_{21}$ *and* $A_3'K_{22}$ *are zero matrices,*

(5.5) (b) $K_1'K_1=I_{\mu_1}$, $K_2'K_2=I_{\mu_2}$,

where

$$K_1'=[K_{11}' \mid K_{12}'] , \qquad K_2'=[K_{21}' \mid K_{22}'] . \tag{5.6}$$

PROOF. (i) *Sufficiency.* Using (5.2), (5.4) and (5.5), it follows that the columns of K_1 form an orthogonal basis of $\overline{W}_1$, and therefore may be taken to be the vectors $\boldsymbol{\theta}_{11}, \cdots, \boldsymbol{\theta}_{1\mu_1}$. Now $\theta=\boldsymbol{c}_1'\boldsymbol{\xi}_1+\boldsymbol{c}_2'\boldsymbol{\xi}_2$. Since $\theta \in \Theta$, l_0 exists, and $l_0=\boldsymbol{g}_1'\boldsymbol{y}_1+\boldsymbol{g}_2'\boldsymbol{y}_2$. Recall that $\boldsymbol{g}_2$ belongs to the column space of B_2. Hence $\boldsymbol{g}_2^{(1)}$ belongs to the column space of (A_3). Also the vectors $\boldsymbol{\theta}_{11}^{(2)}, \cdots, \boldsymbol{\theta}_{1\mu_1}^{(2)}$ are (from (5.6)) the columns of K_{12}. Since $A_3'K_{12}$ is zero, the condition C^* (of theorem 3.1) is satisfied for $\boldsymbol{g}_2^{(1)}$. Similarly for $\boldsymbol{g}_1^{(2)}$, and the proof is completed.

(ii) *Necessity.* If C^* holds for all $\boldsymbol{g}_1 \in W_1$ and $\boldsymbol{g}_2 \in W_2$ (i.e., for all $\theta \in \Theta$), then clearly

$$K_1=[\boldsymbol{\theta}_{11}, \cdots, \boldsymbol{\theta}_{1\mu_1}] , \qquad K_2=[\boldsymbol{\theta}_{21}, \cdots, \boldsymbol{\theta}_{2\mu_2}] ,$$

will satisfy (5.4)–(5.6). Hence the lemma.

The next natural question is: 'When do K_1 and K_2 exist?'. This is answered in

THEOREM 5.2. *A n. & s.c. that (for $p=2$) a design under GIM model be orthogonal is*

$$R(B_1)=R(A_1)+R(A_3)\ , \tag{5.7}$$

$$R(B_2)=R(A_2)+R(A_3) \tag{5.8}$$

where R denotes rank.

PROOF. (i) *Necessity.* Let

$$\gamma_1=R(A_1)\ , \qquad \gamma_2=R(A_2)\ , \qquad \gamma_3=R(A_3)\ . \tag{5.9}$$

Let C_j $(j=1, 2, 3)$ be a matrix of size $n_j\times(n_j-\gamma_j)$ such that $R(C_j)=n_j-\gamma_j$, $C_j'C_j=I_{n_j-\gamma_j}$, and $A_j'C_j=O_{m,n_j-\gamma_j}$. Clearly, such matrices do exist. Now consider the matrices K_{ij} of the last lemma. If they exist and satisfy (5.4), then there exist H_j $(j=1, 3)$ of size $(n_j-\gamma_j)\times\mu_1$ such that $K_{11}=C_1H_1$, and $K_{12}=C_3H_3$. But $R(C_1H_1)\leqq n_1-\gamma_1$, $R(C_3H_3)\leqq n_3-\gamma_3$. Hence

$$R(K_1)\leqq R(K_{11})+R(K_{12})\leqq(n_1-\gamma_1)+(n_3-\gamma_3)\ . \tag{5.10}$$

