This research was supported by the National Institutes of Health, Institute of Medical Sciences, Grant No. GM-12868-04

A GENERAL METHOD FOR OBTAINING TEST CRITERIA FOR MULTIVARIATE LINEAR MODELS WITH MORE THAN ONE DESIGN MATRIX AND/OR INCOMPLETE IN RESPONSE VARIATES

by

David G. Kleinbaum

University of North Carolina
Institute of Statistics Mimeo Series No. 614

ABSTRACT

KLEINBAUM, DAVID GEORGE. A General Method for Obtaining Test Criteria for Multivariate Linear Models with More Than One Design Matrix and/or Incomplete in Response Variates.

In this paper a test of a general linear hypothesis for the More General Linear Multivariate Model (MGLMM) is obtained under normality assumptions on the original data by using a test criterion given in a general form by Wald [12]. The asymptotic distribution of the test statistic under the null hypothesis is a central chi-square variable. As a special case of the above test, we obtain a test of the standard MANOVA linear model which is Hotelling's trace criterion. Other special cases of the general model include the HM, GIM and MDM models given by Srivastava [4], [5], [8], [9].

The test statistic uses estimators of the parameters which need only be asymptotically equivalent in probability to the maximum likelihood estimators.

A GENERAL METHOD FOR OBTAINING TEST CRITERIA FOR MULTIVARIATE LINEAR MODELS WITH MORE THAN ONE DESIGN MATRIX AND/OR INCOMPLETE IN RESPONSE VARIATES

David G. Kleinbaum

1. INTRODUCTION AND SUMMARY

The standard multivariate linear model (SM) given by:

$$E(Y) = A\,\xi$$

$$\mathrm{Var}(Y) = I_n \otimes \Sigma$$

where $Y(n \times p)$, $A(n \times m)$, $\xi(m \times p)$, $\Sigma(p \times p)$

contains two assumptions inherent to the experimental situation which are not always met in practice. These are

(i) each response variate is measured on each experimental unit (i.e., on each experimental unit is observed a p-variate vector).

(ii) the design matrix, A, is the same for each response (e.g., the same blocking system is applicable to each variate).

Situation (i) will, in general, not hold whenever it is physically impossible, uneconomical or inadvisable to observe each response variate on each experimental unit. Situation (ii) will not hold when different blocking systems are applicable to different response variates and when some of the response variates are insensitive to certain treatments.

This paper is concerned with testing linear hypotheses under normality assumptions on the original data for multivariate linear models in which assumptions (i) and/or (ii) are relaxed. In this connection, we define a model, called by the author the More General Linear Multivariate

Model (MGLMM)[1]. Special cases of the MGLMM are the HM, GIM and MDM models of Srivastava ([4], [5], [7], [8] and [9]) and the incomplete variable designs of Monahan [2]. A test of a general linear hypothesis for the MGLMM is proposed, which uses a test criterion given in a general form by Wald [12]. The Wald test statistic uses estimators of the unknown parameters which need only be asymptotically equivalent in probability to maximum likelihood estimators. The asymptotic distribution of the test statistic under the null hypothesis is a central chi-square variable. When restricted to the SM model, we obtain Hotelling's trace criterion as a special case.

2. THE MORE GENERAL LINEAR MULTIVARIATE MODEL

Assume there are n experimental units and p-response variates $V_1, \ldots, V_p$ in total. The n experimental units are divided into u disjoint sets of experimental units $S_1, S_2, \ldots, S_u$ with n_j units in S_j. On each unit in the set S_j, we measure q_j ($\leq p$) responses $V_{\ell_{j1}}, V_{\ell_{j2}}, \ldots, V_{\ell_{jq_j}}$. (The remaining $p - q_j$ response variates are not measured in S_j).

---

[1]There are meaningful multivariate linear models which are not special cases of the MGLMM (e.g., growth curve models). Hence we have used "More" instead of "Most" in the model name. Nevertheless, the technique used in this paper for obtaining test statistics can be used for linear models which are not special cases of MGLMM.

Then the MGLMM is given by

$$E(Y_j) = [A_{j\ell_{j1}} \underline{\xi}_{\ell_{j1}} \mid A_{j\ell_{j2}} \underline{\xi}_{\ell_{j2}} \mid \cdots\cdots \mid A_{j\ell_{jq_j}} \underline{\xi}_{\ell_{jq_j}}] \quad n_j \times q_j$$

$$Var(Y_j) = I_{n_j} \otimes B_j' \Sigma B_j$$

$$j = 1,\ldots, u; \; 1 \le q_j \le p, \; \ell_{j1} \le \ell_{j2} \le \cdots \le \ell_{jq_j}$$

and for each i, i = 1,...,p there exists at least one pair (j,k) such that $\underline{\xi}_{\ell_{jk}} = \underline{\xi}_i$, i.e., $\ell_{jk} = i$ (so that the total set of unknown parameters is $(\underline{\xi}_1, \ldots, \underline{\xi}_p, \Sigma)$).

$Y_j(n_j \times q_j)$ is the matrix of observations for the jth set S_j.

$A_{j\ell_{jk}}(n_j \times m_{\ell_{jk}})$ is the design matrix for response $V_{\ell_{jk}}$ in the set S_j, and rank $A_{j\ell_{jk}} = m_{\ell_{jk}}$

$\underline{\xi}_{\ell_{jk}}(m_{\ell_{jk}} \times 1)$ is the parameter vector for response $V_{\ell_{jk}}$ (and $m_{\ell_{jk}} = m_{\ell_{j'k'}}$ whenever $\ell_{jk} = \ell_{j'k'}$)

$B_j(p \times q_j)$ is the incidence matrix for response variates consisting of 0's and 1's and rank $B_j = q_j$.

