density $dP^T_{(0,\rho)}/dP^T_{(0,0)}$ still depends on ρ, which will be explicitly evaluated in the case of $l=0$ in Chapter 6. This will suggest that there is no similar invariant test using (u_2, w_2) as well as (b_{111}, u_1, w_1) in a nontrivial way. It is noted that (b_{111}, u_1, w_1) is the statistic based on the first equation relevant to our problem from the viewpoint of invariance and the usual F test is given by

$$F=(n-k_1)b_{11}'G_{11}^{-1}b_{111}/l(u_1'u_1+w_1'w_1) > c. \tag{4.20}$$

4.3. *F-test.* Theorem 4.1 implies that the F test (4.20) based on the first equation alone is not LBI, much less UMPI. On the other hand, it is a similar invariant test under $\beta_{111}=0$ or it does not depend on any unknown parameter under $\beta_{111}=0$ and the power function is increasing in $\beta_{111}'G_{11}^{-1}\beta_{111}/\sigma_{11}$, which implies its unbiasedness. Here the question remains open whether it is admissible in the class of similar invariant tests.

Next let us find a situation in which an optimality is claimed for the F-test in the invariant model (4.16). In the covariance matrix of d in (4.16), if

$$F_{12}=0 \quad\text{and}\quad H_1'H_2=0 \tag{4.21}$$

holds, b_{111}, u_1 and u_2 become independent. But this alone does not immediately imply the UMPI property since w_1 and w_2 are still correlated. Here a necessary and sufficient condition for (4.21) to hold is easily shown through (4.11) to be

$$\tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}\tilde{X}_2'\tilde{X}_1 = \tilde{X}_1, \tag{4.22}$$

under which $R_2=0$ and so $H_2=0$ obtain. Hence $u_2\equiv 0$ follows. In this case, the problem of testing $\beta_{111}=0$ in the model (4.16) with $u_2\equiv 0$ is left invariant under the group $\mathcal{G}=\mathcal{O}(q_0)\times\mathcal{GT}(2)$ with

$$\mathcal{GT}(2) = \{A=\begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22}\end{pmatrix} \mid a_{11}a_{22}\neq 0\}$$

acting on (b_{111}, u_1, W) by

$$b_{111}\to a_{11}b_{111}, \quad u_1\to a_{11}u_1, \quad W\to CWA \quad\text{for } (C,A)\in\mathcal{G}$$

Under this group, w_2 is deleted by invariance. Thus we obtain

Theorem 4.2. Under condition (4.22), the F test (4.20) based on the first equation is UMPI for testing $\beta_{111}=0$ in the SUR model (4.1).

A special case for the condition (4.22) to hold is the case that $\tilde{X}_1$ is a submatrix of $\tilde{X}_2$, i.e., $\tilde{X}_2=[\tilde{X}_1,\tilde{X}_3]$. According to Revankar (1974), such a situation may arise in the context of a system of two structural equations, say, a system of demand and supply function equations. Thus, e.g., all the k_2 exogenous variables may appear in the "quantity" reduced form equation while, in view of a priori restrictions on the demand-supply equations, only k_1 ($<k_2$) of the k_2 exogenous variables may appear in the "price" reduced form equation. In such a situation, he considers the efficiency of the OLSE relative to the GLSE.

5. GMANOVA Problem Under Special Covariance Structure.

5.1. *Intra-class covariance structure.* Consider a GMANOVA model

$$Y = X_1BX_2 + E, \quad E\sim N(0, I_n\otimes\Omega) \tag{5.1}$$

where $X_1 : n\times k$ is of rank k and $X_2 : q\times p$ of rank q. Here we assume the intra-class covariance structure for Ω:

$$\Omega = \sigma^2(1-\rho)I_p + \sigma^2\rho ee' \tag{5.2}$$

where $e=e_p=(1,\ldots,1)'\in R^p$ and $-(p-1)^{-1}<\rho<1$. Under this model, let us consider the GMANOVA hypothesis

$$H: X_3 B X_4 = X_0 \tag{5.3}$$

where $X_3 : m_3 \times k$ is of rank m_3 and $X_4 : q \times r_4$ of rank r_4. An additional assumption is made on X_2 and X_4.

Assumption 5.1. Either (I) $X_2 e = 0$ or (II) $M_2 e = e$ and $X_4'(X_2X_2')^{-1}X_2 e = 0$ where $M_2 = X_2'(X_2X_2')^{-1}X_2$.

Under this assumption, the GLSE (generalized least square estimator) given by

$$\hat{B}_0(\Omega) = (X_1'X_1)^{-1}X_1'Y\Omega^{-1}X_2'(X_2\Omega^{-1}X_2')^{-1} \tag{5.4}$$

is identically equal to the OLSE (ordinary LSE) $\hat{B}_0(I)$ and the restricted GLSE given by

$$\tilde{B}(\Omega) = \hat{B}_0(\Omega) - U[X_3\hat{B}_0(\Omega)X_4 - X_0]V(\Omega) \tag{5.5}$$

is identically equal to the restricted OLSE $\tilde{B}(I)$ where

$$U = (X_1'X_1)^{-1}X_3'[X_3(X_1'X_1)^{-1}X_3']^{-1} \quad \text{and}$$

$$V(\Omega) = [X_4'(X_2\Omega^{-1}X_2')^{-1}X_4]^{-1}X_4'(X_2\Omega^{-1}X_2')^{-1}. \tag{5.6}$$

This is easily seen by using the following facts (see also Section 2 of Chapter 1).

Lemma 5.1. Let $\alpha = \rho[1 + \rho(p-1-e'M_2e)]$. Then

(a) $(X_2\Omega^{-1}X_2')^{-1} = \sigma^2(1-\rho)[(X_2X_2')^{-1} + \alpha(X_2X_2')^{-1}X_2ee'X_2'(X_2X_2')^{-1}]$

(b) $\Omega^{-1}X_2'(X_2\Omega^{-1}X_2')^{-1} = X_2'(X_2X_2')^{-1} - \alpha[ee'X_2'(X_2X_2')^{-1} - M_2ee'X_2(X_2X_2')^{-1}]$

Now let

$$a = \mathrm{tr}H, \quad b = e'He \quad \text{and} \quad d = \mathrm{tr}J, \quad \text{where}$$

$$H = (Y - X_1\hat{B}_0(I)X_2)'(Y - X_1\hat{B}_0(I)X_2) \quad \text{and} \tag{5.7}$$

$$J = [\tilde{B}(I) - \hat{B}_0(I)]'[\tilde{B}(I) - \hat{B}_0(I)].$$

Then the following theorem is rather straightforward and the proof is omitted.

Theorem 4.1. Under (I) or (II), the LRT (likelihood ratio test) is given by

$$T \equiv (np - n - kq)pd/m_3r_4[pa - b] > c \tag{5.8}$$

and it is UMPI (uniformly most powerful invariant). Under the assumption (I), the nonnull distribution is $F(m_3r_4,\ n(p-1)-kq : \lambda)$ and under (II), it is $F(m_3r_4,\ n(p-1)-q(k-1) : \lambda)$ where

$$\lambda = \frac{1}{2\sigma^2(1-\rho)}\mathrm{tr}[X_3BX_4 - X_0]'[X_3(X_1'X_1)^{-1}X_3']^{-1} \times [X_3BX_4 - X_0][X_4'(X_2X_2')^{-1}X_4]^{-1}. \tag{5.9}$$

The null distribution under each case are obtained by setting $\lambda = 0$.

In proof, a usual canonical version of the problem will simplify the argument.

Example 4.1. (See Example 1.1 in Chapter 1). Let $y_1(t)$ and $y_2(t)$ denote the GNP's or incomes per capita of Japan and USA respectively where $y(t) = (y_1(t), y_2(t))'$ is observed at time $t = 1, \ldots, n$ and $y(t)$'s are assumed to be independently distributed as $N(\mu(t), \Sigma)$. Further, $\mu_i(t)$ is assumed to obey a polynomial of $k-1$ degree:

$$\mu_i(t) = \beta_{i0} + \beta_{i1}t + \cdots + \beta_{ik-1}t^{k-1} \quad (i = 1, 2).$$

As is shown in Example 1.1 of Chapter 1, this model is written as a MANOVA model:

$$Y = X_1B + E, \quad E \sim N(0,\ I_n \otimes \Sigma), \quad \Sigma \in \mathcal{S}_+(2)$$

where $Y = [y(1), \ldots, y(n)]' : n \times 2$, $X_1 = (x_{ij}) : n \times k$ with $x_{ij} = i^{j-1}$ and $B = (\beta_{ij}) : k \times 2$ (see (1.16) of Chapter 1). Here it is assumed that the variances of $y_i(t)$'s are the same so that $\Sigma=\sigma^2(1-\rho)I+\sigma^2\rho e_2e_2'$, where the positive (or negative) correlation ρ may be related to the interaction through the trades. Under this assumption, we may test the hypothesis that the growth processes of GNP's or per-capita incomes of the both countries follow the same pattern except the intercepts. Then the hypothesis is expressed as

$$X_3BX_4=0$$

where $X_3=[0,\ I_{k-1}] : (k-1)\times k$ and $X_4(-1,1)'$. This is the hypothesis that $\beta_{1j}=\beta_{2j}$ $(j=1,\dots,k-1)$ but β_{10} and β_{20} may be different. Now to verify the assumption (II), note that $X_2=I_2$. Hence $M_2e_2=e_2$ and $X_4'(X_2X_2')^{-1}X_2e=X_4'e_2=(-1,1)e_2=0$. Therefore (II) is satisfied, and so the test based on (5.8) is the LRT and UMPI. In this case, $q=p=2$, $m_3=k-1$ and $r_4=1$. So the null distribution is $F(k-1,\ n-2(k-1))$. Even if the equality of the variances is not assumed, it follows from a result in Section 7 of Chapter 3 that the LRT in this case is UMPI because $X_2=I$ implies $p_3=0$ and because $p_2=1$, where $p_2=r_4$ and $p_3=p-q$ (see Cahpter 3). However, if we start with 3 countries, say $y(t)=(y_1(t),\ y_2(t),\ y_3(t))'$, the above argument holds with $X_2=I_3$ and $X_4=-\begin{pmatrix}-1 & 1 & 0\\ 0 & -1 & 1\end{pmatrix}'$ provided Σ is of the intra-class covariance structure. And the LRT is UMPI.

5.2. *Rao's covariance structure.* More generally let us consider the GMANOVA hypothesis $X_3BX_4=X_0$ in the GMANOVA model (5.1) under Rao's covariance matrix:

$$\Omega = X_2'\Gamma X_2 + \bar{X}_2'\Delta\bar{X}_2 \tag{5.10}$$

where $p>q$, $\Gamma\in\mathcal{S}_+(q)$, $\Delta\in\mathcal{S}_+(p-q)$, and $\bar{X}_2$ is a $(p-q)\times p$ matrix of rank $p-q$ satisfying $X_2\bar{X}_2'=0$. To obtain a canonical form, define

$$C = [X_2'(X_2X_2')^{-1},\ \bar{X}_2'(\bar{X}_2\bar{X}_2')^{-1}].$$

Then post-multiplying (5.1) by C, we obtain

$$\overset{q}{[\tilde{Y}_1},\ \overset{p-q}{\tilde{Y}_2]} = [X_1B, 0] + [\tilde{E}_1, \tilde{E}_2],\quad \tilde{E}\sim N\left(0,\ I_n\otimes\begin{bmatrix}\Gamma & 0\\ 0 & \Delta\end{bmatrix}\right)$$

where $YC=\tilde{Y}=[\tilde{Y}_1, \tilde{Y}_2]$ and $EC=\tilde{E}=[\tilde{E}_1, \tilde{E}_2]$. Hence

$$\begin{cases}\tilde{Y}_1\sim N(X_1B, I_n\otimes\Gamma),\quad \tilde{Y}_2\sim N(0, I_n\otimes\Delta),\\ \tilde{Y}_1 \text{ and } \tilde{Y}_2 \text{ are independent.}\end{cases} \tag{5.11}$$

This is a canonical form of the GMANOVA model (5.1) under the covariance matrix (5.10). Here letting the group $\mathcal{G}_1=\mathcal{O}(n)\times\mathcal{G}l(p-q)$ act on $\tilde{Y}_2$ by $\tilde{Y}_2\to K\tilde{Y}_2A'$ where $(K, A)\in\mathcal{G}_1$, we get rid of $\tilde{Y}_2$ since $\mathcal{G}_1$ acts transitively on the space of $\tilde{Y}_2$. Thus, the problem of testing $X_3BX_4=X_0$ in the GMANOVA model (5.1) under (5.10) is reduced to the problem of testing it in the MANOVA model $\tilde{Y}_1=X_1B+\tilde{E}_1$, and the latter problem has been treated in 7.4 and 8.1 of Chapter 3. That is, the test with critical region $\operatorname{tr} T_1(I+T_1)^{-1}>c$ is LBI, where T_1 is given by (7.15) in Chapter 3 with Y replaced by $\tilde{Y}_1$, and when $\min(m_3,r_4)=1$, it is UMPI.

5.3. *Covariance structure under a growth curve model.* Let us consider Example 1.2 in Section 1 of Chapter 1 where the model is expressed as

$$Y = X_1BX_2 + E,\quad E\sim N(0, I_n\otimes[I_7\otimes\Sigma]) \tag{5.12}$$

where $Y : n\times 7p$, $B=[B_1, B_2] : k\times 2p$, $X_1 : n\times k$ is given by (1.16) in Chapter 1 and

$$X_2=\begin{pmatrix}I & I & I & I & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & I & I & I\end{pmatrix} : 2p\times 7p. \tag{5.13}$$

Under this model, we first consider the hypothesis

(5.14) $X_3BX_4=0$ with $X_3=[0,I_{k-1}] : (k-1)\times k$, $X_4=\begin{bmatrix}I\\ -I\end{bmatrix} : 2p\times p$, which means that the growth pattern between the two groups is the same except the initial levels or intercepts. It is noted that since the covariance matrix of each row of E is of the structure $I_7 \otimes \Sigma$ in (5.12), the problem is not completely a GMANOVA problem treated in Chapter 3. To obtain a canonical form of this problem, let $\overline{X}_2 = [F_{ij}]$ be a $5p \times 7p$ matrix with $F_{ij} : p \times p$ ($i=1,\dots,5 : j=1,\dots,7$) such that $F_{ij}=0$ or I or $-I$, $\overline{X}_2'\overline{X}_2 = I_{5p}$ and $X_2\overline{X}_2'=0$. In fact, such a matrix $\overline{X}_2$ exists. Then letting $C=[X_2', \overline{X}_2'] : 7p \times 7p$ and post multiplying (5.12) by C, we obtain

$$(5.15) \qquad \tilde{Y}_i = a_i X_1 B_i + \tilde{E}_i, \quad \tilde{E}_i \sim N(0, I_n \otimes \Sigma) \ (i=1,2)$$

$$(5.16) \qquad \tilde{Y}_j = \tilde{E}_j, \qquad \tilde{E}_j \sim N(0, I_n \otimes \Sigma) \ (j=3,\dots,7)$$

where $YC = \tilde{Y} = [\tilde{Y}_1, \dots, \tilde{Y}_7]$ with $\tilde{Y}_i : n \times p$, $EC = \tilde{E} = [\tilde{E}_1, \dots, \tilde{E}_7]$ with $\tilde{E}_i : n \times p$, $a_1=4$, $a_2=3$ and $\tilde{Y}_i$'s are independent. Here (5.16) is expressed as a MANOVA model;

$$(5.17) \qquad \begin{pmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \end{pmatrix} = \begin{pmatrix} 4X_1 & 0 \\ 0 & 3X_1 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} + \begin{pmatrix} \tilde{E}_1 \\ \tilde{E}_2 \end{pmatrix}, \quad \begin{pmatrix} \tilde{E}_1 \\ \tilde{E}_2 \end{pmatrix} \sim N(0, I_{2n} \otimes \Sigma).$$

Hence by a usual method, it is also expressed as

$$(5.18) \qquad \begin{array}{c} \begin{pmatrix} Y_1^* \\ Y_2^* \end{pmatrix} \end{array} \begin{array}{l} 2k \\ 2n-2k \end{array} \sim N\left(\begin{pmatrix} \Theta^* \\ 0 \end{pmatrix}, I_{2n} \otimes \Sigma\right)$$

where for some $H_1 \in \mathcal{G}l(2k)$ and $K_1 \in \mathcal{O}(2n)$.

$$(5.19) \qquad \Theta^* = H_1 \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad \begin{pmatrix} 4X_1 & 0 \\ 0 & 3X_1 \end{pmatrix} = K_1 \begin{bmatrix} H_1 \\ 0 \end{bmatrix}$$

Consequently from (5.16) and (5.18), the model (5.12) is finally expressed as a canonical form of a MANOVA model

$$(5.20) \qquad \begin{pmatrix} Z_1^* \\ Z_2^* \end{pmatrix} \begin{array}{l} 2k \\ 7n-2k \end{array} \sim N\left(\begin{pmatrix} \Theta^* \\ 0 \end{pmatrix}, I_{7n} \otimes \Sigma\right)$$

where $Z_1 = Y_1^*$, $Z_2 = [Y_2^{*\prime}, \tilde{Y}_3', \dots, \tilde{Y}_7']'$. On the other hand, the hypothesis (5.14) is correspondingly rewritten as

$$0 = \tilde{X}_3 \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \tilde{X}_3 H_1^{-1} \Theta^* = [H_2, 0] K_2 \Theta^*$$

for some $H_2 \in \mathcal{G}l(k-1)$ and $K_2 \in \mathcal{O}(2k)$, where $\tilde{X}_3 = [0, I_{k-1}, 0, -I_{k-1}] : (k-1) \times 2k$. Therefore the problem is now to test $\Theta_1 = 0$ under the model

$$(5.21) \qquad \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} \begin{array}{l} k-1 \\ 2k-k+1 \\ 7n-2k \end{array} \sim N\left(\begin{pmatrix} \Theta_1 \\ \Theta_2 \\ 0 \end{pmatrix}, I_{7n} \otimes \Sigma\right) \quad \text{where } \Theta = K_2 \Theta^*$$

which is nothing but a MANOVA problem or a special case of the GMANOVA problem treated in Chapter 3.