But in (5.5) we want $R(K_1)=\mu_1$. Hence $\mu_1\leqq(n_1-\gamma_1)+(n_3-\gamma_3)$. On the other hand, $\mu_1=(n_1+n_3)-\rho_1$. Hence $\rho_1\geqq\gamma_1+\gamma_3$. But

$$\rho_1=R(B_1)\leqq R(A_1)+R(A_3)=\gamma_1+\gamma_3\ .$$

Hence $\rho_1=\gamma_1+\gamma_3$. Similarly $\rho_2=\gamma_2+\gamma_3$, then the result is proved.
(ii) *Sufficiency.* To show this simply take

$$K_{11}=[C_1\,|\,O_{n_1,n_3-\gamma_3}]\ , \qquad K_{12}=[O_{n_3,n_1-\gamma_1}\,|\,C_3]\ .$$

In view of (5.7), K_{11} and K_{12} each has μ_1 columns, and $R(K_{11})=R(C_1)=n_1-\gamma_1$, $R(K_{12})=n_3-\gamma_3$. Also

$$A_1'K_{11}=[A_1'C_1\,|\,O_{m,n_3-\gamma_3}]=O_{m,\mu_1}\ .$$

Similarly $A_3'K_{12}=O_{m,\mu_1}$. Also, clearly

$$R(K_1)=R\left[\begin{array}{c|c} C_1 & O \\ \hline O & C_3 \end{array}\right]=R(C_1)+R(C_3)=\mu_1\ ,$$

so that K_1 is full rank. Thus the properties required for K_1 in lemma 5.1 are satisfied. Similarly for K_2. This completes the proof.

It may appear at first sight that it is necessary for (5.7) to hold that B_1 is of the form

$$B_1=\left[\begin{array}{c|c} A_1^* & O \\ \hline O & A_3^* \end{array}\right]\ .$$

However this is not true; for example, consider the case where

$$B_1=\left[\begin{array}{c} A_1 \\ \hline A_3 \end{array}\right]=\left[\begin{array}{c|c|c|c|c|c|c} \boldsymbol{c}_1 & \boldsymbol{c}_1 & \boldsymbol{c}_2 & \boldsymbol{c}_2 & \boldsymbol{c}_1+\boldsymbol{c}_2 & 2\boldsymbol{c}_1 & \boldsymbol{c}_2-\boldsymbol{c}_1 \\ \hline \boldsymbol{d}_1 & \boldsymbol{d}_2 & \boldsymbol{d}_1 & \boldsymbol{d}_3 & \boldsymbol{d}_1 & \boldsymbol{d}_1+\boldsymbol{d}_2 & \boldsymbol{d}_3-\boldsymbol{d}_2 \end{array}\right],$$

where $\boldsymbol{c}_i$'s and $\boldsymbol{d}_j$'s are vectors such that

$$R[\boldsymbol{c}_1, \boldsymbol{c}_2]=2 , \qquad R[\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3]=3 .$$

Here it may be checked that $R(A_1)=2$, $R(A_3)=3$, and $R(B_1)=5$ so that (5.7) is satisfied.

We next treat the case of general p in

THEOREM 5.3. *Under GIM model, with p responses, a n. & s.c. for the orthogonality of the design is:*

$$R(B_r)=R(B_{rs})+R(B_{rs}^*) , \qquad r\neq s ; \quad r, s=1, \cdots, p , \tag{5.10}$$

where B_{rs} is obtained by taking those rows of B_r which correspond to sets in which both variables V_r and V_s are measured, and B_{rs}^* by putting together the remaining rows of B_r. (Thus, for example, when $p=2$, $r=1$ and $r=2$, we have $B_{rs}=A_1$ and $B_{rs}^*=A_3$.)

PROOF. (i) *Necessity.* This follows from the case $p=2$ by considering the estimation problem for the set Θ_{rs} (say) of all piece-wise estimable functions θ which involve only two fixed responses, say V_r and V_s.