If the variates V_r, V_s are not measured together in any set S_k, we set $\sigma_{rs} \equiv 0$ and $\sigma_{sr} \equiv 0$ in $\Sigma = ((\sigma_{ij}))$.

For testing the hypothesis $H_o: \bigcap_{j=1}^{p} C_j \underline{\xi}_j = \underline{0}$ where $C_j \; (c_j \times m_j)$ is of full rank = c_j we will assume that

(1) Y_j and Y_j' are independent if $j \neq j'$

(2) the rows of Y_j are also independent and distributed as a q_j-variate multinormal vector with variance covariance matrix $B_j' \Sigma B_j$.

From (1) and (2) we can write the log of the likelihood function for MGLMM as

$$\log \phi = - \sum_{j=1}^{u} \frac{n_j q_j}{2} \log 2\pi - \sum_{j=1}^{u} \frac{n_j}{2} \log |B_j' \Sigma B_j|$$

$$- \frac{1}{2} \sum_{j=1}^{u} \mathrm{tr}[B_j' \Sigma B_j]^{-1} [Y_j - E[Y_j]]' [Y_j - E[Y_j]]$$

We must use log ϕ to obtain our test statistic. The technique used is described in section 4.

2.1. Special Cases of MGLMM

2.1.1. Multiple Design Multivariate Model - MDM (Srivastava [4], [8], [9]) $u = 1$; $n_1 = n$; $q_1 = p$; $Y_1 \equiv Y$; $B_1 \equiv I_p$. Thus we get

$$E(Y) = [A_1 \underline{\xi}_1 \quad A_2 \underline{\xi}_2 \quad \cdots \quad A_p \underline{\xi}_p]_{(n \times p)}$$

$$\mathrm{Var}(Y) = I_n \otimes \Sigma$$

2.1.2. Hierarchical Model - HM (Srivastava [4], [8], [9])

$u = p$; $q_j = j$, $j=1,\ldots,p$; $\ell_{jk} = k$, $k=1,\ldots,j$;

$$B_j = \begin{bmatrix} I_j \\ 0_{p-j,j} \end{bmatrix}$$

Thus we get

$$E[Y_j] = [A_{j1} \underline{\xi}_1 \quad A_{j2} \underline{\xi}_2 \quad \cdots \quad A_{jj} \underline{\xi}_j]_{(n_j \times j)}$$

$$\mathrm{Var}(Y_j) = I_{n_j} \otimes B_j' \Sigma B_j$$

In [8] and [9], Srivastava assumes $A_{j1} = A_{j2} = \ldots = A_{jj} \equiv A_j$.

Note that if we let

U_r = the set of all experimental units on which the response variate V_r is measured

then for the HM model we have

$$U_1 \supset U_2 \supset \ldots \supset U_p$$

The HM model thus deals with a situation in which the response variates can be graded in descending order of importance, and, further, if V_r is more important than V_s, then V_r is observed on each experimental unit on which V_s is observed.

2.1.3. <u>General Incomplete Multivariate Model - GIM</u> (Srivastava [5], [7], [8]; Monahan [2])

$A_{j\ell_{j1}} = A_{j\ell_{j2}} = \ldots = A_{j\ell_{jq_j}} \equiv A_j$, j = 1 ,..., u

$\underline{\xi}_{\ell_{jk}}$ is $(m \times 1)$, where m is independent of j and ℓ_{jk}.

Thus

$$E(Y_j) = A_j[\underline{\xi}_{\ell_{j1}} \quad \underline{\xi}_{\ell_{j2}} \quad \cdots \quad \underline{\xi}_{\ell_{jq_j}}]_{(n_j \times q_j)} (= A_j \xi B_j)$$

$$\mathrm{Var}(Y_j) = I_{n_j} \otimes B_j' \Sigma B_j$$

Monahan ([2]) considers the special case $q_j \equiv q$; $n_j \equiv n/u$; $A_j \equiv A$, $j = 1,\ldots,u$.

3. REVIEW OF THE LITERATURE

A number of different approaches have been used to obtain tests of hypotheses for the MDM, HM and GIM models. The tests obtained are generally impractical with regard to computation and require more assumptions with regard to the hypotheses and design matrices than the usual estimability requirement. For example, the necessary and sufficient conditions given by Srivastava in [9] for reparameterizing the MDM model to the SM model (for which there are several good test criteria) is that the vector spaces spanned by the columns of the different design matrices are all the same. In [5], Srivastava restricts the GIM model (actually a slightly more general form of the GIM model in which $\underline{\xi}_j$ is $(m_j \times 1)$ and A_j is $(n_j \times m_j)$) to be what he calls a "strongly regular" design. In [4], he proves to be not strongly regular a particular example of the kind of design likely to be considered in practice. In [7], however, he works through a GIM model which is strongly regular and specifies the general computational method for reparameterization. Nevertheless, the class of "strongly regular designs" needs to be better described in terms of GIM models likely to be used in practice. This problem seems formidable since the computations needed for validating the "strongly regular" property are quite complex. Furthermore, the reparameterized form of the model has a Σ matrix which is restricted by $|\rho_{ij}| \leq \lambda_{ij} \leq 1$ where λ_{ij} may be < 1 for some $i \neq j$ and ρ_{ij} is the correlation between variates V_i and V_j. The problem of testing a hypothesis in a SM model with the above type of restriction on Σ has not yet been solved and is likely to require a laborious computer procedure.