Chapter 5

TESTS OF INDEPENDENCE IN A GMANOVA MODEL

1. Testing Independence in a GMANOVA Model

1.1. *Problem.* As has been seen in Section 5 of Chapter 4, the GMANOVA problem is reduced through invariance to the MANOVA problem if the covariance matrix Ω is of Rao's covariance structure. On the other hand, as will be shown soon, the problem of testing the hypothesis that Ω is of the structure is reduced to the problem of testing independence between two sets of variates with extra data on the first set. Hence the analysis of our problem also solves the problem of testing independence with incomplete data. In fact, in Section 2 based on this fact, we slightly extend the problem to the problem of testing independence between two sets of variates where independent extra data are available on the both sets, and analyze the extended problem in Section 4 and thereafter. In a line with the arguments made in Chapter 3, the LBI (locally best invariant) tests are derived and the local minimaxity, admissibility and robustness of the tests are considered. The asymptotic expansions of the null distributions are also derived. It is here interesting to point out that the LRT (likelihood ratio test)

completely ignores the extra data in the original problem.

The problem of testing on Rao's covariance structure is in itself regarded as the the problem of a choice between the GLSE (generalized least squares estimator) and the OLSE (ordinary LSE) in a GMANOVA model, which has been discussed in Chapter 1. But we shall review this soon and get a canonical form. Also the problem is related to the problem of testing independence in a classification model with covariates.

1.2. *Testing on Rao's covariance structure.* Let us consider the hypothesis that the covariance matrix Ω in the growth curve model or GMANOVA model

$$(1.1)\qquad Y = X_1BX_2 + E, \quad E \sim N(0, I_n \otimes \Omega)$$

is of the form

$$(1.2)\qquad \Omega = X_2'\Gamma X_2 + \bar{X}_2'\Delta\bar{X}_2$$

where $\Gamma \in \mathcal{S}_+(p_1+p_2)$ and $\Delta \in \mathcal{S}_+(p_3)$. Here we follow the notation adopted in Chapter 3, $X_1 : n\times(n_1+n_2)$ with $\operatorname{rank}(X_1)=n_1+n_2$, $X_2 : (p_1+p_2)\times p$ with $\operatorname{rank}(X_2)=p_1+p_2$, $p=p_1+p_2+p_3$, $B : (n_1+n_2)\times(p_1+p_2)$ and $\bar{X}_2 : p_3\times p$ with $X_2\bar{X}_2'=0$ and $\operatorname{rank}(\bar{X}_2)=p_3$. As has been remarked in Chapter 1, the covariance structure in (1.2) is often called Rao's covariance structure (see Rao (1967)). From Corollary 2.1 in Chapter 1, the covariance structure in (1.2) is a necessary and sufficient condition for the GLSE defined by

$$(1.3)\qquad \hat{B}(\Omega) = (X_1'X_1)^{-1}X_1'Y\Omega^{-1}X_2'(X_2\Omega^{-1}X_2')^{-1}$$

to be identically equal to the OLSE $\hat{B}(I)$. In addition, by Theorem 2.3 in Chapter 1, it is also a necessary and sufficient condition for $S(\Omega)\equiv S(I)$ where

$$S(\Omega) = (Y - X_1\hat{B}(\Omega)X_2)'(Y - X_1\hat{B}(\Omega)X_2)/n.$$

In other words, if Ω in (1.1) is of the form (1.2), $(\hat{B}(I),\ S(I))$ is as efficient as $(\hat{B}(\Omega),\ S(\Omega))$. For from

$$(1.4)\quad \begin{aligned} &\mathrm{tr}(Y-X_1BX_2)'(Y-X_1BX_2)\Omega^{-1} \\ &= \mathrm{tr}(Y-X_1\hat{B}(\Omega)X_2)'(Y-X_1\hat{B}(\Omega)X_2)\Omega^{-1} \\ &\quad + 2\mathrm{tr}(Y-X_1\hat{B}(\Omega)X_2)'X_1(\hat{B}(\Omega)-B)X_2\Omega^{-1} \\ &\quad + \mathrm{tr}X_2'(\hat{B}(\Omega)-B)'X_1'X_1(\hat{B}(\Omega)-B)X_2\Omega^{-1}, \end{aligned}$$

$(\hat{B}(\Omega),S(\Omega))\equiv(\hat{B}(I),S(I))$ is sufficient for (B,Ω) under the normality in (1.1), since the pdf of Y is given by

$$(1.5)\quad \begin{aligned} &f(Y\mid B,\Omega) \\ &=(2\pi)^{-np/2}|\Omega|^{-n/2}\exp[-\frac{1}{2}\mathrm{tr}(Y-X_1BX_2)'(Y-X_1BX_2)\Omega^{-1}]. \end{aligned}$$

Further, (1.4) and (1.5) show that $\hat{B}(I)$ is the MLE (maximum likelihood estimator) of B under (1.2). It is noted that in the MANOVA model where $p_3=0$ and $X_2=I$, $\hat{B}(\Omega)\equiv(X_1'X_1)^{-1}X_1'Y$ and $S(\Omega)\equiv Y'(I-X_1(X_1'X_1)^{-1}X_1')Y/n$, which do not depend on Ω.

On the problem of testing (1.2), very little work has been done. Lee and Geisser (1972) derived the LRT (likelihood ratio test), Eaton and Kariya (1975, 1983) considered the problem via invariance and derived the LBI (locally best invariant) test, and Sarkar (1980) and Chou and Lo (1983) considered the local minimax property of the LBI test.

To obtain a canonical form, let $\bar{X}_1: n\times n_3$ be of rank n_3 and satisfy $\bar{X}_1'X_1=0$ and let

$$C_1=[X_1(X_1'X_1)^{-1/2},\ \bar{X}_1(\bar{X}_1'\bar{X}_1)^{-1/2}] \quad\text{and}$$
$$C_2=[X_2'(X_2X_2')^{-1/2},\ \bar{X}_2'(\bar{X}_2\bar{X}_2')^{-1/2}]$$

Then $C_1\in\mathcal{O}(n)$ and $C_2\in\mathcal{O}(p)$. Further let

$$(1.6)\quad W=C_1'YC_2,\quad \Sigma=C_2'\Omega C_2,\quad \mu=(X_1'X_1)^{1/2}B(X_2X_2')^{1/2}$$

and partition W and Σ as

$$W=\begin{array}{cc} & \begin{array}{cc} q & p-q \end{array} \\ & \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \begin{array}{c} m \\ n_3 \end{array} \end{array},\qquad \Sigma=\begin{array}{cc} & \begin{array}{cc} q & p-q \end{array} \\ & \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{array}{c} q \\ p-q \end{array} \end{array}$$

where $q=p_1+p_2$ and $m=n_1+n_2$. With this relabelling, we have

$$(1.7)\quad W\sim N\begin{pmatrix} \mu & 0 \\ 0 & 0 \end{pmatrix},\ I_n\otimes\Sigma).$$

The null hypothesis that Ω is of the form (1.2) becomes $H_0: \Sigma_{12}=0$ because with Ω in (1.2)

$$\Sigma=C_2'\Omega C_2=\begin{pmatrix} (X_2X_2')^{1/2}\Gamma(X_2'X_2)^{1/2} & 0 \\ 0 & (\bar{X}_2\bar{X}_2')^{1/2}\Delta(\bar{X}_2\bar{X}_2')^{1/2} \end{pmatrix}$$

This testing problem is clearly invariant under the translation

$$W_{11}\to W_{11}+A \quad\text{with}\quad A: m\times q$$

since B or μ is completely unknown, and a maximal invariant under this group of translations is $\{W_{12},(W_{21},W_{22})\}$. Obviously W_{12} is independent of (W_{21},W_{22}) and

$$(1.8)\quad W_{12}\sim N(0,\ I_m\otimes\Sigma_{22})\quad\text{and}\quad (W_{21},W_{22})\sim N(0,\ I_{n_3}\otimes\Sigma).$$

Here regarding (W_{21},W_{22}) as a complete data for testing $H_0: \Sigma_{12}=0$, (1.8) is a situation in which an additional data W_{12} which marginally carries the informationon on Σ_{22} is available. This fact motivates an extension of the problem made in the next section.

1.4. *MANOVA model with a GMANOVA restriction.* A related problem is the problem described in Section 2.5 of Chapter 1 where a MANOVA model with a linear restriction

$$(1.9)\quad Y=X_1B+E,\quad E\sim N(0,I\otimes\Omega),\quad R_1BR_2=R_0$$

was considered. And the problem is to test

$$\Omega = R_2\Gamma R_2' + P_2\Delta P_2'. \tag{1.10}$$

Here $X_1 : n\times k$ of rank k, $B : k\times p$, $\Omega\in\mathcal{S}_+(p)$, $R_1 : r_1\times k$ of rank r_1, $R_2 : p\times r_2$ of rank $r_2<p$, $\Gamma\in\mathcal{S}_+(r_2)$, $\Delta\in\mathcal{S}_+(p-r_2)$ and $P_2 : p\times(p-r_2)$ which satisfies $P_2'R_2=0$ and $P_2'P_2=I_{p-r_2}$. By Theorem 2.6 in Chapter 1, the covariance structure (1.10) is a necessary and sufficient condition for the BLUE $\tilde{B}(\Omega)$ to be identically equal to the OLSE $\tilde{B}(I)$ under the model (1.9), where

$$\tilde{B}(\Omega) = \hat{B}_0 - (X_1'X_1)^{-1}X_1'[R_1\hat{B}_0R_2 - R_0]R_2'(R_2\Omega^{-1}R_2')^{-1}R_2\Omega^{-1}$$

and $\hat{B}_0=\hat{B}(I)=(X_1'X_1)^{-1}X_1'Y$. To reduce the problem to a canonical form, we let $\bar{X}_1$ denote an $n\times(n-k)$ matrix satisfying $\bar{X}_1\bar{X}_1'=I-X_1(X_1'X_1)^{-1}X_1'$ and $\bar{X}_1'\bar{X}_1=I_{n-k}$ and let P_1 be a $(k-r_1)\times k$ matrix satisfying $P_1P_1'=I_{k-r_1}$, $P_1R_1'=0$. Then Y is in one-one correspondence with $(\hat{B}_0, Z_2'Y)$, which is in turn in one-one correspondence with

$$W\equiv\begin{pmatrix}W_1\\W_2\\W_3\end{pmatrix}\equiv\begin{pmatrix}W_{11} & W_{12}\\W_{21} & W_{22}\\W_{31} & W_{32}\end{pmatrix}\equiv\begin{array}{c} \begin{pmatrix}R_1\hat{B}_0R_2 & R_1\hat{B}_0P_2\\P_1\hat{B}_0R_2 & P_1\hat{B}_0P_2\\\bar{X}_1'YR_2 & \bar{X}_1'YP_2\end{pmatrix}\begin{matrix}r_1\\k-r_1\\n-k\end{matrix} \\ \begin{matrix}r_2 & \qquad p-r_2\end{matrix}\end{array} \tag{1.11}$$

Since

$$\begin{pmatrix}\hat{B}_0\\\bar{X}_2'Y\end{pmatrix} \sim N\left(\begin{pmatrix}B\\0\end{pmatrix}, \begin{pmatrix}(X_1'X_1)^{-1} & 0\\0 & I\end{pmatrix}\otimes\Omega\right),$$

$$W \sim N\left(\begin{pmatrix}R_0 & R_1BP_2\\P_1BR_2 & P_1BP_2\\0 & 0\end{pmatrix}, \begin{pmatrix}R_1(X_1'X_1)^{-1}R_1' & R_1(X_1'X_1)^{-1}P_1' & 0\\P_1(X_1'X_1)^{-1}R_1' & P_1(X_1'X_1)^{-1}P_1' & 0\\0 & 0 & I_{n-k}\end{pmatrix}\otimes\Sigma\right)$$

where

$$\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\\Sigma_{21} & \Sigma_{22}\end{pmatrix}\equiv\begin{pmatrix}R_2'\Omega R_2 & R_2'\Omega P_2\\P_2'\Omega R_2 & P_2'\Omega P_2\end{pmatrix}.$$

Therefore since (1.10) is equivalent to $\Sigma_{12}=R_2'\Omega P_2=0$, the problem of testing (1.10) in the model (1.9) is equivalent to the problem

of testing the independence $\Sigma_{12}=0$ based on W. Discarding (W_{12}, W_{21}, W_{22}) by invariance as in 1.3, since R_0 is known, which we assume to be zero henceforth, the problem is now to test $\Sigma_{12}=0$ based on W_{11} and (W_{31}, W_{32}) or equivalently $\tilde{W}_{11}$ and (W_{31}, W_{32}) where $\tilde{W}_{11}=(R_1(X_1'X_1)^{-1}R_1')^{-1/2}W_{11}$. Under this situation, the model is reduced

$$\tilde{W}_{11}\sim N(0, I_{r_1}\otimes\Sigma_{11}),\quad (W_{31}, W_{32})\sim N(0, I_{n-k}\otimes\Sigma)\ \text{and}\ \tilde{W}_{11}\ \text{and}\ (W_{31}, W_{32})\ \text{are independent}, \tag{1.12}$$

hence the problem is nothing but the problem treated in the above (see (1.8)).

An example of interest is found in the following example.

Example 1.1. In Examples 1.5 and 2.2 of Chapter 1 and Example 2.2 of Chapter 4, it is observed that the model considered in the problem of testing the equality of means with covariates is an extended GMANOVA model and given by $Y=X_1B+E$ with $R_1BR_2=0$ (see Example 2.2 of Chapter 4). And as has been pointed out in Example 2.2 of Chapter 1, the problem of testing the hypothesis that Σ is of the form (1.10) is equivalent to the problem of testing $\Sigma_{12}=0$, under which no information is expected through covariates. In particular, when $\Sigma_{12}=0$ in the original problem, the discriminant function making use of covariates will be less efficient than the usual discriminant function ignoring them. This motivates this testing problem and as has been just seen, the problem is reduced to the problem of testing $\Sigma_{12}=0$ in the model (1.12).

2. The Problem of Testing Independence With Extra Data.

2.1. *Introduction.* In Section 1, it is observed that the problem of testing on Rao's covariance structure in a GMANOVA model is reduced to the problem of testing independence with extra data.

Based on this fact, we here extend the problem. Because of their common occurence in practice, there has been a continuing interest in inference problems where there is missing or extra data. The causes for the data to be missing or extra will not be discussed explicitly here, but will be implicit in our assumptions concerning the likelihood function of the data (see Section 3 of Chapter 4). For an illuminating discussion of such issues, the readers are referred to Rubin (1963).

The problems treated in this section concern data on p coordinates which are partitioned into two groups of p_1 and p_2 coordinates— so $p_1+p_2=p$ and $1\leq p_i<p$ $(i=1,2)$. Let $X_1,\dots,X_n$ be a random sample from $N_p(\mu,\Sigma)$ and let $X_{i1},\dots,X_{im_i}$ be a random sample from $N_{p_i}(\mu_i,\Sigma_{ii})$ $(i=1,2)$. That is, the model we consider here is

$$\begin{aligned}&\tilde{X}: n\times p \sim N_p(e_n\mu', I_n\otimes\Sigma)\\ (2.1)\qquad &\tilde{X}_i: m_i\times p_i \sim N_{p_i}(e_{m_i}\mu_i', I_{m_i}\otimes\Sigma_{ii})\quad (i=1,2)\\ &\tilde{X},\ \tilde{X}_1 \text{ and } \tilde{X}_2 \text{ are independent}\end{aligned}$$

where $\tilde{X}'=[X_1,\dots,X_n]$, $\tilde{X}_i'=[X_{i1},\dots,X_{im_i}]$ $(i=1,2)$, $e_k=(1,\dots,1)'\in R^k$,

$$(2.2)\qquad \mu=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\begin{matrix}p_1\\ p_2\end{matrix} \quad\text{and}\quad \Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\begin{matrix}p_1\\ p_2\end{matrix}.$$

Thus, there are n "complete" observations, m_1 "extra" observations on the first p_1 coordinates, and m_2 "extra" observations on the last p_2 coordinates. When m_1 (or equivalently m_2) is zero, then the data is in a triangularly partitioned form. Under the assumption of multivariate normality, Bhargava (1962) derived maximum likelihood estimators (MLE's) and likelihood ratio tests (LRT's) for a number of problems when the data has a general triangular form. This triangular form permits the explicit calculation of MLE's and LRT's along with the relevant distribution theory. Morrison and Bhoj (1973) discuss the power of the LRT for testing a mean vector

is zero when $m_1=0$. Ordinarily, likelihood methods are proposed for problems with missing data especially when the normal distribution is involved. However, in some situations, the likelihood equations cannot be solved explicitly. The articles by Hartley and Hocking (1971) and Kariya, Krishnaiah and Rao (1983) provide good overviews of the subject and extensive bibliographies.

In this section, we discuss the problem of testing $\Sigma_{12}=0$ in the model (2.1), and in the next section we derive a locally best invariant test. When $m_1=0$ and $m_2>0$ (or $m_1>0$ and $m_2=0$), this test is different from the LRT. (When $m_1>0$ and $m_2>0$, the LRT is not derived explicitly). There the LRT does not utilize the "extra" data at all and is identical to the LRT when $m_1=m_2=0$. As we see next, the model and the problem are extended versions of those stated in Section 1 in association with Rao's covariance structure, which is why we treat the present problem in this book.

2.2. *Canonical form.* We transform the data $\tilde{X}$, $\tilde{X}_1$ and $\tilde{X}_2$ into a canonical form. Let Γ be an $n\times n$ orthogonal matrix with first row $e_n'/\sqrt{n}$. Then the transpose of the first row of the matrix $\Gamma\tilde{X}$ has a $N(\sqrt{n}\mu,\Sigma)$ distribution and is independent of the remaining $(n-1)$ rows which are independently and identically distributed as $N(0,\Sigma)$. Let $U\in R^p$ be the transpose of the first row of $\Gamma\tilde{X}$ multiplied by $1/\sqrt{n}$ and let $W:(n-1)\times p$ be the remaining $(n-1)$ rows of $\Gamma\tilde{X}$. Then U and W are independent with

$$U \sim N(\mu, \tfrac{1}{n}\Sigma) \quad\text{and}\quad W\sim N(0, I_{(n-1)}\otimes\Sigma).$$

Transforming $\tilde{X}_i$ in a similar manner leads to $U_i\in R^{p_i}$ and W_i: $(m_i-1)\times p_i$ which are independent and satisfy

$$U_i\sim N(\mu_i, \tfrac{1}{m_i}\Sigma_{ii}) \quad\text{and}\quad V_i \sim N(0, I_{(m_i-1)}\otimes\Sigma_{ii}).\ (i=1,2)$$

In summary, the complete and partial data $\tilde{X}$, $\tilde{X}_1$ and $\tilde{X}_2$ can be relabeled to yield U, U_1, U_2 and W, W_1, W_2 with the given distributions.