(ii) *Sufficiency.* Recall the condition C^* in theorem 3.1. We notice that it relates to a statement concerning a pair of variables (V_r and V_s say) at a time. Now, for any fixed (r, s), C^* is satisfied as in the case $p=2$. Hence C^* is satisfied as a whole. This completes the proof.

Under the HM model, the conditions in the above theorem simplify to a single one, as indicated in

THEOREM 5.4. *The condition for a design to be OMD under the HM model is*

$$R(A_1' \,|\, A_2' \,|\, \cdots \,|\, A_p')=R(A_1)+ \cdots +R(A_p) . \tag{5.11}$$

The proof is left to the reader.

6. Multivariate block-treatment designs

It can be shown that many situations which apparently look more general can be brought under the scope of the above models by suitable devices. For example, let there be two responses V_1 and V_2, and three sets S_1, S_2, S_3, and let the expectation equations be

$$(6.1)\qquad E(\boldsymbol{y}_1)=A_1^*\left[\frac{\boldsymbol{\beta}_1}{\boldsymbol{\tau}_1}\right], \qquad E(\boldsymbol{y}_2)=A_2^*\left[\frac{\boldsymbol{\beta}_2}{\boldsymbol{\tau}_2}\right],$$
$$E(\boldsymbol{y}_{31}\mid\boldsymbol{y}_{32})=A_3^*\left[\begin{array}{cc}\boldsymbol{\beta}_3 & \boldsymbol{\beta}_4\\ \hline \boldsymbol{\tau}_1 & \boldsymbol{\tau}_2\end{array}\right].$$

This will arise, for example, when there is a block design over the units in each set, such that (i) A_i is the design matrix for S_i, $i=1, 2, 3$, (ii) the vectors $\boldsymbol{\beta}_i$ are parameter vectors corresponding to blocks, and therefore correspond to nuisance parameters, (iii) $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ are respectively the vectors of treatment effects for responses V_1 and V_2. Let the dimensions of the different vectors be as shown below:

$$(6.2)\qquad \boldsymbol{\tau}_1(v\times 1), \quad \boldsymbol{\tau}_2(v\times 1),$$
$$\boldsymbol{\beta}_1(b_1\times 1), \quad \boldsymbol{\beta}_2(b_2\times 1), \quad \boldsymbol{\beta}_3(b_3\times 1), \quad \boldsymbol{\beta}_4(b_4\times 1),$$

where we suppose b_i's are not all equal. Therefore, as it stands now, the model (6.1) is not a special case of GIM model. However, let

$$(6.3)\qquad \boldsymbol{\xi}_1'=[\boldsymbol{\eta}_1'\mid\boldsymbol{\beta}_1'\mid\boldsymbol{\beta}_3'\mid\boldsymbol{\tau}_1'], \qquad \boldsymbol{\xi}_2'=[\boldsymbol{\eta}_2'\mid\boldsymbol{\beta}_2'\mid\boldsymbol{\beta}_4'\mid\boldsymbol{\tau}_2'],$$

$$(6.4)\qquad A_1^*=[A_{11}\mid A_{12}], \qquad A_2^*=[A_{21}\mid A_{22}], \qquad A_3^*=[A_{31}\mid A_{32}],$$

$$(6.5)\qquad A_1=[O\mid A_{11}\mid O\mid A_{12}], \qquad A_2=[O\mid A_{21}\mid O\mid A_{22}],$$
$$A_3=[O\mid O\mid A_{31}\mid A_{32}],$$

where (i) $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ are new nuisance parameters of sizes such that the sizes of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are equal, (ii) the partitioning of A_i^* in (6.4) is induced by the vectors $\boldsymbol{\beta}_i$ and $\boldsymbol{\tau}_j$ in (6.1), and (iii) O in (6.5) denotes zero matrices of appropriate sizes such that the products in the following equations are conformable:

$$(6.6)\qquad E(\boldsymbol{y}_1)=A_1\boldsymbol{\xi}_1, \qquad E(\boldsymbol{y}_2)=A_2\boldsymbol{\xi}_2, \qquad E(\boldsymbol{y}_{31}\mid\boldsymbol{y}_{32})=A_3(\boldsymbol{\xi}_1\mid\boldsymbol{\xi}_2).$$

Thus the device of introducing the nuisance parameters $\boldsymbol{\eta}_i$ brings it under GIM model.