In [11] Trawinski (i.e., Monahan) and Bargmann summarize the main results of Monahan's thesis [2] in which is obtained a likelihood ratio test for the very special case of the GIM model in which $u = p$, $n_j \equiv n$, $q_j \equiv q$ and $A_j \equiv A$, $j=1,\ldots, p$. The maximum likelihood equations require for solution an iterative technique based on the Newton-Raphson method. In demonstration of the method for a particular set of data, the initial estimate Σ_o of Σ used in the iteration turned out to be very comparable to the final estimate obtained after several iterations. The Σ_o estimate (or a more general form of Σ_o) is a likely candidate to be studied for use in the Wald statistic proposed by this author. It is a pooled estimate, in which the $(r - s)$ element $\hat{\sigma}^o_{rs}$ is obtained by averaging estimates of σ_{rs} from all the sets S_j in which both V_r and V_s are measured.

In [10], Srivastava has suggested several different union-intersection type tests for the HM model, one of which was obtained in his earlier paper [4] and which involves a generalization of J. Roy's step-down procedure [3]. The model that he uses in [4] includes the MDM model as a special case. The testing procedure involves making a sequence of independent F tests and rejecting the hypotheses if any of the F tests are rejected. As is the case in general with union-intersection type tests, little can be said about the properties, asymptotic or otherwise, of the step-down method.

The Wald statistic proposed in this paper was used by Allen [1] in connection with nonlinear multivariate models, a special case of which is the SM model. Allen showed that the likelihood ratio criterion and the Wald criterion for his nonlinear model are asymptotically equivalent. In small samples, however, Wald's criterion proved not to be a good approximation of a chi-square variable. Nevertheless, this may have been due to the use of a linear approximation of the test criterion (which is not necessary for the linear MGLMM).

In [8], Srivastava considers distribution-free estimation of linear functions of the design parameters for the HM, GIM and MDM models. Srivastava gives necessary and sufficient conditions for $\theta = \sum_{i=1}^{p} \underline{c}_i' \underline{\xi}_i$ to have a BLUE. These are conditions on the vector spaces spanned by the columns of the design matrices corresponding to the different response variates. The proofs have been simplified considerably by this author.

4. DERIVATION OF RESULTS

4.1 Lemma. Let $\phi(\underline{X}_1,\ldots,\underline{X}_n,\underline{\theta})$ be the joint density of independent random variables $\underline{X}_1,\ldots,\underline{X}_n$, where $\underline{X}_i \sim p_i(\underline{X},\underline{\theta})$, $i=1,\ldots,n$. The $\underline{X}_i$ are vectors which may or may not be of the same length. Assume that the total number of parameters, say u, in $\underline{\theta}$ is independent of n.

Let $\hat{\underline{\theta}}$ = M.L.E. of $\underline{\theta}$

and $$B_n^*(\underline{\theta}) = \left(\left(-1/n\, E_{\underline{\theta}}\left(\frac{\partial^2 \log\phi}{\partial\theta_i \partial\theta_j}\right)\right)\right)_{(u \times u)}$$

Then under suitable regularity conditions on the $p_i(\underline{X},\underline{\theta})$ we have that

(i) $\mathcal{L}(\sqrt{n}\, B_n^{*1/2}(\underline{\theta})\ (\hat{\underline{\theta}} - \underline{\theta})) \longrightarrow MN_u(\underline{0}, I)$

(ii) $\mathcal{L}(n(\hat{\underline{\theta}} - \underline{\theta})'\, B_n^*(\underline{\theta})\ (\hat{\underline{\theta}} - \underline{\theta})) \longrightarrow \chi_u^2$

(iii) $\mathcal{L}(n(\hat{\underline{\theta}} - \underline{\theta})'\, B_n^*(\hat{\underline{\theta}})\ (\hat{\underline{\theta}} - \underline{\theta})) \longrightarrow \chi_u^2$

provided the following conditions are satisfied:

$$(1)\quad \frac{1}{n^{1+\delta}} \sum_{i=1}^{n} E \left| \frac{\partial^2 \log P_i(\underline{X}_i, \underline{\theta})}{\partial\theta_j \, \partial\theta_{j'}} \right|^{1+\delta} \longrightarrow 0 \text{ as } n\to\infty \text{ for all } j, j=1,\ldots,u, \text{ and } 0<\delta\leq 1$$

and

$$(2)\quad \left\{ nE\left[-\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log\phi}{\partial\theta_j^2} \right] \right\}^{-(1+\frac{\delta}{2})} \sum_{i=1}^{n} E \left| \frac{\partial \log P_i(\underline{X}_i, \underline{\theta})}{\partial\theta_j} \right|^{2+\delta} \longrightarrow 0 \text{ as } n\to\infty \text{ for } \delta>0, j=1,\ldots,u$$

Comments: This is a direct generalization of a well known theorem for maximum likelihood estimators in which the $\underline{X}_i$ are i.i.d. random variables. The statistic given in (iii) was first suggested for use in hypothesis testing by Wald [12]. Conditions (1) and (2) are not severe restrictions for the models discussed in this paper; the conditions merely require that each design point be repeated often enough and each distinct variate (e.g., blood pressure) observed often enough in large samples.