With the above discussion in mind, we now redescribe the canonical form for our extra data problem. The mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ being partitioned as (2.2), we consider independent random vectors $Y \in R^p$, $X_i \in R^{p_i}$, $i=1,2$ and independent random matrices $\boldsymbol{Z} : n \times p$, $\boldsymbol{Z}_i : m_i \times p_i$, $i=1,2$ such that

$$
(2.3)\qquad
\begin{aligned}
&Y : p \times 1 \sim N(\mu, c\Sigma) \\
&X_i : p_i \times 1 \sim N(\mu_i, c_i\Sigma_{ii}) \ (i=1,2) \\
&\boldsymbol{Z} : n \times p \sim N(0, I_n \otimes \Sigma) \\
&\boldsymbol{Z}_i : m_i \times p_i \sim N(0, I_{m_i} \otimes \Sigma_{ii}) \ (i=1,2)
\end{aligned}
$$

where c, c_1 and c_2 are known positive constants.

Under this model, the problem of testing $\Sigma_{12}=0$ is considered. Observations which are represented in the form (2.3) will be said to be in canonical form. Throughout this Chapter it is assumed that sample sizes are large enough so maximum likelihood estimators exist, that is, we assume $\min\{n+m_1, n+m_2\} > \max\{p_1, p_2\}$ in the canonical model (2.3).

It is noted that the canonically reduced problem considered in Section 1 is a special case of this problem. In fact, when Y and X_i's are absent and when $\boldsymbol{Z}_2$ is absent, the problem here is reduced to the problem in Section 1.

3. LBI Test

3.1. *Derivation of LBI test.* Let us consider the problem of testing for independence based on data in canonical form. The canonical form will be of the type described by (2.3), but for simplicity our main discussion will be concerned with the following data: Consider three independent random matrices $\boldsymbol{Z} : n \times p$, $\boldsymbol{Z}_1 : m_1 \times p_1$ and $\boldsymbol{Z}_2 : m_2 \times p_2$ with $p_1 + p_2 = p$ satisfying

$$
(3.1)\qquad
\begin{cases}
\boldsymbol{Z} \sim N(0, I_n \otimes \Sigma) \\
\boldsymbol{Z}_i \sim N(0, I_{m_i} \otimes \Sigma_{ii}) \ (i=1,2).
\end{cases}
$$

Here, the unknown covariance matrix $\boldsymbol{\Sigma}$ has been partitioned into $\Sigma_{ij} : p_i \times p_j$ for $i, j=1,2$. The data (3.1) arises from the data described in Section 2 by assuming that the mean vector μ is known and when $m_2=0$, it also corresponds to the canonical model in the problem of testing on Rao's covariance structure stated in Section 1. The problem is here to test

$$H_0 : \Sigma_{12}=0 \quad \text{versus} \quad H_1 : \Sigma_{12} \neq 0.$$

This testing problem is invariant under the group

$$\mathcal{G} = \left\{ g = A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},\ A_i \in \mathcal{G}l(p_i),\ i=1,2 \right\},$$

The action of $\mathcal{G}$ on a sample point is

$$g(\boldsymbol{Z}, \boldsymbol{Z}_1, \boldsymbol{Z}_2) = (\boldsymbol{Z}A', \boldsymbol{Z}_1A_1', \boldsymbol{Z}_2A_2')$$

with the action on Σ being

$$g(\Sigma) = A\Sigma A'.$$

A maximal invariant parameter is the vector $\delta = (\delta_1^2, \ldots, \delta_q^2)'$ with $q = \min\{p_1, p_2\}$ where $\delta_1^2 \geq \ldots \geq \delta_q^2$ are the q-largest eigenvalues of $\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}$. The main result of this section is that there exists an LBI test of the H_0 versus H_1. To describe this result, let $\mathcal{D}_\alpha$ be the class of all the size α $\mathcal{G}$-invariant test functions. Also, set

$$\boldsymbol{Z}'\boldsymbol{Z} = V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{Z}_i'\boldsymbol{Z}_i = S_{ii} \ (i=1,2).$$

The following result is a result corresponding to Theorem 8.1 for the GMANOVA problem in Chapter 3.

Theorem 3.1. Let $\tau=\sum_1^q\delta_i^2$. For $\phi\in\mathscr{D}_\alpha$, the power function of ϕ at δ, say $\pi(\phi,\delta)$, has the form

$$\pi(\phi,\delta)=\alpha+B(\phi)\tau+o(\delta,\phi) \tag{3.2}$$

where

$$\lim_{\delta\to 0}\sup_\phi|o(\delta,\phi)|=0 \quad\text{and}\quad B(\phi)=E_0\Big[\frac{1}{2}\phi U_0\Big],$$

and

$$U_0=\frac{(n+m_1)(n+m_2)}{p_1p_2}\operatorname{tr}(V_{11}+S_{11})^{-1}V_{12}(V_{22}+S_{22})^{-1}V_{21} + n-\sum_{i=1}^2\Big(\frac{n+m_i}{p_i}\Big)\operatorname{tr}(V_{ii}+S_{ii})^{-1}V_{ii}. \tag{3.3}$$

The size α test which rejects for $U_0>k$ is an LBI size α test.

Proof. The representation (3.2) is established in 3.3. That rejecting for $U_0>k$ gives an LBI test follows immediately from (3.2) by maximizing $B(\phi)$ and applying the Generalized Neyman-Pearson Lemma (see Section 1 of Chapter 2).

In the discussion below, the situation treated by Theorem 3.1 for the data (3.1) will be called **case (0)**. We now turn to a brief discussion of some other cases of interest.

Case (i): This refers to case (0) when $m_2=0$ so that the data matrix S_{22} is not available. A direct analogue of Theorem 3.1 shows that the test which rejects for large values of

$$U_1\equiv\frac{(n+m_1)n}{p_2p_1}\operatorname{tr}V_{22}^{-1}V_{21}(V_{11}+S_{11})^{-1}V_{12} -\frac{n+m_1}{p_1}\operatorname{tr}(V_{11}+S_{11})^{-1}V_{11} \tag{3.4}$$

is an LBI test for testing $H_0:\Sigma_{12}=0$ versus $H_1:\Sigma_{12}\neq 0$. This case corresponds to the problem of testing on Rao's covariance structure

described in Section 1.

Case (ii): In this case we consider the data given in (2.3) and as usual, let $V=Z'Z$, and $S_{ii}=Z_i'Z_i$ $(i=1,2)$. Further, let

$$U=\begin{pmatrix}U_1\\U_2\end{pmatrix}\equiv\begin{pmatrix}(c+c_1)^{-1/2}(Y_1-X_1)\\(c+c_2)^{-1/2}(Y_2-X_2)\end{pmatrix}$$

and set $b^2=c^2/(c+c_1)(c+c_2)$. Define the statistic U_2 by

$$\begin{aligned}U_2=&\frac{(n+m_1+1)(n+m_2+1)}{p_1p_2}\\&\times\operatorname{tr}(V_{11}+S_{11}+U_1U_1')^{-1}V_{12}(V_{22}+S_{22}+U_2U_2')^{-1}V_{21}\\&+b^2\operatorname{tr}(V_{11}+S_{11}+U_1U_1')^{-1}U_1U_2'(V_{22}+S_{22}+U_1U_2')^{-1}U_2U_1'\\&+n+b^2\\&-\frac{n+m_1+1}{p_1}\operatorname{tr}(V_{11}+S_{11}+U_1U_1')^{-1}(V_{11}+b^2U_1U_1')\\&-\frac{n+m_2+1}{p_2}\operatorname{tr}(V_{22}+S_{22}+U_2U_2')^{-1}(V_{22}+b^2U_2U_2').\end{aligned} \tag{3.5}$$

Rejecting $H_0:\Sigma_{12}=0$ for large values of U_2 is an LBI test.

Case (iii): Again consider the data as given in (2.3) but assume that the mean of X_i is unrelated to the mean of Y, $i=1,2$. In this case, the LBI test of $H_0:\Sigma_{12}=0$ rejects for large values of U_0 given by (3.3).

Since the marginal distribution of $\operatorname{tr}(V_{ii}+S_{ii})^{-1}V_{ii}$ does not depend on any parameter $(i=1,2)$ in (3.3), one may ignore the two ancillary statistics $\operatorname{tr}(V_{ii}+S_{ii})^{-1}V_{ii}$ $(i=1,2)$ in (3.3) and reject H_0 for large values of

$$\tilde{U}_0=\operatorname{tr}(V_{11}+S_{11})^{-1}V_{12}(V_{22}+S_{22})^{-1}V_{21}.$$

The exact null distribution of U_0 and $\tilde{U}_0$ is difficult to derive but the asymptotic null distributions up to order n^{-1} are derived in Section 5. For the limiting distributions, letting n and m_1 tend to

∞ with $\frac{m_i}{n} \to \beta_i$, $i=1,2$, it is not too hard to show that nU_0 converges in distribution to a random variable with a scaled chi-square distribution. In particular,

$$n\tilde{U}_0 \xrightarrow{d} (1+\beta_1)(1+\beta_2)\chi^2_{p_1 p_2}$$

when n is large.

3.2. *LRT*. Without loss of generality, the model (3.1) is considered. When m_1 and m_2 are positive, it seems difficult to derive explicitly the LRT (likelihood ratio test) for testing $\Sigma_{12}=0$. However, in case (i) where $m_2=0$, the LRT is derived as follows.

Proposition 3.1. When $m_2=0$, the LRT for testing $\Sigma_{12}=0$ rejects for small values of

$$U_4 = |I - V_{11}^{-1} V_{12} V_{22}^{-1} V_{21}| \tag{3.6}$$

Proof. As is well known, the distribution of V is decomposed into

$$\begin{aligned} &w \equiv V_{11}^{-1/2} V_{12} \text{ given } V_{11} \sim N(V_{11}^{1/2}\beta,\ I \otimes \Sigma_{22.1}) \quad \text{with } \beta = \Sigma_{11}^{-1}\Sigma_{12} \\ &V_{11} \sim W_{p_1}(\Sigma_{11}, n) \quad \text{and} \\ &V_{22.1} = V_{22} - V_{21} V_{11}^{-1} V_{12} \sim W_{p_2}(\Sigma_{22.1},\ n-p_1) \end{aligned}$$

where (w, V_{11}) is independent of $V_{22.1}$. Hence the joint density of $(w, V_{11}, V_{22.1}, S_{11})$ is of the form

$$h_1(w \mid V_{11}, \beta, \Sigma_{22.1}) h_2(V_{22.1} \mid \Sigma_{22.1}) h_3(V_{11} \mid \Sigma_{11}) h_4(S_{11} \mid \Sigma_{11}) = L(\beta, \Sigma_{22.1}, \Sigma_{11}).$$

Hence the maximum likelihood is

$$\begin{aligned} \sup L(\beta, \Sigma_{22.1}, \Sigma_{11}) = &\sup_{(\beta, \Sigma_{22.1})} h_1(w \mid V_{11}, \beta, \Sigma_{22.1}) h_2(V_{22.1} \mid \Sigma_{22.1}) \\ &\sup_{\Sigma_{11}} h_3(V_{11} \mid \Sigma_{11}) h_4(S_{11} \mid \Sigma_{11}), \end{aligned}$$

which implies that the LRT is independent of S_{11} because

$$\frac{\sup L(\beta, \Sigma_{22.1}, \Sigma_{11})}{\sup L(0, \Sigma_{22.1}, \Sigma_{11})} = \frac{\sup h_1(w \mid V_{11}, \beta, \Sigma_{22.1}) h_2(V_{22.1} \mid \Sigma_{22.1})}{\sup h_1(w \mid V_{11}, 0, \Sigma_{22.1}) h_2(V_{22.1} \mid \Sigma_{22.1})}.$$

Further this is nothing but the LRT for independence based on the complete part of data in (3.1), yielding the result.

It is rather surprising that the LRT ignores the data S_{11} in (3.1). Of course, when $m_1 = m_2 = 0$, rejecting for small values of U_4 gives the LRT. However, when m_1 is very large, the value of Σ_{11} is essentially known but the likelihood ratio criterion ignores this information. Indeed, if Σ_{11} is known, the LRT based on V alone is also that given in Proposition 3.1 (see Section 3 of Chapter 1).

3.3. *The Case* $p_1=1$ and $m_2=0$ is here considered. In this case

$$T_1 = V_{11}^{-1} V_{12} V_{22.1}^{-1} V_{21} = w V_{22.1}^{-1} w' \quad \text{and} \quad T_2 = S_{11}/V_{11}$$

are shown to form a maximal invariant and a maximal invariant parameter is $\lambda = \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22.1}^{-1}\Sigma_{21} = \Sigma_{11}\beta\Sigma_{22.1}^{-1}\beta'$, where $w = V_{11}^{-1/2} V_{12}$ and $\beta = \Sigma_{11}^{-1}\Sigma_{12}$. Hence the problem of testing $\Sigma_{12}=0$ or $\beta=0$ is similar in structure to the special case of the GMANOVA problem treated in Section 7 of Chapter 3. In fact, first of all, even though the maximal invariant parameter is one dimmensional, a UMPI test does not exist and the LRT which rejects for large values of T_1 is different from the LBI test whose critical region is, by (3.3) with $p_1=1$ and $m_2=0$, given by

$$\begin{aligned} b &\equiv n(V_{11}+S_{11})^{-1} V_{12} V_{22}^{-1} V_{21} - p_2(V_{11}+S_{11})^{-1} V_{11} \\ &= (1+T_2)^{-1}[nT_1(1+T_1)^{-1} - p_2] > c. \end{aligned}$$

This is of a similar form as the LBI test in the GMANOVA problem. Second, as in the case of the GMANOVA problem terated in Section 7 of Chapter 3, the maximal invariant consists of two dimmensional statistics T_1 and T_2 in which T_2 is ancillary, and the LRT ignores the ancillary statistic. Third, in the class of conditional level α invariant tests given V_{11}, the LRT is UMPI. For from (3.7), it follows that

$$(n-p_2)T_1/p_2 \text{ given } V_{11} \sim F(p_2, n-p_2 : V_{11}\lambda)$$
$$T_2 \text{ given } V_{11} \sim (\Sigma_{11}/V_{11})\chi^2(m_1)$$

and conditional on V_{11}, T_1 and $S_{22}=V_{11}T_2$ are independent, where $F(a,b:\lambda)$ denotes the noncentral F-distribution with noncentral parameter λ and degrees of freedom (a,b). Hence from the monotone likelihood ratio property of $F(p_2,\ n-p_2 : V_{11}\lambda)$ the conditional UMPI property of the LRT follows for each given V_{11}.

In Marden (1981), the admissibility of invariant tests in this case is also considered. Let $f_\lambda\ (T_1,T_2)$ be the density of (T_1,T_2) under λ and let

$$R_\lambda(T_1,T_2)=(1+\lambda)^{(n+m_1)/2}f_\lambda(T_1,T_2)/f_0(T_1,T_2) \quad \text{and}$$
$$d(T_1,T_2)=\int_0^1\left[\frac{R_\lambda-1}{\lambda}\right]\pi^0(d\lambda)+\int_1^\infty R_\lambda\pi^1(d\lambda)$$

where π^0 and π^1 are finite measures on $[0,1]$ and $[1,\infty]$ and $(R_\lambda-1)/\lambda$ is continuously extended to the null hypothesis $\lambda=\beta'\Sigma_{22\cdot1}^{-1}\beta=0$ by

$$(R_\lambda-1)/\lambda|_{\lambda=0}\equiv\text{LBI}.$$

Define $\mathscr{B}$ to be the class of tests ϕ with critical region of the form $d(T_1,T_2)>c$ for some π^0 and π^1. Then as in the case of the GMANOVA problem, the following proposition holds.

Proposition 3.2. (Marden (1981)) The class of tests $\mathscr{B}$ is minimal complete for testing $\lambda=0$ versus $\lambda>0$ based on (T_1,T_2).

Using this result, it is shown that the LRT of size α is inadmissible if $\alpha>\alpha^*$ and it is admissible if $0<\alpha\leq\alpha^*$ where α^* is given by $p((n-p_2)T_1/p_2>1)=\alpha^*$. While the LBI test is admissible because it is a unique best test aroud $\lambda=0$.

Finally, consider the special case of $p_1=p_2=1$ so $p=2$ and the data

is given by (3.1). A minimal sufficient statistic is (V,S_{11},S_{22}) where

$$V=(V_{jk})=Z'Z \quad \text{and} \quad S_{ii}=Z_i'Z_i\ (i=1,2).$$

In this case, the problem is to test $\rho=0$ where ρ is the bivariate correlation coefficient. The testing problem is invariant under scale changes and a maximal invariant statistic is (t_1,t_2,t_3) where

$$t_1=\frac{V_{12}^2}{V_{11}V_{22}},\quad t_2=\frac{S_{11}}{V_{11}},\quad t_3=\frac{S_{22}}{V_{22}}.$$

When $m_2=0$, the LRT rejects for large values of t_1 while the LBI test involves both t_1 and t_2. But, when both m_1 and m_2 are positive there is a complication. The statistics t_2 and t_3 are marginally ancillary but (t_2,t_3) is not an ancillary statistic. It is not clear how to condition in this case, but rejecting for large t_1 will not be appropriate.