Interest may lie here in the estimation of functions θ of the form $\boldsymbol{c}_1'\boldsymbol{\tau}_1+\boldsymbol{c}_2'\boldsymbol{\tau}_2$. Conditions for this can be obtained from theorem 4.1. However, for lack of space here, the theory of multivariate designs satisfying condition C^* is deferred to later communications, where the applications of the theory to areas like multivariate response surfaces, time series, etc., will also be considered. Included in particular will be the theory of multivariate block designs, where the concepts of relationship algebras of designs (see, for example [4] and [7]) are gainfully employed.

7. Acknowledgement

I am thankful to the referees for pointing out some misprints or omissions and an error in the statement of lemma 2.3, in the original version of the paper.

UNIVERSITY OF NEBRASKA*

REFERENCES

[1] T. W. Anderson, *Introduction to Multivariate Analysis*, John Wiley, New York, 1956.

[2] R. C. Bose, *Mimeographed Notes on Analysis of Variance*, Institute of Statistics, University of North Carolina, 1958.

[3] D. R. Cox, *Planning of Experiments*, John Wiley, New York, 1960.

[4] A. T. James, "The relationship algebra of a PBIB design," *Ann. Math. Statist.*, 28 (1957), 349-361.

[5] O. Kempthorne, *The design and Analysis of Experiments*, John Wiley, New York, 1952.

[6] F. M. Lord, "Estimation of parameters from incomplete data," *J. Amer. Statist. Ass.*, 50 (1955), 870-876.

[7] J. Ogawa and G. Ishii, "The relationship algebra and the analysis of variance of a PBIB design," *Ann. Math. Statist.*, 36 (1965), 1815-1828.

[8] C. R. Rao, "Analysis of dispersion with incomplete observations on one of the characters," *J. Roy. Statist. Soc.*, Ser. B, 18 (1956), 259.

[9] S. N. Roy, *Some Aspects of Multivariate Analysis*, John Wiley, New York, 1957.

[10] S. N. Roy and J. N. Srivastava, *Hierarchical and p-block Multiresponse Designs and Their Analysis*, Mahalanobis dedicatory volume, Indian Statistical Institute, Calcutta, 1962.

[11] H. Scheffe, *Analysis of Variance*, John Wiley, New York, 1959.

[12] H. Smith, R. Gnanadesikan and J. B. Hughes, "Multivariate analysis of variance," *Biometrics*, 18 (1962), 22-41.

[13] J. N. Srivastava, "A multivariate extension of the Gauss-Markov theorem," *Ann. Inst. Statist. Math.*, 17 (1965), 63-66.

[14] J. N. Srivastava, "Some generalizations of multivariate analysis of variance," *Multivariate Analysis*, (edited by P. R. Krishnaiah), Academic Press, New York, (1966), 129-148.

[15] J. N. Srivastava, "On a general class of designs for multiresponse experiments," *N. C. Inst. of Statist.*, Mimeo Series No. 402, 1964.

[16] J. N. Srivastava, "On the extension of Gauss-Markov theorem to complex multivariate linear models," (abstract), *Ann. Math. Statist.*, 37 (1966), 1860.

[17] J. N. Srivastava, "Incomplete multiresponse designs," *Sankhyā* A, 28 (1966), 377-388.

[18] I. M. Trawinski, *Incomplete-Variable Designs*, (unpublished thesis), V. P. I., Blacksburg, Virginia, 1961.

[19] I. M. Trawinski and R. E. Bargmann, "Maximum likelihood estimation with incomplete multivariate data," *Ann. Math. Statist.*, 35 (1964), 647-657.

---

* Now at Colorado State University