4.2 Corollary. Given the same situation as in Lemma 4.1, and suppose that $\underline{\theta}' = (\underline{\xi}', \underline{\sigma}')$ so that $\hat{\underline{\xi}}$ = M.L.E. of $\underline{\xi}$ and $\hat{\underline{\sigma}}$ = M.L.E. of $\underline{\sigma}$.

Let

$$B_n^*(\underline{\xi}, \underline{\sigma}) = \left(\left(-1/n \, E_{\underline{\theta}} \left(\frac{\partial^2 \log\phi}{\partial\xi_i \partial\xi_j} \right) \right)\right) \quad (u \times u)$$

Then if (1) and (2) of Lemma 4.1 hold, and, in addition,

$$E_{\underline{\theta}} \left[\frac{\partial^2 \log\phi}{\partial\xi_i \partial\sigma_j} \right] = 0 \quad \text{for all } i,j$$

we have that

(i) $\mathcal{L}(\sqrt{n}\ B_n^*(\underline{\xi}, \underline{\sigma})(\hat{\underline{\xi}} - \underline{\xi})) \longrightarrow MN(\underline{0}, I)$

(ii) $\mathcal{L}(n(\hat{\underline{\xi}} - \underline{\xi})'\ B_n^*(\hat{\underline{\xi}}, \hat{\underline{\sigma}})(\hat{\underline{\xi}} - \underline{\xi})) \rightarrow \chi_u^2$

(iii) $\mathcal{L}(n(\underline{\xi}^* - \underline{\xi})\ B_n^*(\underline{\xi}^*, \underline{\sigma}^*)(\underline{\xi}^* - \underline{\xi})) \rightarrow \chi_u^2$ where $\underline{\xi}^*$ is asymptotically equivalent to $\hat{\underline{\xi}}$, i.e., $\sqrt{n}(\underline{\xi}^* - \hat{\underline{\xi}}) \underset{p}{\longrightarrow} 0$, and $\underline{\sigma}^* \underset{p}{\rightarrow} \underline{\sigma}$.

4.3 Theorem Given the standard MANOVA model:

$$E(Y) = A\ \xi$$

$$\mathrm{Var}(Y) = I_n \otimes \Sigma$$

where $Y(n \times p)$, $A(n \times m)$, $\xi(m \times p)$, $\Sigma(p \times p)$, rank $A = m$.

Suppose the rows of Y are distributed p-variate multinormal.

Let $H_o : C\ \xi = 0$ where $C(q \times m)$ is of rank q.

Let $\hat{\xi} = [\hat{\underline{\xi}}_1 \ldots \hat{\underline{\xi}}_p]$ and $\hat{\Sigma}$ be the usual M.L.E. of ξ and Σ respectively.

Then under H_o we have

$$\begin{bmatrix} C\hat{\underline{\xi}}_1 \\ \vdots \\ C\hat{\underline{\xi}}_p \end{bmatrix}' \left\{ \begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix} [\Sigma^{-1} \otimes A'A]^{-1} \begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix}' \right\}^{-1} \begin{bmatrix} C\hat{\underline{\xi}}_1 \\ \vdots \\ C\hat{\underline{\xi}}_p \end{bmatrix}$$

is distributed asymptotically as χ_{qp}^2.

Proof:

$$\log\phi(Y, \xi, \Sigma) = -\frac{np}{2}\log 2\pi - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\mathrm{tr}\ \Sigma^{-1}(Y - A\xi)'(Y - A\xi)$$

$$\frac{\partial \log\phi}{\partial \xi} = -\frac{1}{2}\frac{\partial \mathrm{tr}\ \Sigma^{-1}(Y-A\xi)'(Y-A\xi)}{\partial \xi} = A'(Y-A\xi)\Sigma^{-1}$$

$$\therefore \quad \frac{\partial \log\phi}{\partial \xi_{k\ell}} = (k\text{-}\ell)^{th} \text{ element of } A'(Y-A\xi)\Sigma^{-1}$$

$$= \sum_{z=1}^{p}\sum_{u=1}^{n} a_{uk}y_{uz}\,\sigma^{z\ell} - \sum_{z=1}^{p}\sum_{u=1}^{n}\sum_{w=1}^{m} a_{uk}a_{uw}\xi_{wz}\,\sigma^{z\ell}\ .$$

where $A(n \times m) = ((a_{uw}))$, $\xi(m \times p) = ((\xi_{wz}))$, $\Sigma^{-1}(p \times p) = ((\sigma^{z\ell}))$,

$Y(n \times p) = ((y_{uz}))$.

Now

$$\frac{\partial\left[\frac{\partial \log\phi}{\partial\xi}\right]}{\partial\sigma_{qr}} = A'(Y - A\xi)\ \frac{\partial\Sigma^{-1}}{\partial\sigma_{qr}}$$

so that

$$E\left[\frac{\partial\left[\frac{\partial \log\phi}{\partial\xi}\right]}{\partial\sigma_{qr}}\right] = A'E(Y - A\xi)\ \frac{\partial\Sigma^{-1}}{\partial\sigma_{qr}} = 0\ (m \times p)$$

and

$$E\left[\frac{\partial^2 \log\phi}{\partial\xi_{k\ell}\partial\sigma_{qr}}\right] = 0 \quad \forall \quad \begin{array}{l} k = 1,\ldots,m \\ q,\ r,\ \ell = 1,\ldots,p \end{array} .$$