3.4. *Proof of Theorem* 3.1. The proof of Theorem 3.1 is similar to the proof of Theorem 8.1 in Chapter 3. However, for comparison, we use Andersson's (1983) representation theorem concerning the density of a maximal invariant rather than Wijsman's (1967) theorem. Let $\mathscr{X}=\mathscr{X}_0\times\mathscr{X}_1\times\mathscr{X}_2$ denote the sample space of the data given by (3.1) where $\mathscr{X}_i$ is the linear space of $n_i\times p_i$ real matrices $(i=0,1,2)$. Here the notation

$$n_0=n,\quad n_1=m_1,\ n_2=m_2 \quad \text{and} \quad \Sigma_{00}=\Sigma$$

is used and we shall write $x=(x_0,x_1,x_2)\in\mathscr{X}$ for $(Z,Z_1,Z_2)\in\mathscr{X}$. The Lebesgue measure on $\mathscr{X}$ will be denoted by $dx=dx_0dx_1dx_2$, and from (3.1), the density of x is given by

$$(3.7)\qquad f(x|\Sigma)=\prod_{i=0}^2(\sqrt{2\pi})^{-n_i/2}|\Sigma_{ii}|^{-n_i/2}\exp(-\frac{1}{2}\mathrm{tr}x_i\Sigma_{ii}^{-1}x_i').$$

The group $\mathscr{G}$ defined above acts on $x\in\mathscr{X}$ by $gx=(x_0A',\ x_1A_1',\ x_2A_2')$ where $g=A=\begin{bmatrix}A_1 & 0\\ 0 & A_2\end{bmatrix}$ with $A_i\in\mathscr{G}l(p_i)$. A left invariant measure on $\mathscr{G}$ is

$$(3.8)\quad \nu(dg)=\nu_1(dA_1)\nu_2(dA_2) \text{ with } \nu_i(dA_i)=|A_iA_i'|^{-p_i/2}dA_i \ (i=1,2).$$

Further, dx is relatively invariant with multiplier

$$\chi_0(g)=\prod_{i=1}^{2}|A_iA_i'|^{(n_0+n_i)/2}.$$

To establish Theorem 3.1, we will use the following well known argument. For any invariant test function ϕ, the power function of ϕ at a maximal invariant parameter point δ is

$$(3.9)\qquad \pi(\phi,\delta)=\int\phi dP_\delta=\int\phi\left(\frac{dP_\delta}{dP_0}\right)dP_0$$

where P_δ is the probability measure of a maximal invariant. Under the conditions given in Theorem 3.1 of Chapter 2, the ratio dP_δ/dP_0 is given by

$$(3.10)\qquad r_\delta(x)=\frac{\int_{\mathscr{G}} f(gx\,|\,\Sigma(\delta))\chi_0(g)\nu_l(dg)}{\int_{\mathscr{G}} f(gx\,|\,\Sigma(0))\chi_0(g)\nu_l(dg)}$$

where

$$\Sigma(\delta)\equiv\begin{pmatrix} I_{p_1} & \Delta \\ \Delta' & I_{p_2}\end{pmatrix}$$

and Δ is $p_1\times p_2$ with $\Delta_{ii}=\delta_{ii}$ for $i=1,\dots,p_2$, and $\Delta_{ij}=0$ for $i\neq j$. Without loss of generality we have taken $p_1\geq p_2$. Before calculating r_δ, the assumptions in Theorem 3.1 of Chapter 2 required for the representation (3.10) need to be checked. That $\mathscr{X}$ and $\mathscr{G}$ are locally compact sigma compact spaces is clear and obviously $\mathscr{G}$ acts topologically on $\mathscr{X}$. Thus, it must be shown that the action of $\mathscr{G}$ on $\mathscr{X}$ is proper. The next paragraph is devoted to a discussion of this point.

First, the action of $\mathscr{G}$ on $\mathscr{X}$ is not proper. To see this, let $K(g,x)=(gx,x)$ be a map from $\mathscr{G}\times\mathscr{X}$ into $\mathscr{X}\times\mathscr{X}$ and take $C=\{0,0\}\in\mathscr{X}\times\mathscr{X}$ so

$$K^{-1}(C)=\mathscr{G}\times\{0\}$$

which is not compact in $\mathscr{G}\times\mathscr{X}$. However, the work of Wijsman (1967) and the discussion in Farrell (1976) suggests that a set of Lebesgue measure zero needs to be removed from $\mathscr{X}$. For $x=(x_0,x_1,x_2)$, write

$$y_1=\begin{pmatrix} x_0^{(1)} \\ x_1\end{pmatrix},\quad y_2=\begin{pmatrix} x_0^{(2)} \\ x_2\end{pmatrix}$$

where y_i is $(n_0+n_i)\times p_i$ $(i=1,2)$ and x_0 has been partitioned into $x_0^{(i)}: n_0\times p_i$ $(i=1,2)$. Now, the action of $\mathscr{G}$ on $\mathscr{X}$ becomes

$$g(y_1,y_2)\equiv(y_1A_1',\ y_2A_2')$$

where

$$g=\begin{pmatrix} A_1 & 0 \\ 0 & A_2\end{pmatrix}\in\mathscr{G}.$$

Let $\mathscr{Y}_i$ be the real linear space of all $(n_0+n_i)\times p_i$ matrices so $\mathscr{G}=\mathscr{G}l(p_1)\times\mathscr{G}l(p_2)$ acts on $\mathscr{Y}_1\times\mathscr{Y}_2\equiv\mathscr{Y}$ coordinatewise. Furthermore, Lebesgue measure dx on $\mathscr{X}$ corresponds to Lebesgue measure dy on $\mathscr{Y}$. Since the product group $\mathscr{G}l(p_1)\times\mathscr{G}l(p_2)$ acts coordinatewise on the product space $\mathscr{Y}_1\times\mathscr{Y}_2$ and $dy=dy_1\times dy_2$, to discuss the issues surrounding the action of $\mathscr{G}$ on $\mathscr{Y}$, it suffices to discuss the action of $\mathscr{G}l(p_1)$ on $\mathscr{Y}_1$ (see Lemma 3.2 in Chapter 2). Let $\mathscr{N}_1$ be the elements of $\mathscr{Y}_1$ which have rank less than p_1. Since $n_0+n_1\geq p_1$, $\mathscr{N}_1$ has Lebesque measure zero and the space $\mathscr{Y}_1^*\equiv\mathscr{Y}_1-\mathscr{N}_1$ is an open set in $\mathscr{Y}_1$ which is acted on by $\mathscr{G}l(p_1)$. The claim is that $\mathscr{G}l(p_1)$ acts properly on $\mathscr{Y}_1^*$. To see this, let $C\subseteq\mathscr{Y}_1^*\times\mathscr{Y}_1^*$ be compact. It is clear that we can find a compact set $B_1\subseteq\mathscr{Y}_1^*$ such that $C\subseteq B_1\times B_1$. Since the mapping K defined above is continuous, it follows that if $K^{-1}(B_1\times B_1)$ is compact then $K^{-1}(C)$ is compact. However,

$$K^{-1}(B_1\times B_1)=\{(g,y)\,|\,(g(y),y)\in B_1\times B_1\}.$$

For $u \in \mathcal{Y}_1$, let

$$||u||^2 = \mathrm{tr} u'u$$

so $||\cdot||$ is a norm on $\mathcal{Y}_1$. Also for $h \in R^{p_1^2}$, let

$$||h||_1^2 = \mathrm{tr} hh'$$

so $||\cdot||_1^2$ is a norm on $R^{p_1^2}$. Since $K^{-1}(B_1 \times B_1)$ is a closed subset of $\mathcal{Gl}(p_1) \times \mathcal{Y}_1^* \subseteq R^{p_1^2} \times \mathcal{Y}_1$, to show $K^{-1}(B_1 \times B_1)$ is compact, it suffices to show that $||g||_1^2 + ||y||^2$ remains bounded for $(g, y) \in K^{-1}(B_1 \times B_1)$. The compactness of $B_1 \subseteq \mathcal{Y}_1^*$ implies that there are two constants a_1 and b_1 such that

$$||y||^2 \leq a_1 < +\infty, \quad y \in B_1$$

and

$$\lambda_p(y'y) \geq b_1 > 0, \quad y \in B_1$$

where $\lambda_p(y'y)$ is the smallest eigenvalue of $y'y$. For $(g, y) \in K^{-1}(B_1 \times B_1)$,

$$\begin{aligned} a_1 &\geq ||g(y)||^2 = \mathrm{tr}[(yg')'yg'] \\ &= \mathrm{tr} y'yg'g = \mathrm{tr}(y'y - b_1 I_{p_1})g'g + b_1 \mathrm{tr} g'g \geq b_1 ||g||_1^2. \end{aligned}$$

The inequality follows since for $y \in B_1$, $y'y - b_1 I_p$ is nonnegative definite so $\mathrm{tr}(y'y - b_1 I_{p_1})g'g \geq 0$. Therefore, for $(g, y) \in K^{-1}(B_1 \times B_1)$,

$$||g||_1^2 + ||y||^2 \leq a_1 + a_1 b_1^{-1}$$

so $K^{-1}(B_1 \times B_1)$ is compact. Hence the action of $\mathcal{Gl}(p_1)$ on $\mathcal{Y}_1^*$ is proper. Now, let $\mathcal{X}^* \subseteq \mathcal{X}$ be defined by

$$\mathcal{X}^* = \left\{(x_0, x_1, x_0) \,\middle|\, \begin{pmatrix} x_0^{(i)} \\ x_i \end{pmatrix} \text{has rank } p_i,\ i = 1, 2\right\}.$$

It follows immediately that $\mathcal{X} - \mathcal{X}^*$ has Lebesgue measure zero and $\mathcal{G}$ acts properly on $\mathcal{X}^*$.

We now proceed with the evaluation of $r_\delta(x)$ given in (3.10). For $x = (x_0, x_1, x_2) \in \mathcal{X}^*$, let

$$(3.11) \qquad x_0'x_0 \equiv V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}, \quad x_i'x_i \equiv S_{ii},\ i = 1, 2.$$

and

$$(3.12) \qquad T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}$$

$$\equiv \begin{pmatrix} (V_{11}+S_{11})^{-1/2} & 0 \\ 0 & (V_{22}+S_{22})^{-1/2} \end{pmatrix} V \begin{pmatrix} (V_{11}+S_{11})^{-1/2} & 0 \\ 0 & (V_{22}+S_{22})^{-1/2} \end{pmatrix}$$

where the notation of 3.1 is being used. Note that $0 < T < I_p$ in the sense of positive definiteness. Let E_i denote expectation with respect to the distribution on $\mathcal{Gl}(p_i)$ given by

$$(3.13) \qquad Q_i(dA) = c(\alpha_i, p_i)|A_i A_i'|^{\alpha_i/2} \exp[-\tfrac{1}{2} \mathrm{tr} A_i A_i'] \nu_i(dA_i)$$

where $\alpha_1 = n + m_1$, $\alpha_2 = n + m_2$, $c(\alpha_i, p_i)$ is a normalization constant and $\nu_i(dA_i) = dA_i / |A_i'A_i|^{p_i/2}$. A bit of algebra and a change of variable show that

$$(3.14) \qquad r_\delta(x) = |\Sigma(\delta)|^{-n/2} E_1 E_2 \{\exp[-\tfrac{1}{2} \mathrm{tr} TA'(\Sigma^{-1}(\delta) - I_p)A]\}$$

where

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad A_i \in \mathcal{Gl}(p_i) \quad (i = 1, 2),$$

and E_i is expectation on A $(i = 1, 2)$. Define $\Psi : p \times p$ by

$$(3.15) \qquad \Psi \equiv \Sigma^{-1}(\delta) - I_p = \begin{pmatrix} (I - \Delta\Delta')^{-1} - I & -(I - \Delta\Delta')^{-1}\Delta \\ -(I - \Delta'\Delta)^{-1}\Delta' & (I - \Delta'\Delta)^{-1} - I \end{pmatrix} = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{pmatrix}$$

where Ψ_{ij} is $p_i \times p_j$ $(i = 1, 2)$. Then we have

$$(3.16)\qquad \mathrm{tr}TA'(\Sigma^{-1}(\delta)-I_p)A = \mathrm{tr}T_{11}A_1'\Psi_{11}A_1 + 2\mathrm{tr}T_{12}A_2'\Psi_{21}A_1 + \mathrm{tr}T_{22}A_2'\Psi_{22}A_2.$$

We now make the following claim: For δ small, $\tau=\sum t^2\delta_i^2$, and all T satisfying $0<T<I_p$ (in the sense of positive definiteness),

$$(3.17)\qquad \begin{cases} \exp[-\frac{1}{2}\mathrm{tr}T_{11}A_1'\Psi_{11}A_1]=1-\frac{1}{2}\mathrm{tr}T_{11}A_1'\Delta\Delta'A_1+R_1(T_{11},A_1,\Delta) \\ \exp[-\frac{1}{2}\mathrm{tr}T_{22}A_2'\Psi_{22}A_2]=1-\frac{1}{2}\mathrm{tr}T_{22}A_2'\Delta'\Delta A_2+R_2(T_{22},A_2,\Delta) \\ \exp[-\mathrm{tr}T_{12}A_2'\Psi_{21}A_1]=1+\mathrm{tr}T_{12}A_2'\Delta'A_1+\frac{1}{2}(\mathrm{tr}T_{12}A_2'\Delta'A_1)^2 \\ \qquad\qquad +R_3(T_{12},A_1,A_2,\Delta) \end{cases}$$

where the error terms R_1, R_2 and R_3 satisfy

$$(3.18)\qquad \begin{cases} |R_1(T_{11},A_1,\Delta)|\le H_1(A_1)\varphi_1(\tau) \\ |R_2(T_{22},A_2,\Delta)|\le H_2(A_2)\varphi_2(\tau) \\ |R_3(T_{12},A_1,A_2,\Delta)|\le H_3(A_1)H_4(A_2)\varphi_3(\tau) \end{cases}$$

Further, the inequalities in (3.18) hold for all $T, 0<T<I_p$, the functions H_i are integrable (A_1,A_2) and

$$\lim_{\tau\to 0}\varphi_i(\tau)/\tau=0,\quad i=1,2,3.$$

The arguments leading to (3.17) and (3.18) are similar to those in Section 8 of Chapter 3 and are omitted. The following identities are used in the evaluation of (3.17) :

$$(3.19)\qquad \begin{cases} E_1\mathrm{tr}T_{11}A_1'\Delta\Delta'A_1 = \dfrac{n+m_1}{p_1}(\mathrm{tr}T_{11}T_{11}')\tau \\ E_2\mathrm{tr}T_{22}'A_2\Delta'\Delta A_2 = \dfrac{n+m_1}{p_2}(\mathrm{tr}T_{22}T_{22}')\tau \end{cases}$$

$$(3.20)\qquad \begin{cases} E_1E_2\mathrm{tr}T_{12}A_2'\Delta'A_1 = 0 \\ E_1E_2(\mathrm{tr}T_{12}A_2'\Delta'A_1)^2 = \dfrac{n+m_1}{p_1}\,\dfrac{n+m_2}{p_2}(\mathrm{tr}T_{12}T_{12}')\tau \end{cases}$$

Note that $|\Sigma(\delta)|^{-n/2}=1+\frac{n}{2}\tau+o(\delta)$ where $\lim_{\delta\to 0} o(\delta)/\tau=0$. Substituting this and the expressions in (3.19) into (3.14) leads to the expression:

$$(3.21)\qquad r_\delta(x) = 1+\frac{1}{2}U_0\tau+o(T,\delta)$$

where U_0 is defined in (3.3). The remainder term is uniformly bounded in $T, 0<T<I_p$, and satisfies

$$\lim_{\delta\to 0}\sup_T \frac{o(T,\delta)}{\tau}=0.$$

The identities in (3.19) and the results expressed in (3.18) are used to establish (3.21).

Now, let ϕ be any level α invariant test of $H_0:\delta=0$ versus $H_1:\delta\neq 0$. Substituting (3.21) into (3.9) yields

$$\pi(\phi,\delta)=\int\phi\left(\frac{dP_\delta}{dP_0}\right)dP_0=\int\phi[1+\frac{1}{2}U_0\tau+o(T,\delta)]dP_0$$
$$=\alpha+\frac{1}{2}E_0(\phi U_0)\tau+o(\phi,\delta)$$

where the remainder term satisfies

$$\lim_{\delta\to 0}\sup_\phi \frac{o(\phi,\delta)}{\tau}=0.$$

This proves Theorem 3.1.

4. Local Minimaxity of the LBI Test.

4.1. *The local minimaxity* of the LBI test derived in Section 3 is considered in this section. It was first proved by Chou and Lo (1983) where using Stein's representation theorem for the probability ratio of a maximal invariant and taking Schwartz's approach (1967), the sufficient condition for local minimaxity provided by Giri and Kiefer (1964) was verified. Here using Kariya's approach (1978), it is verified. The definition of local minimaxity and the sufficient condition are given in Section 9 of Chapte 3. We follow the notation in Section 3. Let us denote the LBI test statistic by

$$(4.1)\qquad U_0 \equiv \frac{(n+m_1)(n+m_2)}{p_1 p_2}\mathrm{tr}(V_{11}+S_{11})^{-1}V_{12}(V_{22}+S_{22})^{-1}V_{21} + n - \sum_{i=1}^{2}\left(\frac{n+m_i}{p_i}\right)\mathrm{tr}(V_{ii}+S_{ii})^{-1}V_{ii}$$

and let

$$(4.2)\qquad \lambda \equiv \lambda(\Sigma) = \mathrm{tr}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \sum_{i=1}^{q}\delta_i^2 \quad (q=\min(p_1,p_2)).$$

Theorem 4.1. The LBI test ϕ_0 with critical region $U_0>c$ is locally minimax with respect to the contour in (4.2) as $\lambda\to 0$ under the model (3.1). That is,

$$\lim_{\lambda\to 0}\frac{\inf_{\Sigma}\pi(\phi_0,\Sigma)-\alpha}{\sup_{\phi\in Q_\alpha}\inf_{\Sigma}\pi(\phi,\Sigma)-\alpha}=1,$$

where $\pi(\phi,\Sigma)=E[\phi|\Sigma]$ is the power function of ϕ and Q_α is the class of size α test.

It is noted that the class of tests Q_α in which we claim the local minimaxity of ϕ_0 is not restricted to the class of invariant tests of size α. It is the class of all tests of size α for testing $\Sigma_{12}=0$ in the model(3.1). However, this theorem shows the local minimaxity of ϕ_0 simply on the contour

$$C_\lambda=\{\Sigma\in\mathcal{S}_+(p)\,|\,\lambda(\Sigma)=\lambda\}.$$

In the MANOVA model, Schwartz (1967) defined a local family of contours and showed the local minimaxity of the Pillai test for independence which is LBI and defined by

$$U_3=\mathrm{tr}\,V_{11}^{-1}V_{12}V_{22}^{-1}V_{21}>c.$$

It is also remarked that as in the case of the GMANOVA problem, if an invariant test is to be locally minimax, it must be the one which is LBI.

4.2. *The proof of Theorem* 4.1 is similar to the proof of Theorem 9.1 in Chapter 3 and so outlined here. By the Hunt-Stein theorem, it suffices to show that ϕ_0 is locally minimax in the class of tests invariant under the group

$$(4.3)\qquad G=\left\{g\,|\,g=\begin{pmatrix}B_1 & 0\\ 0 & B_2\end{pmatrix},\ B_i\in\mathcal{GT}(p_i)\ (i=1,2)\right\},$$

where $\mathcal{GT}(p_i)$ denotes the group of $p_i\times p_i$ nonsingular lower triangular matrices. Let a left invariant measure on B_i be $\mu_i(dB_i)$ and let

$$(4.4)\qquad \mu(dg)=\mu_1(dB_1)\mu_2(dB_2)$$

and G acts on $(\boldsymbol{Z},\boldsymbol{Z}_1,\boldsymbol{Z}_2)$ and Σ by $g(\boldsymbol{Z},\boldsymbol{Z}_1,\boldsymbol{Z}_2)=(\boldsymbol{Z}g,\boldsymbol{Z}_1B_1,\boldsymbol{Z}_2B_2)$ and $g(\Sigma)=g'\Sigma g$ respectively. Then replacing $\Sigma(\delta)$ by $\Sigma(\xi)$ and r_δ by $r_{\Sigma(\xi)}$, (3.10) holds with $\chi_0(g)=\prod_{i=1}^{2}|B_i'B_i|^{(n+m_i)/2}$ and

$$(4.5)\qquad \Sigma=\begin{pmatrix}I, & \xi\\ \xi', & I\end{pmatrix}\ \text{with}\ \xi=\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}\quad(\Sigma_{ii}^{-1/2}\in\mathcal{GT}(p_i)\,:\,i=1,2).$$

where the condition for Wijsman's or Andersson's representation of the probability ratio is checked similarly to the case of Section 9 in Chapter 3 or the case of Section 3. Hence defining T by (3.12) and $Q_i(dA_i)$ by (3.13) with $\nu_i(dA_i)$ replaced by $\mu(dg)$, we obtain the expression like (3.14);

$$(4.6)\qquad r_{\Sigma(\xi)}=|\Sigma(\xi)|^{-n/2}E_1E_2\,\{\exp[-\tfrac{1}{2}\mathrm{tr}TB(\Sigma(\xi)^{-1}-I)B']\}$$

where E_i is expectation on B_i $(i=1,2)$. Further defining Ψ by (3.15) with $\Sigma(\delta)$ and Δ replaced by $\Sigma(\xi)$ and ξ respectively, (3.16), (3.17) and (3.18) hold with A_i replaced by B_i's, where τ is replaced by λ in (4.2). This implies that the probability ratio is expressed as

$$R_{\Sigma(\xi)}=|\Sigma(\xi)|^{-n/2}\{1-\frac{1}{2}\textstyle\sum_{i=1}^{2}E_i[\operatorname{tr}T_{ii}B_i\Psi_{ii}B_i'] \tag{4.7}$$
$$-E_1E_2[\operatorname{tr}T_{12}B_2\Psi_{21}B_1'-\frac{1}{2}(\operatorname{tr}T_{12}B_2\Psi_{21}B_1')^2]\}+o(\lambda)$$

where $o(\lambda)$ is uniform in T (note $0<T<I$).

Under this preparation, we first verify Assumption 3 of Lemma 9.1 in Chapter 3. Noting $|\Sigma(\xi)|^{-n/2}=1+\frac{n}{2}\lambda+o(\lambda)$, it suffices to show that the inside of { } in (4.7) is $1+\frac{1}{2}[U_5-\frac{n}{2}]\lambda+o(\lambda)$ for *some* ξ. In Section 3, since $A_i\in\mathcal{G}l(p_i)$ was chosen for $B_i\in\mathcal{G}\mathcal{T}(p_i)$, this was true for all A. We show that there exists a probability measure η on ξ such that

$$E_\eta E_i[\operatorname{tr}T_{ii}B_i\Psi_{ii}B_i']=\frac{n+m_i}{p_i}\lambda\operatorname{tr}T_{ii}'T_{ii}+o(\lambda)\qquad(i=1,2) \tag{4.8}$$

where $o(\lambda)$ is uniform. Define a map $h:\mathcal{O}(p_1)\times\mathcal{E}\times\mathcal{O}(p_2)\to R^{p_1p_2}$ by

$$\xi\equiv h(P,D,Q)=PDQ \tag{4.9}$$

where $\mathcal{E}$ is the set of $p_1\times p_2$ diagonal matrices and conditional on D, let

$$\eta(d\xi\,|\,D)=\nu(dP)\tau(dQ) \tag{4.10}$$

when ν and τ are respectively the invariant probability measures on $\mathcal{O}(p_1)$ and $\mathcal{O}(p_2)$. Then $\lambda=\operatorname{tr}\xi\xi'=\operatorname{tr}DD'$ and using $\Psi_{11}=(I-\xi\xi')^{-1}-I=\xi\xi'+o(\lambda)$ and $\Psi_{22}=(I-\xi'\xi)^{-1}-I=\xi'\xi+o(\lambda)$,

$$E_\eta[\operatorname{tr}T_{ii}B_i\Psi_{ii}B_i']=\frac{\lambda}{p_i}\operatorname{tr}T_{ii}'T_{ii}B_iB_i'+o(\lambda)$$

where Lemma 8.2 in Chapter 3, is used, implying (4.8) from $B_iB_i'\sim W(I,\ n+m_i)$. Similarly using $\Psi_{21}=(I-\xi'\xi)^{-1}\xi'=\xi'+o(\lambda)$,

$$E_\eta E_1E_2[\operatorname{tr}T_{12}B_2\Psi_{21}B_1'-\frac{1}{2}(\operatorname{tr}T_{12}B_2\Psi_{21}B_1)^2]$$
$$=-\frac{(n+m_1)(n+m_2)}{2p_1p_2}\lambda\operatorname{tr}T_{12}'T_{12}+o(\lambda)$$

Hence for any D such that $\operatorname{tr}DD'=\lambda$, the Assumption 3 is verified.

The verification of Assumptions 1 and 2 of Lemma 9.1 in Chapter 3 are similar to that in Theorem 9.1 there and so omitted here.

5. Asymptotic Null Distributions

5.1. *The LRT and Pillai's test.* In 5.1, for convenience, we consider what we called in Section 3 the case (0) where the model is given by (3.1). In the case (0) with $m_1\neq 0$ and $m_2=0$, which we referred to the case (1) in Section 3, there is additional data only on the second p_1 coordinates and the LRT for independence

$$U_4=-2(n-\frac{1}{2}(p+1))|I-V_{11}^{-1}V_{12}V_{22}^{-1}V_{21}|>c \tag{5.1}$$

completely ignores the additional data and it is the same as the LRT for independence in the case of $m_1=0$ and $m_2=0$ where no additional data is available at all, which we may call the complete case. Here the suffix 4 is used correspondingly to the GMANOVA problem. In this complete case, as is shown in Schwartz (1967a), the Pillai test

$$U_3=n\operatorname{tr}V_{11}^{-1}V_{12}V_{22}^{-1}V_{21}>c \tag{5.2}$$

is the LBI. Hence even in the case (1), we may use this test and such related tests as Lawley-Hotelling test etc, based on the fact that the LRT in the complete case is still the LRT for the case (1). The asymptotic null and nonnull distributions of these tests up to order n^{-2} are derived by Sugiura and Fujikoshi (1969) and Fujikoshi (1970). Here the results are introduced up to order n^{-1}.

(1) The LRT U_4:

$$P(U_4\leq x)=G_f(x)+o([n-\frac{1}{2}(p+1)]^{-1}) \tag{5.3}$$

where $f=p_1p_2$ and G_f is the distribution function of χ^2-distribution with degrees of freedom f, i.e., $\chi^2(f)$.

(2) The Pillai test U_3 :

$$P(U_3 \le x) = G_f(x) + \frac{f\gamma}{4n}[-G_f(x) + 2G_{f+2}(x) - G_{f+4}(x)] + o(n^{-1}) \tag{5.4}$$

where $\gamma = p_1 + p_2 + 1 = p + 1$.

It is noted that in the case (0) where $m_1 \neq 0$ and $m_2 \neq 0$, the LRT is not explicitly derived. Hence in the case (0), it is less advisable to use these tests.

5.2. *The LBI test and a related test.* First we shall treat the case (0), where the model is given by (3.1) with $m_1 \ge 0$ and $m_2 \ge 0$, and derive the asymptotic null distribution up to order n^{-1} of the LBI test statistic defined by

$$\begin{aligned} U_5 = {} & n\operatorname{tr}(V_{11}+S_{11})^{-1}V_{12}(V_{22}+S_{22})^{-1}V_{21} \\ & -\textstyle\sum_{i=1}^2\Big[\Big(1+\frac{m_i}{n}\Big)\Big(\frac{p_1p_2}{p_i}\Big)\Big/\Big(1+\frac{m_1}{n}\Big)\Big(1+\frac{m_2}{n}\Big)\Big] \\ & \times \operatorname{tr}(V_{ii}+S_{ii})^{-1}V_{ii} \end{aligned} \tag{5.5}$$

and the asymptotic null distribution up to order n^{-1} of the test statistic defined by

$$U_6 = n\operatorname{tr}(V_{11}+S_{11})^{-1}V_{12}(V_{22}+S_{22})^{-1}V_{21}. \tag{5.6}$$

Concerning the asymptotic orders of m_1 and m_2, the following cases are distinguished: (a) $m_i = O(1)$ $(i=1,2)$, (b) $m_1 = O(1)$, $m_2 = O(n)$ and (c) $m_i = O(n)$ $(i=1,2)$.

Case (a) $m_i = O(1)$ or $\lim_{n\to\infty}(m_i/n) = 0$ $(i=1,2)$: Define

$$U_5^a = U_5 + 2p_1p_2(n - m_1 - m_2)/n - (p_2m_1 + p_1m_2)/n \tag{5.7}$$

Then

$$\begin{aligned} P(U_5^a \le x) = {} & G_f(x) + \frac{f\gamma}{4n}[-G_f(x) + 2G_{f+2}(x) - G_{f+4}(x)] \\ & + \frac{(m_1+m_2)}{2n}[G_f(x) - G_{f+2}(x)] + o(n^{-1}). \end{aligned} \tag{5.8}$$

Further for U_6 in (5.6),

$$P(U_6 \le x) = P(U_5^a \le x) + o(n^{-1})$$

Case (b) $m_1 = O(1)$ and $\lim_{n\to\infty}(m_2/n) = \beta_2$ with $0 < \beta_2 < \infty$: Define

$$U_5^b = U_5 + 2p_1p_2\Big(1+\frac{m_2}{n}\Big)^{-1}\Big(1-\frac{m_1}{n}\Big) \tag{5.9}$$

Then

$$\begin{aligned} P(U_5^b \le x) = {} & G_f(x) + \frac{f\gamma}{4n}[-G_f(x) + 2G_{f+2}(x) - G_{f+4}(x)] \\ & + \frac{m_1}{2n}\Big(1+\frac{m_2}{n}\Big)^{-1}[G_f(x) - G_{f+2}(x)] + o(n^{-1}). \end{aligned} \tag{5.10}$$

Further for U_6 in (5.6),

$$P(U_6 \le x) = P(U_5^b \le x) + o(n^{-1})$$

The case where $\lim_{n\to\infty}(m_1/n) = \beta_1$ with $0 < \beta_1 < \infty$ and $m_2 = O(1)$ is similar.

(c) $\lim_{n\to\infty}(m_i/n) = \beta_i$ with $0 < \beta_i < \infty$ $(i=1,2)$: Define

$$U_5^c = \Big[1 + \Big(1+\frac{m_1}{n}\Big)^{-1}\Big(1+\frac{m_2}{n}\Big)^{-1}\Big]^{-1}\Big[U_5 + p_1p_2\textstyle\sum_{i=1}^2\Big(1+\frac{m_i}{n}\Big)\Big] \tag{5.11}$$

Then

$$\begin{aligned} P(U_5^c \le x) & = G_f(x) + \frac{f\gamma}{4n}[-G_f(x) + 2G_{f+2}(x) - G_{f+4}(x)] + o(n^{-1}) \\ & = P(U_3 \le x) + o(n^{-1}) \end{aligned} \tag{5.12}$$

Further for U_6 in (5.6),

$$P(U_6 \le x) = P(U_5^c \le x) + o(n^{-1}).$$

The proofs of these results are given in 5.3.

It is noted that all the terms denoted by $o(n^{-1})$ in the above expressions are in fact $O(n^{-2})$. As has been observed in Case (c), the asymptotic distribution up to order n^{-1} of the LBI test statistic U_5^c is the same as that of the Pillai test statistic U_3 in the complete case where no additional data is available. The case (c) sometimes ocuurs in missing data problems. On the other hand, an example of Case (a) is the invariant problem of testing on Rao's covariance structure in a growth curve model described in 1.2 where rank $(X_1)=n_1+n_2$ is fixed. In particular, the problem of testing independence in the classification model with covariates is an example of the case (c).

5.3. *Proofs.* Case (a) : $m_i=O(1)$ $(i=1,2)$. To prove (5.8), let

$$W_{ii}=\sqrt{n}\Big(\frac{V_{ii}}{n}-I\Big) \quad \text{or} \quad \frac{V_{ii}}{n}=I+\frac{1}{\sqrt{n}}W_{ii} \tag{5.13}$$

so that

$$\Big(\frac{V_{ii}}{n}\Big)^{-1}=I-\frac{1}{\sqrt{n}}W_{ii}+\frac{1}{n}W_{ii}^2+o_p(n^{-1}) \tag{5.14}$$

Then, since $m_i=O(1)$ implies $S_{ii}/n=O_p(n^{-1})$, this implies that U_6 in (5.6) is expanded as

$$\begin{aligned}(5.15)\quad U_6&=n\mathrm{tr}\Big[I+\Big(\frac{V_{11}}{n}\Big)^{-1}\Big(\frac{S_{11}}{n}\Big)\Big]^{-1}\Big(\frac{V_{11}}{n}\Big)^{-1}\Big(\frac{V_{12}}{n}\Big)\Big(\frac{V_{22}}{n}\Big)^{-1}\\
&\quad\times\Big[I+\Big(\frac{S_{22}}{n}\Big)\Big(\frac{V_{22}}{n}\Big)^{-1}\Big]^{-1}\Big(\frac{V_{21}}{n}\Big)\\
&=n\mathrm{tr}\Big(I-\frac{S_{11}}{n}\Big)\Big(\frac{V_{11}}{n}\Big)^{-1}\Big(\frac{V_{12}}{n}\Big)\Big(\frac{V_{22}}{n}\Big)^{-1}\Big[I-\frac{S_{22}}{n}\Big]\Big(\frac{V_{21}}{n}\Big)+o_p(n^{-1})\\
&=U_3-n\mathrm{tr}\Big(\frac{S_{11}}{n}\Big)V_{11}^{-1}V_{12}V_{22}^{-1}V_{21}-n\mathrm{tr}\Big(\frac{S_{22}}{n}\Big)V_{22}^{-1}V_{21}V_{11}^{-1}V_{12}+o_p(n^{-1}).\end{aligned}$$

Further, the second term in (5.5) is expanded as

(5.16)

$$\begin{aligned}R&\equiv\textstyle\sum_{i=1}^2\Big[\Big(1+\frac{m_i}{n}\Big)\Big(\frac{p_1p_2}{p_i}\Big)\Big/\Big(1+\frac{m_1}{n}\Big)\Big(1+\frac{m_2}{n}\Big)\Big]\mathrm{tr}(V_{ii}+S_{ii})^{-1}V_{ii}\\
&=p_2\Big(1-\frac{m_2}{n}\Big)\mathrm{tr}\Big(I-\frac{S_{11}}{n}\Big)+p_1\Big(1-\frac{m_1}{n}\Big)\mathrm{tr}\Big(I-\frac{S_{22}}{n}\Big)+o_p(n^{-1})\\
&=2p_1p_2-\frac{1}{n}p_1p_2(m_1+m_2)-\frac{1}{n}p_2\mathrm{tr}S_{11}-\frac{1}{n}p_1\mathrm{tr}S_{22}+o_p(n^{-1})\end{aligned}$$

Therefore from (5.7), (5.15) and (5.16),

$$U_5^a=U_6+\frac{1}{n}p_2(\mathrm{tr}S_{11}-m_1)+\frac{1}{n}p_1(\mathrm{tr}S_{22}-m_2).$$

and so using the independence between V and (S_{11},S_{22}), the characteristic function of U_5^a is evaluated as

$$\begin{aligned}\phi_5^a(t)&=E\{\exp(itU_3)\Big[1-it\Big(\frac{m_1}{n}+\frac{m_2}{n}\Big)U_3\Big]\}+o(n^{-1})\\
&=\phi_3(t)-\frac{t}{n}(m_1+m_2)\frac{\partial}{\partial t}\phi_3(t)+o(n^{-1})\end{aligned}$$

where $\phi_3(t)=E[\exp(itU_3)]$and $E(S_{ii})=m_iI$ $(i=1,2)$ is used. Since the result (5.3) implies $\phi_3(t)=(1-2it)^{-f/2}+o(n^{-1})$ and since $-2it(1-2it)^{-1}=1-(1-2it)^{-1}$,

$$2t\frac{\partial}{\partial t}\phi_3(t)=-(2it)(f/2)[(1-2it)^{-f/2}-(1-2it)^{-1-f/2}]$$

Substituting this into $\phi_5^a(t)$ above and inverting it yields (5.8). Similarly from (5.15), the characteristic function of U_6 is evaluated as

$$\phi_6(t)=E\{\exp(itU_3)[1-\frac{it}{n}(m_1+m_2)U_3]\}+o(n^{-1})$$

which is nothing but $\phi_5(t)$ up to order n^{-1}, yielding the desired result.