Also

$$\frac{\partial\left[\frac{\partial \log\phi}{\partial\xi_{k\ell}}\right]}{\partial\xi_{qr}} = -\sum_{u=1}^{n} a_{uk}a_{uq}\sigma^{r\ell} = -\sigma^{r\ell}\ \{(q^{th} \text{ col. of } A)\cdot(k^{th} \text{ col. of } A)\}$$

$$= -\sigma^{r\ell}\ \{(q\text{-}k)^{th} \text{ element of } A'A\}, \text{ which is a constant independent of } Y.$$

Thus

$$B(\xi,\ \Sigma) = \left[\left[-E\left(\frac{\partial^2 \log\phi}{\partial\xi_{qr}\partial\xi_{k\ell}}\right)\right]\right] = [\Sigma^{-1} \otimes A'\ A]_{(mp \times mp)}$$

Now from Lemma 4.2(i) we have

$$\sqrt{n}\left\{\begin{bmatrix} \hat{\underline{\xi}}_1 \\ \vdots \\ \hat{\underline{\xi}}_p \end{bmatrix} - \begin{bmatrix} \underline{\xi}_1 \\ \vdots \\ \underline{\xi}_p \end{bmatrix}\right\} \text{ is asymptotically } MN_{mp}\ (\underline{0}\ , n[\Sigma^{-1} \otimes A'A]^{-1})$$

Similarly

$$\sqrt{n}\left\{\begin{bmatrix} C\,\hat{\underline{\xi}}_1 \\ \vdots \\ C\,\hat{\underline{\xi}}_p \end{bmatrix} - \begin{bmatrix} C\,\underline{\xi}_1 \\ \vdots \\ C\,\underline{\xi}_p \end{bmatrix}\right\} \text{is asymptotically } MN_{qp}\left(\underline{0}, n\begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix}[\Sigma^{-1} \otimes A'A]^{-1}\begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix}'\right)$$

Thus if H_o is true we amy use Lemma 4.2(iv) to obtain

$$W = \begin{bmatrix} C\,\hat{\underline{\xi}}_1 \\ \vdots \\ C\,\hat{\underline{\xi}}_p \end{bmatrix}'\left\{\begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix}[\Sigma^{-1} \otimes A'A]^{-1}\begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix}'\right\}^{-1}\begin{bmatrix} C\,\hat{\underline{\xi}}_1 \\ \vdots \\ C\,\hat{\underline{\xi}}_p \end{bmatrix}$$

is asymptotically distributed as χ^2_{pq} .

q.e.d.

We will show later that W is equivalent to Hotelling's trace criterion.

4.4 Main Theorem (A test of H_o : $\bigcap_{j=1}^{p} C_j\,\underline{\xi}_j = \underline{0}$ for MGLMM).

Suppose we have observation matrices $Y_j(n_j \times q_j)$, j=1,...,u from a MGLMM. Suppose also that Y_j and $Y_{j'}$ are independent if $j \neq j'$ and that for a given j, the rows of Y_j are independent and distributed q_j - variate multinormal.

Let $\hat{\Sigma}$ be an estimate of Σ such that $\hat{\Sigma} \underset{p}{\to} \Sigma$ and $\hat{\underline{\xi}}_i$, i=1,...,p be estimates of $\underline{\xi}_i$, i=1,...,p such that $\hat{\underline{\xi}}_i \underset{p}{\to} \underline{\xi}_i$, i=1,...,p, as in Corollary 4.2(iii).

Let

$$nB(\xi,\Sigma) = \left(\left(\sum_{j_{rs}} \sigma^{rs}_{j_{rs}} A'_{j_{rs}r} A_{j_{rs}s}\right)\right)_{\left(\sum_{r=1}^{p} m_r \times \sum_{r=1}^{p} m_r\right)}$$

where

1) j_{rs} runs over all integers k for which both responses V_r and V_s

are measured in S_k.

2) $\sigma_{j_{rs}}^{rs}$ is the element of $[B'_{j_{rs}} \Sigma B_{j_{rs}}]^{-1}$ which corresponds to the variate pair (V_r, V_s).

(Note (1) $\sigma_{j_{rs}}^{rs} \equiv \sigma_{j_{rs}}^{sr}$

(2) If the variates V_r and V_s are not measured together in any S_k, then we set

$$\sum_{j_{rs}} \sigma_{j_{rs}}^{rs} A'_{j_{rs}r} A_{j_{rs}s} \equiv 0_{(m_r \times m_s)} .$$

and $\sum_{j_{rs}} \sigma_{j_{rs}}^{sr} A'_{j_{rs}s} A_{j_{rs}r} \equiv 0_{(m_s \times m_r)}$.)

Then under H_o: $\bigcap_{j=1}^{p} C_j \underline{\xi}_j = \underline{0}$

where

$C_j (c_j \times m_j)$ is of full rank $(= c_j)$

we have

$$W = n \begin{bmatrix} C_1 \hat{\underline{\xi}}_1 \\ C_2 \hat{\underline{\xi}}_2 \\ \vdots \\ C_p \hat{\underline{\xi}}_p \end{bmatrix}' \left\{ \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_p \end{bmatrix} B^{-1} (\hat{\xi}, \hat{\Sigma}) \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_p \end{bmatrix}' \right\}^{-1} \begin{bmatrix} C_1 \hat{\underline{\xi}}_1 \\ C_2 \hat{\underline{\xi}}_2 \\ \vdots \\ C_p \hat{\underline{\xi}}_p \end{bmatrix}$$

is asymptotically distributed as a central chi-square variable with

$\sum_{j=1}^{p} c_j$ d.f.