Case (b) : $m_1=O(1)$ and $m_2=O(n)$. Let $T_{ii}=\sqrt{m_i}\Big(\frac{S_{ii}}{m_i}-I\Big)$ so that

$$\frac{S_{ii}}{m_i}=I+\frac{1}{\sqrt{m_i}}T_{ii}+o_p(n^{-1}) \quad (i=1,2) \tag{5.17}$$

Then from (5.14)

$$(5.18)\quad \left(I+\frac{m_2}{n}\left(\frac{V_{22}}{n}\right)^{-1}\left(\frac{S_{22}}{m_2}\right)\right)^{-1}$$
$$=\left(1+\frac{m_2}{n}\right)^{-1}\left[I-\left(1+\frac{m_2}{n}\right)^{-1}\left(\frac{n}{m_2}\right)^{1/2}\frac{1}{n}W_{22}T_{22}+o_p(n^{-1})\right]^{-1}$$
$$=\left(1+\frac{m_2}{n}\right)^{-1}\left[I+\left(1+\frac{m_2}{n}\right)^{-1}\left(\frac{n}{m_2}\right)^{1/2}\frac{1}{n}W_{22}T_{22}\right]+o_p(n^{-1})$$

Hence in the same way as in (5.15),

$$U_6=U_3-n\left(1+\frac{m_2}{n}\right)^{-1}\mathrm{tr}\left(\frac{S_{11}}{n}\right)V_{11}^{-1}V_{12}V_{22}^{-1}V_{21}$$
$$+\left(1+\frac{m_2}{n}\right)^{-1}\left(\frac{n}{m_2}\right)^{1/2}\mathrm{tr}\,V_{22}^{-1}V_{21}V_{11}^{-1}V_{12}W_{22}T_{22}+o_p(n^{-1})$$

and R in (5.16) in this case becomes

$$R=p_2\left(1+\frac{m_2}{n}\right)^{-1}\mathrm{tr}\left(I-\frac{S_{11}}{n}\right)$$
$$+p_1\left(1-\frac{m_1}{n}\right)\left(1+\frac{m_2}{n}\right)^{-1}\mathrm{tr}\left[I+\left(1+\frac{m_2}{n}\right)^{-1}\left(\frac{n}{m_2}\right)^{1/2}\frac{1}{n}W_{22}T_{22}\right]+o_p(n^{-1})$$
$$=2p_1p_2\left(1+\frac{m_2}{n}\right)^{-1}-p_2\left(1+\frac{m_2}{n}\right)^{-1}\mathrm{r}\left(\frac{S_{11}}{n}\right)-p_1p_2\left(1+\frac{m_2}{n}\right)^{-1}\frac{m_1}{n}$$
$$+p_1\left(1+\frac{m_2}{n}\right)^{-2}\left(\frac{n}{m_2}\right)^{1/2}\frac{1}{n}\mathrm{tr}\,W_{22}T_{22}+o_p(n^{-1})$$

Since $U_6^b=U_3-R+2p_1p_2\left(1+\frac{m_2}{n}\right)^{-1}\left(1-\frac{m_1}{n}\right)$ with U_3 in (5.2) and since S_{22} or T_{22} is independent of V with $E(T_{22})=0$, the characteristic function of U_6^b is evaluated as

$$\phi_6^b(t)=E\{\exp(itU_3)[1-it\frac{m_1}{n}\left(1+\frac{m_2}{n}\right)^{-1}U_3]\}+o(n^{-1})$$

Hence a similar argument as in the case (a) yields (5.10) and (5.11) as well.

Case (c) : $\lim_{n\to\infty}(m_i/n)=\beta_i$ $(0<\beta_i<\infty : i=1,2)$. Using (5.17) and (5.18),

$$U_6=U_3+\left(1+\frac{m_1}{n}\right)^{-1}\left(1+\frac{m_2}{n}\right)^{-1}U_3$$

$$+\left(1+\frac{m_1}{n}\right)^{-1}\left(\frac{n}{m_1}\right)^{1/2}\mathrm{tr}\,V_{11}^{-1}V_{12}V_{22}^{-1}V_{21}W_{11}T_{11}$$
$$+\left(1+\frac{m_2}{n}\right)^{-1}\left(\frac{n}{m_2}\right)^{1/2}\mathrm{tr}\,V_{22}^{-1}V_{21}V_{11}^{-1}V_{12}W_{22}T_{22}+o_p(n^{-1})$$

while R in (5.15) in this case becomes

$$R=p_1p_2\sum_{i=1}^{2}\left(1+\frac{m_i}{n}\right)^{-1}$$
$$+\sum_{i=1}^{2}\left(\frac{p_1p_2}{p_i}\right)\left(1+\frac{m_i}{n}\right)^{-1}\left(\frac{n}{m_i}\right)^{1/2}\frac{1}{n}\mathrm{tr}\,W_{ii}T_{ii}+o_p(n^{-1}).$$

Hence using the independence between T_{ii} and V and $E(T_{ii})=0$, the characteristic function of U_6^c in (5.11) is evaluated as

$$\psi_6^c(t)=E[\exp(itU_3)]+o(n^{-1}),$$

implying (5.12).

Chapter 6

TESTS FOR INDEPENDENCE IN A TWO EQUATIONS SUR MODEL.

1. Motivation And Model

1.1. *SUR Model.* When the error terms in two different regression equations are correlated, making use of the correlation and estimating the two equations simultaneously, Zellner (1962, 1963) proposed two GLSE's (generalized least squares estimators). However, since the covariance matrix of the two equations is estimated, the OLSE is better than the GLSE when the correlation of the two equations is close enough to zero. This chapter considers the problem of testing the independence between the two equations and derives LBI tests for a one-sided alternative and a two-sided alternative. In the context of this book, first of all a SUR model is regarded as an extended GMANOVA model as has been shown in Chapter 1. Second, as will be seen below, the problem of testing the independence of the two equations is regarded as an extended form of the problem of testing independence in a missing data model treated in Chapter 5. To formulate our problems, let

$$(1.1)\qquad y_i=\tilde{X}_i\beta_i+\varepsilon_i,\quad E(\varepsilon_i)=0 \text{ and } E(\varepsilon_i\varepsilon_j')=\sigma_{ij}I_n\quad (i=1,2)$$

be a SUR model of two equations, where y_i is an $n\times 1$ vector and $\tilde{X}_i$ is an $n\times k_i$ fixed matrix of rank $k_i(i=1,2)$. Then the relation (1.1) can be rewritten as a form of a multivariate regression model with prior information on the structure of the coefficient matrix

$$(1.2)\qquad Y=XB+E,\quad E\sim N(0,I_n\otimes\Sigma)$$

where normality is assumed here for ε_i's,

$$(1.3)\qquad Y=[y_1,y_2],\quad X=[\tilde{X}_1,\tilde{X}_2],\quad \Sigma\in\mathcal{S}_+(2),\quad E=[\varepsilon_1,\varepsilon_2] \text{ and}$$
$$B=\begin{bmatrix}\beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{bmatrix}: (k_1+k_2)\times 2 \quad\text{with}\quad \beta_{12}=0 \text{ and } \beta_{21}=0.$$

By letting

$$y=\begin{pmatrix}y_1\\ y_2\end{pmatrix},\quad \beta=\begin{pmatrix}\beta_{11}\\ \beta_{22}\end{pmatrix},\quad \bar{X}=\begin{pmatrix}\tilde{X}_1 & 0\\ 0 & \tilde{X}_2\end{pmatrix}: 2n\times(k_1+k_2) \quad\text{and}\quad u=\begin{pmatrix}u_1\\ u_2\end{pmatrix},$$

it can be also rewritten as

$$(1.4)\qquad y=\bar{X}\beta+\varepsilon,\quad \varepsilon\sim N(0,\Sigma\otimes I_n).$$

Now under the multivariate regression model (1.2), the prior restrictions $\beta_{12}=0$, $\beta_{21}=0$ on the coefficient matrix B are respectively expressed as

$$(1.5)\qquad X_3BX_4=0 \quad\text{and}\quad X_5BX_6=0,$$

where $X_3=[I_{k_1},0]$, $X_4=(0,1)'$, $X_5=[0,I_{k_2}]$ and $X_6=(1,0)'$. Therefore in the view of (1.2) and (1.5), the SUR model is an extended GMANOVA model described in Chapter 1. Further it has been also shown in Chapter 1 that a necessary and sufficient condition for a GLSE $b(\Sigma\otimes I)$ to be identically equal to the OLSE $b(I\otimes I)$ is $\sigma_{12}=0$ unless $k_1=k_2$ and $M_1=M_2$, where

$$(1.6)\qquad b(\Sigma\otimes I)=(\bar{X}'[\Sigma\otimes I]^{-1}\bar{X})^{-1}\bar{X}'[\Sigma\otimes I]^{-1}y \quad\text{and}$$

$$M_i = \tilde{X}_i(\tilde{X}_i'\tilde{X}_i)^{-1}\tilde{X}_i' \quad (i=1,2)$$

In this section, we consider the problem of testing

(1.7) $\qquad H: \sigma_{12}=0$ or $\rho \equiv \dfrac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}=0$ versus $K_1: \rho>0$

and

(1.8) $\qquad H: \rho=0$ versus $K_2: \rho\neq 0.$

The null hypothesis H means that the two equations in (1.1) are independent. As will be shown soon, a canonical form of this problem is formulated as a problem of testing independence with additional data though the additional data are serially correlated.

Before an analysis is made, some references are reviewed. Zellner (1962, 1963) examined the gain in efficiency of a GLSE $b(\Sigma\otimes I)$ over the OLSE $b(I\otimes I)$ where Σ is replaced by either $\hat{\Sigma}=(\hat{\sigma}_{ij})$ or $V=(v_{ij})$ with

(1.9) $$\hat{\sigma}_{ij}=e_i'e_j/n, \quad e_i=N_iy_i, \quad N_i=I-M_i$$
$$V=Y'N_0Y/n \quad \text{and} \quad N_0=I-X(X'X)^+X'.$$

Here $(X'X)^+$ denotes the Penrose inverse. Of course e_i's are the OLS residual for each equation so that $(\hat{\sigma}_{ij})$ is covariance matrix based on these residuals. On the other hand, V is the sample covariance matrix based on the model (1.2) where the prior restrictions $\beta_{12}=0$ and $\beta_{21}=0$ are ignored. When $\tilde{X}_1'\tilde{X}_2=0$, Zellner showed that the GLSE is more efficient than the OLSE except when the correlation ρ in (1.7) between the two equations is low and/or the sample size is small. For the case $\tilde{X}_1'\tilde{X}_2\neq 0$, Kmenta and Gilbert (1968) confirmed it through a numerical method, and Revankar (1974) verified it analytically for a special case and Mehta and Swamy (1976) verified it in a general case. These results imply that when ρ is close enough to zero, the OLSE is

more efficient than the GLSE. Motivated by these facts, Kariya (1981) derived the LBI tests for the problems (1.7) and (1.8). Since there exists no uniformly most powerful invariant (UMPI) test, the local sensitivity of these tests near $\rho=0$ is relevant to the problem of choice between the GLSE and the OLSE. In a multivariate situation, Kariya, Fujikoshi and Krishnaiah (1983) also derived an LBI test and the asymptotic null and nonnull distributions. The argument here is mainly based on Kariya (1981).

2. LBI Tests.

2.1. *Canonical form via invariance.* The problems stated above are analyzed through invariance. (see also Section 4 of Chapter 4) Let $\mathcal{G}_1=R_+\times R_+\times R^{k_1}\times R^{k_2}$ with $R_+=\{x\in R\,|\,x>0\}$ and $\mathcal{G}_2=R_*\times R_*\times R^{k_1}\times R^{k_2}$ with $R_*=\{x\in R\,|\,x\neq 0\}$. Then $\mathcal{G}_1$ and $\mathcal{G}_2$ leave respectively the problems (1.7) and (1.8) with the actions

(2.1) $$y_i \rightarrow a_iy_i+\tilde{X}_ig_i$$
$$\beta_i \rightarrow a_i\beta_i+g_i \quad \text{and} \quad \sigma_{ij} \rightarrow a_ia_j\sigma_{ij},$$

where $(a_1,a_2,g_1,g_2)\in\mathcal{G}_1$ for the one-sided problem (1.7) and $(a_1,a_2,g_1,g_2)\in\mathcal{G}_2$ for the two-sided problem (1.8). Then it is easy to see that a maximal invariant under $\mathcal{G}_i$ is a function of the residuals $e_j=N_jy_j$ $(j=1,2)$ defined by (1.9) and that a maximal invariant parameter under $\mathcal{G}_1$ is the correlation coefficient ρ, while under $\mathcal{G}_2$, it is ρ^2. Here define the idempotent matrix;

(2.2) $\qquad R_j=X(X'X)^+X'-M_j$ and $r_j=\mathrm{rank}(X)-k_j$ $(j=1,2)$

and let H_j be an $n\times r_j$ matrix such that

(2.3) $\qquad H_jH_j'=R_j$ and $H_j'H_j=I_{r_j}$ $(j=1,2)$.

Further let L_0 be an $n\times q_0$ matrix such that

(2.4) $$N_0=L_0L_0' \quad \text{and} \quad L_0'L_0=I_{q_0}$$

where $q_0=n-\mathrm{rank}(X)$ and N_0 is given in (1.9), and define

(2.5) $$L_j=[L_0\ H_j] : n\times q_j \quad \text{with} \quad q_j=n-k_j\ (j=1,2).$$

Then $L_j'L_j=I_{q_j}$ and $L_jL_j'=N_j$ with N_j in (1.9). Hence defining

(2.6) $$W=L_0'Y \quad \text{and} \quad u_j=H_j'y_j\ (j=1,2),$$

the model is reduced by the invariance of $y_j\to y_j+\tilde{X}_jg_j$ $(j=1,2)$ to the model

(2.7) $$u=\begin{pmatrix}u_1\\u_2\end{pmatrix} \sim N(0, \begin{pmatrix}\sigma_{11}I_{r_1} & \sigma_{12}H_1'H_2\\ \sigma_{21}H_2'H_1 & \sigma_{22}I_{r_2}\end{pmatrix})$$
$$W\equiv(w_1,w_2) \sim N(0,I_{q_0}\otimes\Sigma)$$
u and W are independent

In other words, from the invariance viewpoint, the model (1.2) with (1.3) is equivalent to the model (2.7), which is related to a missing data model. In fact, if $H_1'H_2=0$, (2.7) is reduced to

$$u_i\sim N(0,I_{r_i}\otimes\sigma_{ii})\ (i=1,2) \quad \text{and} \quad W\sim N(0,I_{q_0}\otimes\Sigma)$$

where u_1,u_2 and W are independent. This is the model (3.1) in Chapter 5 with $n=q_0$, $m_i=r_i$ $(i=1,2)$ and $p_1=p_2=1$. Here u may be regarded as an additional data to W.

2.2. *LBI tests.* Let $V=W'W=(v_{ij})$ with $v_{ij}=w_i'w_j$, which is equal to V in (1.9), and let $S=(s_{ij})$ with $s_{ij}=u_i'H_i'H_ju_j$ where u_i's are given by (2.6). Then from (2.3), (2.4) and (2.5), it is readily seen that for the residuals e_j's in (1.9),

(2.8) $$e_i=L_i\begin{pmatrix}w_i\\u_i\end{pmatrix} \quad \text{and} \quad e_i'e_j=v_{ij}+s_{ij}\ (i,j=1,2).$$

Noticing this fact, we here state the main results but the proofs are deferred to the end of this section.

Theorem 2.1. For testing $H:\rho=0$ versus $K_1:\rho>0$, an LBI test is given by the critical region

(2.9) $$U_1\equiv[e_1'e_2/(e_1'e_1e_2'e_2)^{1/2}]>c_1$$

where $e_i=N_iy_i$ $(i=1,2)$. For the alternative $K_1':\rho<0$, the inequality is reversed.

Theorem 2.2. For testing $H:\rho=0$ versus $K_2:\rho\neq0$, an LBI test is given by the critical region

(2.10) $$U_2\equiv q_1q_2U_1^2-q_1[e_1'N_2e_1/e_1'e_1]-q_2[e_2'N_1e_2/e_2'e_2]>c_2$$

where U_1 is defined by (2.9).

In terms of the canonical model (2.7), U_1 in (2.9) is expressed as

$$U_1=(v_{12}+s_{12})/[(v_{11}+s_{11})(v_{22}+s_{22})]^{1/2}.$$

and if $H_1'H_2=0$ or $R_1R_2=0$, in which case the problem is a special case treated in Section 3 of Chapter 5, U_2 in (2.10) is shown to be the LBI test derived there. If $M_1M_2=M_1$ or if $M_1M_2=M_2$ or if $M_1M_2=0$, $R_1R_2=0$ holds. Equivalently if $\tilde{X}_1$ (or $\tilde{X}_2$) is a submatrix of $\tilde{X}_2$ (or $\tilde{X}_1$) or if $\tilde{X}_1'\tilde{X}_2=0$, the LBI test in (2.10) is nothing but the LBI test in Theorem 3.1 in Chapter 5.

By an intuitive or constructive approach, the test statistic U_1 in Theorem 2.1 is also obtained since U_1 is the correlation between the two OLS residuals e_i's. In this sense, Theorem 2.1 shows that the test based on this correlation is LBI. On the other hand, it is difficult not only to derive the test statistic U_2 in (2.10) through an intuitive approach but also to interpret it. One may propose for the alternative K_2 the critical region

(2.11) $$U_1^2>c_3.$$

Because the second and third terms in U_2 do not simultaneously become constant unless $N_1=N_2$ or $M_1=M_2$, the test defined by (2.11) is different from the test in (2.10). Further we can think of a third test based on a "complete" part $V=(v_{ij})$ in (2.7);

$$U_3 \equiv [v_{12}^2/v_{11}v_{22}] > c_4. \tag{2.12}$$

As is easily shown, this is the LRT for testing H versus K_2 *in the case that* the prior information $\beta_{12}=0$ and $\beta_{21}=0$ on the coefficient B is not available in the multivariate regression model (1.2). In this case the LRT is UMPI (e.g., see Giri (1977) page 194~195). However, in our problem where the prior information is available, as asserted in Theorem 2.2, the test (2.10) locally dominates the test (2.12) in power. The null distribution of $(n-k_1-k_2-1)U_3/(1-U_3)$ is $F(1, n-k_1-k_2-1)$, F distribution with degrees of freedom 1 and $n-k_1-k_2-1$, while the distribution of U_2 is complicated (see Section 3).

It is remarked that there exist no UMPI tests in our problems. For a most powerful invariant test for testing $\rho=0$ versus $\rho=\rho_0$ (fixed) cannot be free from the fixed ρ_0.

2.3. *Distributions of maximal invariants.* Using L_j's in (2.5), let

$$z_j=L_j'y_j \quad (j=1,2). \tag{2.13}$$

Note $z_j'=(w_j',u_j')$ from (2.6) and (2.7). Also it is easy to see that maximal invariants under the groups $G_i (i=1,2)$ are functions of $(z_1/||z_1||, z_2/||z_2||)$, and that they are regarded as maximal invariants under the groups $G_1=R_+\times R_+$ and $G_2=R_*\times R_*$ acting on z_i's by

$$z \rightarrow Az$$

with

$$z=\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad \text{and} \quad A=\begin{pmatrix} a_1 I_{q_1} & 0 \\ 0 & a_2 I_{q_2} \end{pmatrix}, \tag{2.14}$$

where $(a_1,a_2)\in G_1$ for the one-sided problem (1.7) and $(a_1,a_2)\in G_2$ for the two sided problem (1.8). Conversely any maximal invariant under G_i acting on (z_1,z_2) in this way is easily shown to be a maximal invariant under G_i with $z_j=L_j'y_j$ $(i=1,2)$. Hence considering the problems in terms of $y=(y_1',y_2')'$ acted upon by the groups G_i's is equivalent to considering them in terms of $z=(z_1',z_2')'$ acted upon by the groups G_i's. Further, it is also easy to see that maximal invariant parameters under G_1 and G_2 are ρ and ρ^2 respectively. Hence without loss of generality, we assume $\sigma_{11}=\sigma_{22}=1$.