Proof: This follows as a direct extension of Theorem 4.3. Detailed proofs are given for the MDM, GIM and HM special cases.

4.5 Special Cases of the Main Theorem

4.5.1 MDM:

$$W = \begin{bmatrix} C_1 \hat{\underline{\xi}}_1 \\ \vdots \\ C_p \hat{\underline{\xi}}_p \end{bmatrix}' \left\{ \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_p \end{bmatrix} \begin{bmatrix} \hat{\sigma}^{11} A_1'A_1 \cdots & \hat{\sigma}^{1p} A_1'A_p \\ \vdots & \\ \hat{\sigma}^{p1} A_p'A_1 \cdots & \hat{\sigma}^{pp} A_p'A_p \end{bmatrix}^{-1} \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_p \end{bmatrix}' \right\}^{-1} \begin{bmatrix} C_1 \hat{\underline{\xi}}_1 \\ \vdots \\ C_p \hat{\underline{\xi}}_p \end{bmatrix}$$

where $\Sigma^{-1} = ((\sigma^{rs}))$.

4.5.2 HM and GIM (assuming $A_{j\ell_{j1}} = A_{j\ell_{j2}} = \ldots = A_{j\ell_{jq_j}} = A_j$) :

$$W = \begin{bmatrix} C_1 \hat{\underline{\xi}}_1 \\ \vdots \\ C_p \hat{\underline{\xi}}_p \end{bmatrix}' \left\{ \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_p \end{bmatrix} [\sum_{j=1}^{u} \{B_j [B_j' \hat{\Sigma} B_j]^{-1} B_j'\} \otimes A_j'A_j]^{-1} \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_p \end{bmatrix}' \right\}^{-1} \begin{bmatrix} C_1 \hat{\underline{\xi}}_1 \\ \vdots \\ C_p \hat{\underline{\xi}}_p \end{bmatrix}$$

Proof of 4.5.1 :

$$\log \phi = \frac{-np}{2} \log 2\pi - \frac{n}{2} \log|\Sigma| - \frac{1}{2} \operatorname{tr} \Sigma^{-1}[Y-E(Y)]'[Y-E(Y)]$$

where $E(Y) = [A_1 \underline{\xi}_1 \vdots A_2 \underline{\xi}_2 \vdots \cdots \vdots A_p \underline{\xi}_p]$ and

$$\operatorname{tr} \Sigma^{-1}\{Y-[A_1\underline{\xi}_1 \vdots A_2\underline{\xi}_2, \vdots \cdots \vdots A_p\underline{\xi}_p]\}'\{Y-[A_1\underline{\xi}_1 \vdots A_2\underline{\xi}_2 \vdots \cdots \vdots A_p\underline{\xi}_p]\}$$

$$= \sum_{k=1}^{p} \sum_{\ell=1}^{p} \sigma^{k\ell} (\underline{y}_\ell - A_\ell\underline{\xi}_\ell)'(\underline{y}_k - A_k\underline{\xi}_k)$$

since $[Y-E(Y)]'[Y-E(Y)] = (((\underline{y}_\ell - A_\ell\underline{\xi}_\ell)'(\underline{y}_k - A_k\underline{\xi}_k)))$

Now

$$\frac{\partial \log\phi}{\partial \underline{\xi}_w} = \frac{\partial}{\partial \underline{\xi}_w} [-\frac{1}{2} \sum_{k=1}^{p} \sum_{\ell=1}^{p} \sigma^{k\ell} (\underline{y}_\ell - A_\ell\underline{\xi}_\ell)' (\underline{y}_k - A_k\underline{\xi}_k)]$$

$$= -\frac{1}{2} \frac{\partial}{\partial \underline{\xi}_w} [\sigma^{ww}(\underline{y}_w - A_w\underline{\xi}_w)'(\underline{y}_w - A_w\underline{\xi}_w) + 2 \sum_{\substack{\ell=1 \\ \ell\neq w}}^{p} \sigma^{w\ell}(\underline{y}_\ell - A_\ell\underline{\xi}_\ell)'(\underline{y}_w - A_w\underline{\xi}_w)]$$

$$= -\frac{1}{2}\left[-2\sigma^{ww} A_w' (\underline{y}_w - A_w\underline{\xi}_w) - 2 \sum_{\substack{\ell=1\\ \ell\neq w}}^{p} \sigma^{w\ell} A_w'(\underline{y}_\ell - A_\ell\underline{\xi}_\ell)\right]$$

$$= \sigma^{ww} A_w' (\underline{y}_w - A_w\underline{\xi}_w) + \sum_{\substack{\ell=1\\ \ell\neq w}}^{p} \sigma^{w\ell} A_w' (\underline{y}_\ell - A_\ell\underline{\xi}_\ell)$$

$$= \sum_{\ell=1}^{p} \sigma^{w\ell} A_w' (\underline{y}_\ell - A_\ell\underline{\xi}_\ell) .$$

$$\therefore \frac{\partial \log\phi}{\partial \underline{\xi}_{kw}} = \sum_{\ell=1}^{p} \sum_{t=1}^{n} \sigma^{w\ell} a_{(w)tk}\underline{y}_{t\ell} - \sum_{\ell=1}^{p} \sum_{t=1}^{n} \sum_{z=1}^{m_\ell} a_{(w)tk}a_{(\ell)tz}\underline{\xi}_{z\ell}\sigma^{w\ell}$$

where $A_w(n \times m_w) = ((a_{(w)tk}))$ and $\underline{\xi}_\ell(m_\ell \times 1) = \begin{bmatrix} \xi_{1\ell} \\ \vdots \\ \xi_{m_\ell \ell} \end{bmatrix}$

$$\therefore \frac{\partial^2 \log \phi}{\partial \xi_{kw} \partial \xi_{qr}} = \sum_{t=1}^{n} a_{(w)tk}a_{(r)tq}\sigma^{wr} .$$

$$= \{(k\text{-}q)^{th} \text{ element of } A_w'A_r\} \cdot \sigma^{wr} .$$

$$\therefore nB(\xi,\Sigma) = \begin{bmatrix} \sigma^{11}A_1'A_1 & \cdots\cdots & \sigma^{1p}A_1'A_p \\ \sigma^{21}A_2'A_1 & \cdots\cdots & \sigma^{2p}A_2'A_p \\ \vdots & & \vdots \\ \sigma^{p1}A_p'A_1 & \cdots\cdots & \sigma^{pp}A_p'A_p \end{bmatrix} .$$

q.e.d.

Proof of 4.5.2. :

$$\log \phi = -\frac{(\sum_{i=1}^{u} q_i n_i)\log 2\pi}{2} - \frac{1}{2}\sum_{i=1}^{u} n_i \log |B_i' \Sigma B_i|$$

$$-\frac{1}{2}\sum_{i=1}^{u} \operatorname{tr}(B_i' \Sigma B_i)^{-1}(Y_i - A_i\xi B_i)'(Y_i - A_i\xi B_i)$$

where $\xi = [\underline{\xi}_1 \;\; \ldots \;\; \underline{\xi}_p]$

$$\frac{\partial \log\phi}{\partial\xi} = -\frac{1}{2}\sum_{i=1}^{u} \frac{\partial \, trU_i^{-1}P_iP_i'}{\partial\xi} \quad \text{where } P_i = (Y_i - A_i \; B_i)' \text{ , which is } (q_i \times n)$$

$$U_i = B_i' \; \Sigma \; B_i \text{, which is } (q_i \times q_i)$$

$$= \sum_{i=1}^{p} A_i' \; (Y_i - A_i\xi B_i)U_i^{-1}B_i'$$

$$\therefore \frac{\partial \log\phi}{\partial\xi_{k\ell}} = (k-\ell)^{th} \text{ element of } \sum_{i=1}^{p} A_i'(Y_i - A_i\xi B_i) \; U_i^{-1} \; B_i'$$

$$= \sum_{i=1}^{p} \sum_{z=1}^{q_i} \sum_{u=1}^{n_i} \sum_{t=1}^{q_i} a_{iuk}y_{iuk}u^{izt}b_{\ell t}$$

$$- \sum_{i=1}^{p} \sum_{z=1}^{q_i} \sum_{c=1}^{p} \sum_{u=1}^{n_i} \sum_{t=1}^{q_i} \sum_{u=1}^{m} a_{iuk}a_{iuw}\xi_{wc}b_{icz}u^{izt}b_{i\ell t}$$

where $B_i(p \times q_i) = ((b_{wt}))$, $A_i(n_i \times m) = ((a_{iuw}))$

$Y_i(n_i \times q_i) = ((y_{uz}))$, $U_i^{-1}(q_i \times q_i) = ((u^{xt}))$

$$\therefore \frac{\partial^2 \log \phi}{\partial\xi_{k\ell}\partial\xi_{qr}} = - \sum_{i=1}^{p} \sum_{z=1}^{q_i} \sum_{u=1}^{n_i} \sum_{t=1}^{q_i} a_{iuk}a_{iuq}b_{irz}u^{izt}b_{i\ell t} \; .$$

$$= - \sum_{i=1}^{p} \left(\sum_{u=1}^{n_i} a_{iuk}a_{iuq} \right) \left\{ \sum_{z=1}^{q_i} b_{irz} \left(\sum_{t=1}^{q_i} u^{izt}b_{i\ell t} \right) \right\}$$

$$= - \sum_{i=1}^{p} [(k-q)^{th} \text{ element of } A_i'A_i] \left\{ \sum_{z=1}^{q_i} b_{irz}[(z-\ell)^{th} \text{ element of } U_i^{-1}B_i] \right\}$$

$$= - \sum_{i=1}^{p} [(k-q)^{th} \text{ element of } A_i'A_i] \; [(r-\ell)^{th} \text{ element of } B_iU_i^{-1}B_i^{-1}]$$

$$\therefore nB(\xi,\Sigma) = \sum_{i=1}^{p} \{B_iU_i^{-1} \; B_i' \otimes A_i'A_i\}$$

$$\text{i.e., } nB(\xi,\Sigma) = \sum_{i=1}^{p} B_i[B_i' \Sigma B_i]^{-1} B_i' \otimes A_i'A_i$$

q.e.d.