Now from the assumption of normality of y, the distribution of z is normal with mean 0 and covariance matrix

$$\Omega(\rho)=\begin{pmatrix} I_{q_1}, & \rho L_1'L_2 \\ \rho L_2'L_1 & I_{q_2} \end{pmatrix}. \tag{2.15}$$

Here using Wijsman's theorem, we derive the density of a maximal invariant under the transformation $z\rightarrow Az$.

Lemma 2.1. Let $T=t(z)$ be a maximal invariant under G_i acting on z by $z\rightarrow Az$. Then the density of T with respect to the probability measure P_0^T induced by T under $H: \rho=0$ is given by

$$\frac{dP_{i\rho}^T}{dP_{i\rho}^T}=f_i^T(t(z)\,|\,\rho)=\frac{\int_{G_i} f(Az\,|\,\Omega(\rho))\,|A'A|^{1/2}d\nu_i(a_1,a_2)}{\int_{G_i} f(Az\,|\,\Omega(0))\,|A'A|^{1/2}d\nu_i(a_1,a_2)}, \tag{2.16}$$

where f is the normal density of z, A and $\Omega(\rho)$ are given by (2.14) and (2.15) respectively, and ν_i is an invariant measure on $G_i (i=1,2)$. Here $P_{i\rho}^T$ is the probability measure induced by T under ρ and G_i.

Checking the condition required by Wijsman's Theorem is similar to the proof of Theorem 9.1 in Chapter 3 and is checked later. For an invariant measure on G_i, we take

$$d\nu_i = |a_1 a_2|^{-1} da_1 da_2 \quad \text{for } (a_1,a_2)\in G_i \ (i=1,2),$$

where da_j denotes the Lebesgue measure.

First we treat the case of $G_1=R_+\times R_+$. After cancellation of some constants, let $K(\rho)$ be the numerator of the right hand side of (2.16). Then from the assumption of normality,

$$(2.17)\quad K(\rho)=|\Omega(\rho)|^{-1/2}\int_0^\infty\int_0^\infty \exp[-\frac{1}{2}z'A'\Omega(\rho)^{-1}Az]\chi(a)da_1da_2$$

with $\chi(a)=|a_1|^{q_1-1}|a_2|^{q_2-1}$, which is $a_1^{q_1-1}a_2^{q_2-1}$ under G_1. Evaluating $A'\Omega(\rho)^{-1}A$ in (2.17) yields

$$(2.18)\quad K(\rho)=|\Omega(\rho)|^{-1/2}\int_0^\infty\int_0^\infty \chi(a)\exp(-\frac{1}{2}a'H(\rho)a)da_1da_2$$

where $a=(a_1,a_1)'$ and $H(\rho)=(h_{ij})$ is a 2×2 matrix with elements

$$(2.19)\quad \begin{aligned} &h_{ii}\equiv h_{ii}(\rho)=z_i'(I-\rho^2Q_i)z_i \quad (i=1,2) \quad \text{and} \\ &h_{21}=h_{12}\equiv h_{12}(\rho)=-\rho z_1'(I-\rho^2Q_1)^{-1}L_1'L_2z_2. \end{aligned}$$

Here

$$(2.20)\quad Q_1=L_1'L_2L_2'L_1 \quad \text{and} \quad Q_2=L_2'L_1L_1'L_2.$$

Let

$$(2.21)\quad r\equiv r(\rho)=-h_{12}/(h_{11}h_{22})^{1/2} \text{ and } R=\begin{pmatrix}1 & r\\ r & 1\end{pmatrix}.$$

By changing a_i's into $x_1=[h_{22}/(h_{11}h_{22}-h_{12}^2)]^{-1/2}a_1$ and $x_2=[h_{11}/(h_{11}h_{22}-h_{12}^2)]^{-1/2}a_2$ and after a little algebra, we obtain $K(\rho)=\kappa_1(\rho)\kappa_2(\rho)$, where

$$(2.22)\quad \kappa_1(\rho)=2\pi|\Omega(\rho)|^{-1/2}(1-r^2)^{-(q_1+q_2-1)/2}h_{11}^{-q_1/2}h_{22}^{-q_2/2}$$

and

$$(2.23)\quad \kappa_2(\rho)=\int_0^\infty\int_0^\infty[2\pi(1-r^2)^{1/2}]^{-1}x_1^{q_1-1}x_2^{q_2-1}\exp(-\frac{1}{2}x'R^{-1}x)dx_1dx_2$$

where $x=(x_1,x_2)'$. By Kamat (1958) page 318, (2.23) becomes with $c_0=2^{-(q_1+q_2-6)}\pi^{-1}$

$$(2.24)\quad \begin{aligned}\kappa_2(\rho)=c_0[&\Gamma\left(\frac{q_1}{2}\right)\Gamma\left(\frac{q_2}{2}\right)F(-\frac{1}{2}(q_1-1),\ -\frac{1}{2}(q_2-1)\ ;\ \frac{1}{2}\ ;\ r^2)\\ &+2r\Gamma\left(\frac{q_1+1}{2}\right)\Gamma\left(\frac{q_2+1}{2}\right)F(-\frac{1}{2}(q_1-2),\ -\frac{1}{2}(q_2-2)\ :\ \frac{3}{2}\ ;\ r^2)]\end{aligned}$$

where

$$(2.25)\quad \begin{aligned} &F(a,b\ ;\ c\ ;\ x)=\textstyle\sum_{j=0}^\infty \tau_j[x^j/j!] \quad \text{and}\\ &\tau_j=\Gamma(a+j)\Gamma(b+j)\Gamma(c)/\Gamma(c+j)\Gamma(a)\Gamma(b).\end{aligned}$$

Hence the denominator of the right hand side of (2.16) is

$$(2.26)\quad K(0)=2\pi d_{11}^{-q_1/2}d_{22}^{-q_2/2}c_0\Gamma\left(\frac{q_1}{2}\right)\Gamma\left(\frac{q_2}{2}\right)$$

where

$$(2.27)\quad d_{ii}=h_{ii}(0)=z_i'z_i=y_i'N_iy_i=e_i'e_i \ (i=1,2).$$

Since $f_1^T(t(z)|\rho)=K(\rho)/K(0)$, we obtain

Theorem 2.3. Under $G_1=R_+\times R_+$, the density of a maximal invariant $T=t(z)$ with respect to the probability measure P_0^T is given by

$$(2.28)\quad f_1^T(t(z)|\rho)=|\Omega(\rho)|^{-1/2}(1-r^2)^{-(q_1+q_2-1)/2}\left[\frac{h_{11}}{d_{11}}\right]^{-q_1/2}\left[\frac{h_{22}}{d_{22}}\right]^{-q_2/2}\gamma(\rho).$$

where $\gamma(\rho)=\kappa_2(\rho)/c_0\Gamma(q_1/2)\Gamma(q_2/2)$ with $\kappa_2(\rho)$ in (2.24). Naturally $f_1^T(t(z)|0)=1$.

It is noted that as a special case, (2.28) gives the distribution of a maximal invariant in the case of $\tilde{X}_1'\tilde{X}_2=0$ for which the problem becomes a special case of the problem treated in Chapter 5 (see Section 1).

Next we treat the case of $G_2=R_*\times R_*$. In this case, the argument from (2.17) to (2.23) goes through by replacing the region of the integrals by $R_*\times R_*$ and x_i's in (2.23) by $|x_i|$'s. Then (2.24) is replaced by

$$(2.29)\quad \kappa_2'(\rho)=c_1\Gamma\left(\frac{q_1}{2}\right)\Gamma\left(\frac{q_2}{2}\right)F\left(-\frac{1}{2}(q_1-1),\ -\frac{1}{2}(q_2-1)\ ;\ \frac{1}{2}\ ;\ r^2\right)$$

where $c_1=\pi^{-1}2^{(q_1+q_2-2)/2}$ (see Nabeya (1951)).

Theorem 2.4. Under $G_2=R_*\times R_*$, the density $f_2^T(t(z)|\rho)$ of a maximal invariant $T=t(z)$ is given by (2.28), where $\kappa_2(\rho)$ is replaced by $\kappa_2'(\rho)$ in (2.29). It is noted that $f_2^T(t(z)|\rho)$ is a function ρ^2 since r^2 is a function of ρ^2.

2.4. *Proofs of Theorems.* First we note that the power function of an invariant test ϕ is expressed as

$$\pi(\phi,\rho)=\int\phi(t(z))f_i^T(t(z)|\rho)dP_0^T\quad(i=1,2).$$

Since it is shown by Lemma 1.3 in Chapter 2 that the derivatives of π with respect to ρ and ρ^2 can be computed beneath the integral sign, as in Ferguson (1967) page 235—238, for the one-sided problem (1.7), an LBI test is given by the critical region

$$(2.29)\qquad \frac{\partial}{\partial\rho}f_1^T(t(z)|\rho)|_{\rho=0}>cf_1^T(t(z)|0).$$

and for the two-sided problem (1.8), an LBI test is given by the critical region

$$(2.30)\qquad \frac{\partial}{\partial(\rho^2)}f_2^T(t(z)|\rho)|_{\rho=0}>cf_0^T(t(z)|0).$$

Hence from Theorem 2.3, we write f_1^T as $f_1^T(t(z)|\rho)=\prod_{i=1}^5Q_i(\rho)$ where

$$Q_1(\rho)=|\Omega(\rho)|^{-1/2},\quad Q_2(\rho)=(1-r^2)^{-(q_1+q_2-1)/2},$$
$$Q_3(\rho)=(h_{11}/d_{11})^{-q_1/2},\quad Q_4(\rho)=(h_{22}/d_{22})^{-q_2/2}\ \text{ and }\ Q_5(\rho)=\gamma(\rho).$$

By using the formula $(I-A)^{-1}=I+A(I-A)^{-1}$, it can be easily

shown that

$$h_{ii}(0)=d_{ii},\quad \partial h_{ii}(0)/\partial\rho=0\ \text{ and }\ \partial h_{ii}(0)/\partial(\rho^2)=d_{i+2,i+2},$$

where

$$d_{33}=z_1'L_1'L_2L_2'L_1z_1=e_1'N_2e_1\ \text{ and }\ d_{44}=z_2'L_2'L_1L_1'L_2z_2=e_2'N_1e_2,$$

and that

$$h_{12}(0)=0,\quad \frac{\partial h_{12}(0)}{\partial\rho}=d_{12},$$
$$d_{12}=z_1'L_1'L_2z_2=e_1'e_2\ \text{ and }\ \frac{\partial h_{12}(0)}{\partial(\rho^2)}=0.$$

Hence $r(0)=0$, $\partial r(0)/\partial\rho=d_{12}/(d_{11}d_{22})^{1/2}=U_1$ with U_1 in (2.9), $\partial r(0)/\partial(\rho^2)=0$, $\partial Q_i(0)/\partial\rho=0$ $(i=1,\dots,4)$ and $\partial Q_5(0)/\partial\rho=2c_2U_1$ with $c_2=\prod_{i=1}^2[\Gamma((q_i+1)/2)\Gamma(q_i/2)]$. Therefore

$$\partial f_1^T(t(z)|\rho)/\partial\rho=[\prod_{i=1}^4Q_i(0)]\ [\partial Q_5(0)/\partial\rho]=c_2U_1,$$

and substituting this into (2.29) yields Theorem 2.1. To prove Theorem 2.2, note that

$$\partial f_2^T(t(z)|0)/\partial(\rho^2)=\textstyle\sum_{i=1}^5[\prod_{j\neq i}Q_j(0)]\bar{Q}_j(0)\quad\text{with}\quad \bar{Q}_j(0)=\partial Q_j(0)/\partial(\rho^2).$$

From the above results, it can be easily verified that $\bar{Q}_1(0)\prod_{j=2}^5Q_j(0)$ $=$const., $\bar{Q}_2(0)=(q_1+q_2-1)U_1^2$, $\bar{Q}_3(0)=q_1[d_{33}/d_{11}]$, $\bar{Q}_4(0)=q_2[d_{44}/d_{22}]$ and $\bar{Q}_5(0)=(q_1-1)(q_2-1)U_1^2$. Hence we obtain

$$\partial f_2^T(t(z)|0)/\partial(\rho^2)=\text{const}+q_1q_2U_1^2-q_1(d_{33}/d_{11})-q_2(d_{44}/d_{22}).$$

Substituting this into (2.30) yields Theorem 2.2, completing the proofs.

Finally it is verified that the space $R^{q_1+q_2}$ of z in (2.14) is a Cartan $\mathscr{H}$ space where $\mathscr{H}$ is the group whose elements are of the form A in (2.14), which is a Lie subgroup of $\mathscr{G}l(q_1+q_2)$. Since $P(z=0)=0$, $\mathscr{Z}=R^{q_1+q_2}-\{0\}$ is regarded as the sample of z. Here letting for $x\in\mathscr{Z}$

$$V(x)=\{z\in\mathcal{Z}\,|\,||z-x||<||x||/2\},$$

the closure of the set

$$\{A\in\mathscr{H}\,|\,A\,V(x)\cap V(x)\neq\phi\}$$

is compact. Hence $\mathcal{Z}$ is a Cartan $\mathscr{H}$ space (see Definition 3.2 in Chapter 2).

3. The Null Distribution of the LBI Tests.

3.1. *Null Distribution of* U_1. In this section, two approximation procedures are taken for the null distributions of the LBI tests derived in Section 2; Jacobi type approximation and Edgeworth expansion type approximation. First consider the null distribution of the LBI test statistic $U_1=e_1'e_2/[e_1'e_1e_2'e_2]^{1/2}$ in (2.9) for the one sided problem $H:\rho=0$ versus $K_1:\rho>0$. Let L_i's be defined as in (2.5) and let

(3.1) $$z_i=L_i'y_i\ (i=1,2)\quad\text{and}\quad \bar{z}_2=L_1'L_2z_2.$$

Then in terms of L_i and z_i, U_1 is expressed as

(3.2) $$U_1=z_1'L_1'L_2z_2/(z_1'z_1z_2'z_2)^{1/2}$$

Further let

(3.3) $$V_1=z_1'\bar{z}_2(z_1'z_1\bar{z}_2'\bar{z}_2)^{-1/2}\quad\text{and}\quad V_2=[\bar{z}_2'\bar{z}_2/z_2'z_2]^{1/2}.$$

Then $U_1=V_1V_2$. Since under H, z_i's are independently distributed as $N(0,\sigma_{ii}I_{q_i})$, conditional on z_2, we get $z_1'\bar{z}_2\sim N(0,\sigma_{11}\bar{z}_2'\bar{z}_2)$ and so $z_1'\bar{z}_2(\bar{z}_2'\bar{z}_2)^{-1/2}\sim N(0,\sigma_{11})$ unconditionally. This implies that V_1 is independent of V_2 and $(q_1-1)^{1/2}V_1(1-V_1^2)^{-1/2}\sim t(q_1-1)$ or

(3.4) $$V_1^2\sim B(1/2,(q_1-1)/2),$$

where $t(a)$ denotes t-distribution with df(degrees of freedom) a and $B(a,b)$ denotes the beta distribution with df a and b. Since the

distribution of V_1 is symmetric about zero and since $V_2>0$, the distribution of U_1 is symmetric about zero under H. Hence

$$P(U_1>c_1|H)=\frac{1}{2}P(U_1^2>c_1^2|H)=\frac{1}{2}P(V_1^2V_2^2>c_1^2|H).$$

While V_1^2 follows (3.4),

$$V_2^2=z_2'Q_2z_2/z_2'z_2\quad\text{with}\quad Q_2=L_2'L_1L_1'L_2$$

is of the same form as test statistics for serial correlation and its distribution is not easy to treat. Press (1969) reviews this problem. Here as is often done (see, e.g., Press (1969) and Bloch and Watson (1967)), we approximate the distribution of V_2^2 by a beta distribution. From Press (1969) page 195,

(3.5) $$V_2^2/d_1\sim B(a,b)$$

where d_1 is the largest root of Q_2 and

(3.6) $$a=\left[\frac{E(V_2^2)}{d_1}\right]\left\{\left[\frac{E(V_2^2)}{\operatorname{Var}(V_2^2)}\right][d_1-E(V_2^2)]-1\right\}$$

and

(3.7) $$b=\left[1-\frac{E(V_2^2)}{d_1}\right]\left\{\left[\frac{E(V_2^2)}{\operatorname{Var}(V_2^2)}\right][d_1-E(V_2^2)]-1\right\}$$

Since $z_2'z_2$ is a sufficient and complete statistic for σ_{22} and since V_2^2 is independent of σ_{22}, V_2^2 is independent of $z_2'z_2$ (Lehmann (1959) page 162). Here

$$E(V_2^2)=\frac{E(z_2'Q_2z_2)}{E(z_2'z_2)}=\frac{\operatorname{tr}Q_2}{q_2}=\frac{\operatorname{tr}N_1N_2}{q_2}$$

and

$$\operatorname{Var}(V_2^2)=\frac{E(z_2'Qz_2)^2}{E(z_2'z_2)^2}-[E(V_2^2)]^2=\frac{2[q_2\operatorname{tr}Q_2^2-(\operatorname{tr}Q_2)^2]}{q_2^2(q_2+2)}.$$

Substituting these into (3.6) and (3.7), we obtain the approximation (3.5). Hence for $0 \le x \le d_1$

$$(3.8)\quad P(U_1 \le x \mid H) = P(V_1^2 V_2^2 \le x,\ V_2^2 \le x \mid H) + P(V_1^2 V_2^2 \le x,\ V_2^2 > x \mid H)$$

$$= P\left(\frac{V_2^2}{d_1} \le \frac{x}{d_1} \mid H\right) + P\left(\frac{V_1^2 V_2^2}{d_1} \le \frac{x}{d_1},\ \frac{V_2^2}{d_1} > \frac{x}{d_1} \mid H\right)$$

$$= I\left(a,b : \frac{x}{d_1}\right) + \int_{x/d_1}^{1} I\left(\frac{1}{2},\ (q_1-1)/2 : \frac{x}{td_1}\right) b(t : a,b)\,dt$$

where $I(\alpha,\beta : z) = \int_0^z b(t : \alpha,\beta)\,dt$ and $b(t : \alpha,\beta)$ is the density of $B(\alpha,\beta)$. The second term in the last equation above is

$$(3.9)\quad \int_{x/d_1}^{1} \left[Be\left(\frac{1}{2},\ \frac{q_1-1}{2}\right)\right]^{-1} \sum_{j=0}^{\infty} \binom{(q_1-3)/2}{j} (-1)^j \left(j+\frac{1}{2}\right)^{-1} \times \left(\frac{x}{d_1}\right)^{j+1/2} \left(\frac{1}{t}\right)^{j+1/2} b(t : a,b)\,dt$$

$$= \left[Be\left(\frac{1}{2},\ \frac{q_1-1}{2}\right) Be(a,b)\right]^{-1} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \binom{(q_1-3)/2}{j} \binom{a-j-3/2}{i} \times (-1)^{i+j} \left[\left(j+\frac{1}{2}\right)(b+i)\right]^{-1} \left(\frac{x}{d_1}\right)^{j+1/2} \left(1-\frac{x}{d_1}\right)^{i}$$

where $Be(a,b)$ denotes the beta function. In the above computation, t was transformed into $1-t$ and $(1-t)^{a-j-3/2}$ was expanded. Finally we obtain the approximation

$$(3.10)\quad P(U_1 \le x \mid H) \doteqdot \frac{1}{2}[I(a,b : x/d_1) + (3.9)].$$

For a given significance level α, the approximate significant point can be obtained from (3.10) though a numerical table for $x/d_1, a, b$ and α needs to be developed.