5. THE EQUIVALENCE OF THE WALD STATISTIC AND HOTELLING'S TRACE STATISTIC

5.1 Theorem The Wald Statistic (W) obtained for SM is equivalent to Hotelling's trace criterion (T_o^2).

Proof:

$$T_o^2 = trS_HS_E^{-1} \text{ where } S_H = (C\hat{\xi})'[C(A'A)^{-1}C']^{-1}(C\hat{\xi})$$

$$S_E = Y'[I - A(A'A)^{-1}A']\,Y$$

Now

$$\hat{\Sigma} = \frac{1}{n} S_E \text{ is the M.L.E. of } \Sigma$$

Thus

$$\frac{1}{n} T_o^2 = tr\,(C\hat{\xi})'\,[C(A'A)^{-1}C']^{-1}\,(C\hat{\xi})\,\hat{\Sigma}^{-1}$$

Now

$$W = \begin{bmatrix} C\hat{\xi}_1 \\ \vdots \\ C\hat{\xi}_p \end{bmatrix}' \left\{ \begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix} [\hat{\Sigma}^{-1} \otimes A'A]^{-1} \begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix}' \right\}^{-1} \begin{bmatrix} C\hat{\xi}_1 \\ \vdots \\ C\hat{\xi}_p \end{bmatrix}$$

$$= \begin{bmatrix} C\hat{\xi}_1 \\ \vdots \\ C\hat{\xi}_p \end{bmatrix}' \left\{ \begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix} [\Sigma \otimes (A'A)^{-1}] \begin{bmatrix} C & & 0 \\ & \ddots & \\ 0 & & C \end{bmatrix}' \right\}^{-1} \begin{bmatrix} C\hat{\xi}_1 \\ \vdots \\ C\hat{\xi}_p \end{bmatrix}$$

$$= \begin{bmatrix} C\hat{\xi}_1 \\ \vdots \\ C\hat{\xi}_p \end{bmatrix}' [\hat{\Sigma} \otimes C(A'A)^{-1}C']^{-1} \begin{bmatrix} C\hat{\xi}_1 \\ \vdots \\ C\hat{\xi}_p \end{bmatrix}$$

$$= \begin{bmatrix} C\,\hat{\underline{\xi}}_1 \\ \vdots \\ C\,\hat{\underline{\xi}}_p \end{bmatrix}' [\hat{\Sigma}^{-1} \otimes [C(A'A)^{-1}C']^{-1}] \begin{bmatrix} C\,\hat{\underline{\xi}}_1 \\ \vdots \\ C\,\hat{\underline{\xi}}_p \end{bmatrix}$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \hat{\sigma}^{ij} (C\hat{\underline{\xi}}_i)' [C(A'A)^{-1}C']^{-1} (C\hat{\underline{\xi}}_k) \text{ where } \hat{\Sigma}^{-1} = ((\sigma^{ik}))$$

$$= \text{tr } (C\hat{\xi})' [C(A'A)^{-1}C']^{-1} (C\hat{\xi}) \hat{\Sigma}^{-1}$$

q.e.d.

6. CONCLUSION AND FURTHER RESEARCH

In Theorem 4.4, we have obtained a general test criterion for testing linear hypotheses in multivariate linear models which are incomplete in response variates and/or in which different response variates have different design matrices. When restricted to the standard MANOVA model, our test criterion becomes Hotelling's trace criterion. The general test criterion uses estimators of the unknown parameters which are asymptotically equivalent in probability. Maximum likelihood estimators could therefore be used, but even for simple cases of the general model (e.g., Monahan [2]) outside of the standard model, an iterative method is required. Computationally simpler estimators can probably be obtained, nevertheless, since we only require them to be asymptotically equivalent in probability to maximum likelihood estimators. In any case, further work needs to be done here. Also it is necessary to study the small sample properties of the test criterion.

REFERENCES

1. Allen, D.M. 1967. Multivariate analysis of nonlinear models. Institute of Statistics Mimeo Series No. 558, Univ. of North Carolina, Chapel Hill, North Carolina.

2. Monahan, I.P. 1961. Incomplete-variable designs. Unpublished thesis. Virginia Polytechnic Institute, Blacksburg, Virginia.

3. Roy, J. 1958, Step-down procedure in multivariate analysis. Annals of Mathematical Statistics 29, 1177-1187.

4. Roy, S.N. and J.N. Srivastava. 1964. Hierarchical and p-blick multiresponse designs and their analysis. Mahalanobis Dedicatory Volume. Pergamon Press, New York.

5. Srivastava, J.N. 1964. On a general class of designs for multiresponse experiments. Inst. Statistics Mimeo Series No. 402, Univ. of North Carolina, Chapel Hill, North Carolina.

6. Srivastava, J.N. 1965. A multivariate extension of the Gauss-Markov theorem. Ann. Inst. Statist. Math. 17, 63-66.

7. Srivastava, J.N. 1966. Incomplete multiresponse designs. Vol. 26, Sankhya.

8. Srivastava, J.N. 1966. On the extension of the Gauss-Markov theorem to complex multivariate linear models. Aerospace Research Laboratories Publication.

9. Srivastava, J.N. 1966. Some generalizations of multivariate analysis of variance. Symposium on Multivariate Analysis at Dayton, Ohio, 1965, pp. 129-145.

10. Srivastava, J.N. 1968. Some studies of intersection tests in multivariate analysis of variance. Paper presented at Symposium on Multivariate Analysis at Dayton, Ohio, 1968.

11. Trawinski, I.M. and R.E. Bargmann. 1964. Maximum likelihood estimation with incomplete multivariate data. Annals of Mathematical Statistics. 35, 647-658.

12. Wald, A. 1943. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society 54, 426-482.