It is remarked that various types of approximations to the distributions of such statistics as V_2^2 are also surveyed in Durbin and Watson (1971).

Next we treat the two special cases; (1) $\tilde{X}_1'\tilde{X}_1 = 0$ and (2) $N_1N_2 = N_2$. Case (1), it is easy to see that $V_2^2 \sim B((n-k_1-k_2)/2,\ q_2/2)$, while the distribution of V_1^2 remains the same. Hence in the case

of (1), (3.10) holds exactly with $a = (n-k_1-k_2)/2$, $b = q_2/2$ and $d_1 = 1$. In the case of (2), $V_2^2 = 1$ so that $U_1^2 = V_1^2 \sim B(1/2, (q_1-1)/2)$ under H. In the case (1), Zellner (1962) considered the relative efficiency of his estimator over the OLSE. On the other hand, when $\tilde{X}_1$ is a subset of $\tilde{X}_2$, i.e., $\tilde{X}_2 = [\tilde{X}_1, \tilde{X}_3]$, in which the case (2) holds, Revankar (1974, 1976) analyzed the same problem. In Revankar (1974), an example for the case (2) is indicated.

3.2. *The null distribution of* U_2 in (2.10) is more difficult to treat. Here we consider it only for the case $N_1N_2 = N_2$. An asymptotic expansion of the distribution will be given below for a general case. As is observed above, in the case $N_1N_2 = N_2$, $V_2^2 = [e_2'N_1e_2/e_2'e_2] = 1$. Since $[e_1'N_2e_1/e_1'e_1] = [z_1'Q_1z_1/z_1'z_1]$ where $Q_1 = L_1'L_2L_2'L_1$, U_2 in (2.10) can be written as

$$U_2 = q_1\tilde{U}_2 - q_2$$

where

$$(3.11)\quad \tilde{U}_2 = z_1'[q_2Q_0 - Q_1]z_1/z_1'z_1 \quad \text{and} \quad Q_0 = \bar{z}_2(\bar{z}_2'\bar{z}_2)^{-1}\bar{z}_2'$$

Since Q_i's are idempotent under $N_1N_2 = N_2$ $(i=0,1)$ and since $Q_1Q_0 = Q_0$, there exists $P \in \mathcal{O}(q_1)$ such that

$$P[q_2Q_0 - Q_1]P' = \mathrm{diag}\{(q_2-1), -1, \dots, -1, \underbrace{0, \dots, 0}_{q_1-q_2}\}.$$

Here note that $N_1N_2 = N_2$ implies $q_1 \ge q_2$ and that $q_1 = q_2$ and $N_1N_2 = N_2$ imply $N_1 = N_2$. Hence assuming $N_1 \neq N_2$ implies $q_1 > q_2$. Now with

$$v = Pz_1, \quad l = q_2 - 1 \quad \text{and} \quad c_2 = (q_2-1)/2,$$

$$(3.12)\quad \tilde{U}_2 = l(v_1^2/v'v) - [(v_2^2 + \dots + v_{q_1}^2)/v'v]$$

where $v = (v_1, \dots, v_{q_1})'$. Since $v \sim N(0, \sigma_{11}I)$,

$$(T_1, T_2) \equiv (v_1^2/v'v,\ (v_2^2+\ldots+v_{q_2}^2)/v'v) \sim D(1/2, c_2 : c_3)$$

where $c_3=(q_1-q_2)/2$ and $D(\alpha,\beta:\gamma)$ denotes a Dirichelet distribution with density $d(t_1,t_2)=Ct_1^{\alpha-1}t_2^{\beta-1}(1-t_1-t_2)^{\gamma-1}$, $\alpha+\beta+\gamma=1$ (see Wilks (1962) page 172—182). Further since given T_2, $T_1/(1-T_2) \sim B\left(\frac{1}{2}, c_3\right)$(Wilks (1962) page 180) and $T_2 \sim B\left(c_2, c_3+\frac{1}{2}\right)$,

$$(3.13) \qquad F(x) \equiv P(\tilde{U}_2 \le x \mid H) = P\left(\frac{T_1}{1-T_2} \le \frac{x+T_2}{l(1-T_2)}\right)$$
$$= E\left[I\left(\frac{1}{2}, c_3 : \frac{x+T_2}{l(1-T_2)}\right) \mid T_2\right].$$

We evaluate this. For $x \ge l$, $F(x)=1$ since $(x+T_2)/l(1-T_2)>1$. For $l>x>0$

$$F(x)=\int_0^{(l-x)/(1+l)} I\left(\frac{1}{2}, \quad c_3 : \frac{x+t_2}{l(1-t_2)}\right) b\left(t_2 : c_2, c_3+\frac{1}{2}\right) dt_2.$$

For $0>x>-1$

$$F(x)=\int_{-x}^{(l-x)/(1+l)} I\left(\frac{1}{2}, \quad c_3 : \frac{x+t_2}{l(1-t_2)}\right) b\left(t_2 : c_2, c_3+\frac{1}{2}\right) dt_2$$

And for $-1>x$, $F(x)=0$. From these $F(x)$ can be tabulated through a numerical method.

3.3. *Asymptotic expansions of the null distributions.* Based on Kariya, Fujikoshi and Krishnaiah (1984) where a more general case is treated, we introduce the results on the asymptotic expansions of the null distributions of the LBI tests. To treat the LBI test statistic U_1 in (2.9) in the one-sided problem, it is again noted that the null distribution of U_1 is symmetric about zero. Hence U_1^2 is considered in the following form

$$\tilde{U}_1 = -\tilde{q}_0 \log(1-U_1^2) = -\tilde{q}_0 \log[1-(e_1'e_2)^2/e_1'e_1e_2'e_2]$$
$$(3.14) \qquad \tilde{q}_0 = q_0 - a$$
$$a = \frac{3}{2}+k_1+k_2-2\operatorname{rank}(X)+\operatorname{tr}H_1'H_2H_2'H_1$$

where H_i's are defined in (2.3). Let $G_f(x)$ denote the χ^2 distribution function with degrees of freedom f.

Proposition 3.1. $P(\tilde{U}_1 \le x \mid H) = G_1(x) + o(q_0^{-1})$.

The idea for the proof of this result is to use (2.7) and (2.8) and expand the statistic in terms of v_{ij} and s_{ij} (see Kariya, Fujikoshi and Krishnaiah (1984) for details). The choice of $\tilde{q}_0$ in (3.14) corresponds to the Bartlett correction.

For the null distribution of the LBI test U_2 in (2.10) in the two sided problem, we use the form:

$$\tilde{U}_2 = U_2 + 2\left[1+\frac{1}{q_0}\operatorname{tr}H_1'H_2H_2'H_1\right]$$

Proposition 3.2.

$$P(\tilde{U}_2 \le x \mid H) = G_1(x) - \frac{3}{4q_0}[G_1(x)-2G_3(x)+G_5(x)]$$
$$-\frac{1}{2q_0}\operatorname{tr}H_1'H_2H_2'H_2'H_1[G_1(x)-G_3(x)] + o(q_0^{-1}).$$

Example. A theory on investment behaviors of firms explains the investment demand of a firm at t as a function of the market value of the firm represented by the stock price F_{-1} at $t-1$ (which is regarded as a proxy for the expected profit) and the capital stock K_{-1} at $t-1$. In Theil (1971), using this theory, the investment functions of General Electric (G) company and Westinghouse (W) company are estimated as follows:

	OLSE			GLSE(SUR)		
	const.	F_{-1}	K_{-1}	const.	F_{-1}	K_{-1}
G	−10.0	0.027	0.152	−27.2	0.038	0.139
W	−0.5	0.053	0.092	−1.3	0.058	0.064

The data is anually observed for 1935–1954 (n=20) and the covariance matrix for the GLSE $b(\hat{\Sigma})$ is based on $\hat{\Sigma}=(\hat{\sigma}_{ij})$ with $\hat{\sigma}_{ij}=e_i'e_j/n$ as in (1.9) :

$$\hat{\Sigma}=\begin{pmatrix} 660.83 & 176.46 \\ 176.46 & 88.67 \end{pmatrix}$$

The correlation coefficient between the two equations is $\hat{\rho}=\hat{\sigma}_{12}/(\hat{\sigma}_{11}\hat{\sigma}_{22})^{1/2}=0.73$, which is considerably high. Hence the GLSE $b(\hat{\Sigma})$ is expected to be more efficient than the OLSE. To confirm this, we may test the hypothesis $\rho=0$. From the nature of the problem, one-sided testing problem $H: \rho=0$ versus $K: \rho>0$ seems reasonable. From (2.9), $U_1^2=0.5314$. On the other hand, $d_1=1$, $\mathrm{tr}N_1N_2=16.8036$ and $\mathrm{tr}(N_1N_2)^2=16.6312$ are obtained from the data in Theil (1971) page 296. Hence from (3.6) and (3.7), $a=82.567$ and $b=0.965$ are obtained where $q_1=q_2=17$. Here using (3.10), it turns out after evaluation that the test rejects H at 0.1 significance level.

APPENDIX

A Section 3 of Chapter 2

Description of Theorem 3.1. In this appendix we will describe a recent result due to S. Andersson (1982) concerning quotient measures and the representation of densities of maximal invariants. In what follows the notation and terminology given in Nachbin (1965) will be used. Let $\mathcal{G}$ be a locally compact σ-compact topological group which acts topologically on the locally compact σ-compact space $\mathcal{X}$. A left (right) Haar measure on $\mathcal{G}$ is denoted by $\nu_l(\nu_r)$ and the right hand modulus of $\mathcal{G}$ is Δ_r so $\Delta_r\nu_r=\nu_l$. The natural projection π from $\mathcal{X}$ to the quotient space $\mathcal{X}/\mathcal{G}=\mathcal{Q}$ is a convenient and natural choice for a maximal invariant under the action of $\mathcal{G}$ on $\mathcal{X}$. All measures on $\mathcal{X}$ and $\mathcal{Q}$ will be Radon measures.

Let μ be a measure on $\mathcal{X}$ which is relatively invariant with multiplier Δ_r^{-1}, that is,

$$\int_{\mathcal{X}} f(g^{-1}x)\mu(dx)=\Delta_r^{-1}(g)\int_{\mathcal{X}} f(x)\mu(dx)$$

for all μ-integrable f. Ignoring questions of existence of integrals for a moment, consider

$$\text{(A. 1)} \qquad \bar{f}(x)\equiv\int_{\mathcal{G}} f(gx)\nu_r(dg)$$

and note that $\tilde{f}(x)=\tilde{f}(hx)$, $h\in\mathcal{G}$ so $\tilde{f}$ is invariant. Thus, we can write $\tilde{f}(x)=\hat{f}(\pi(x))$, where $\hat{f}$ is defined on $\mathcal{Q}$. If α is a measure on $\mathcal{Q}$, we can then integrate $\hat{f}$ over $\mathcal{Q}$. This integration will be denoted by

$$J_\alpha(f)\equiv\int_{\mathcal{Q}}\left(\int_{\mathcal{G}} f(gx)\nu_r(dg)\right)d\alpha. \tag{A.2}$$

An easy calculation shows that (A.2) is relatively left invariant with multiplier Δ_r^{-1}, that is, $J_\alpha(g_1 f)=\Delta_r^{-1}(g_1)J_\alpha(f)$. A question treated in Andersson (1978) is the following; given μ on $\mathcal{X}$ which is relatively invariant with multiplier Δ_r^{-1}, under what conditions will there exist an α on $\mathcal{Q}$ so that

$$J_\alpha(f)=\int_{\mathcal{X}} f(x)\mu(dx) \tag{A.3}$$

for all μ-integrable f? A sufficient condition for the representation (A.3) to hold is provided by the notion of a proper action.

Definition 3.1: Consider the mapping K of $\mathcal{G}\times\mathcal{X}$ to $\mathcal{X}\times\mathcal{X}$ given by $K(g,x)=(gx,x)$. The action of $\mathcal{G}$ on $\mathcal{X}$ is proper if $K^{-1}(C)$ is compact for each compact subset $C\subseteq\mathcal{X}\times\mathcal{X}$.

Theorem (see Andersson (1982)). Suppose the action of $\mathcal{G}$ on $\mathcal{X}$ is proper. If μ is a relatively left invariant measure on $\mathcal{X}$ with multiplier Δ_r^{-1}, then there exists a measure α on $\mathcal{Q}$ such that

$$\int_{\mathcal{X}} f(x)\mu(dx)=\int_{\mathcal{Q}}\left(\int_{\mathcal{G}} f(gx)\nu_r(dg)\right)d\alpha \tag{A.4}$$

for all μ-integrable f.

For the remainder of this appendix, it is assumed that $\mathcal{G}$ acts properly on $\mathcal{X}$. In some situations, one has a measure μ_0 on $\mathcal{X}$ which is relatively invariant with a multiplier χ_0, that is

$$\int f(g^{-1}x)\mu_0(dx)=\chi_0(g)\int f(x)\mu_0(dx)$$

for all integrable f and $g\in\mathcal{G}$. But a representation of the form

(A.4) is still desired. To obtain such a result, the measure μ_0 needs to be modified. It is asserted in Andersson (1982) that there exists a positive continuous function η_0 on $\mathcal{X}$ such that

$$\eta_0(gx)=\Delta_r(g)\chi_0(g)\eta_0(x) \tag{A.5}$$

for $x\in\mathcal{X}$ and $g\in\mathcal{G}$. Setting $\mu=\eta_0^{-1}\mu_0$, it is easily verified that μ is relatively invariant with multiplier Δ_r^{-1}. Thus, applying the above theorem gives

$$\begin{aligned}\int_{\mathcal{X}} f(x)\mu_0(dx)&=\int_{\mathcal{X}} f(x)\eta_0(x)\mu(dx)\\ &=\int_{\mathcal{Q}}\left(\int_{\mathcal{G}} f(gx)\eta_0(gx)\nu_r(dg)\right)d\alpha\\ &=\int_{\mathcal{Q}}\left(\eta_0(x)\int_{\mathcal{G}} f(gx)\,\chi_0(g)\nu_l(dg)\right)d\alpha.\end{aligned} \tag{A.6}$$

where $\nu_l=\Delta_r\nu_r$ is a left Haar measure on $\mathcal{G}$.

Following Andersson (1982), we will now apply (A.6) to find the density function of a maximal invariant. Let μ_0 be relatively invariant with multiplier χ_0 and suppose that the random variable $X\in\mathcal{X}$ has a density f_0 with respect to μ_0. The random variable $Y\equiv\pi(X)\in\mathcal{Q}$ is maximal invariant. In the notation of (A.6), the claim is that the density of Y with respect to the measure α is φ_0 where

$$\varphi_0(\pi(x))=\eta_0(x)\int_{\mathcal{G}} f_0(gx)\chi_0(g)\nu_l(dg). \tag{A.7}$$

Given (A.6), the verification of (A.7) is identical to the case when $\mathcal{G}$ is compact (see Stein (1966) or Eaton (1983)). To verify that φ_0 is the density of Y, it suffices to show that

$$Ek(Y)=\int_{\mathcal{Q}} k(y)\varphi_0(y)\alpha(dy)$$

for suitably many functions k on $\mathcal{Q}$. But,

$$Ek(Y)=Ek(\pi(x))=\int k(\pi(x))f_0(x)\mu_0(dx)$$

$$= \int_Q \eta_0(x) k(\pi(x)) \int_g f_0(gx)\chi_0(g)\nu_l(dg) d\alpha = \int_Q k(y)\varphi_0(y)\alpha(dy)$$

with the last equality following from the definition of φ_0.

Our applicaiton of (A. 7) concerns the ratio of two densities of a maximal invariant. If f_0 and f_1 are two possible densities of X and φ_0 and φ_1 are the two induced densities of Y, then from (A. 7) the ratio $r(y)=\varphi_1(y)/\varphi_0(y)$ is given by

$$\text{(A. 8)} \qquad r(y)=\frac{\int_g f_1(gx)\chi_0(g)\nu_l(dg)}{\int_g f_0(gx)\chi_0(g)\nu_l(dg)}$$

as long as the denominator is positive.

B. Section 2 of Chapter 3

Proof of Propositions 2.2 *and* 2.3 Clearly $s(Z, V)$ is invariant i.e., $s(h(Z, V))=s(Z, V)$ for $h\in\mathscr{H}$. To show its maximality, suppose $s(Z, V)=s(Z^*, V^*)$, that is,

$$\text{(B. 1)} \qquad (Z_{12}, Z_{13})\begin{pmatrix} V_{22} & V_{23} \\ V_{32} & V_{33}\end{pmatrix}^{-1}(Z_{12}, Z_{13})' = (Z^*_{12}, Z^*_{13})\begin{pmatrix} V^*_{22} & V^*_{23} \\ V^*_{32} & V^*_{33}\end{pmatrix}^{-1}(Z^*_{12}, Z^*_{13})'$$

$$\text{(B. 2)} \qquad \begin{pmatrix} Z_{13} \\ Z_{23}\end{pmatrix} V_{33}^{-1}\begin{pmatrix} Z_{13} \\ Z_{23}\end{pmatrix}' = \begin{pmatrix} Z^*_{13} \\ Z^*_{23}\end{pmatrix} V_{33}^{*-1}\begin{pmatrix} Z^*_{13} \\ Z^*_{23}\end{pmatrix}'.$$

Let $V=SS'$ where $S\in\mathscr{A}$ (e.g., $S\in\mathscr{GU}^+(p)$). Then

$$\text{(B. 3)} \qquad \begin{pmatrix} V_{22} & V_{23} \\ V_{32} & V_{33}\end{pmatrix} = \begin{pmatrix} S_{22} & S_{23} \\ 0 & S_{33}\end{pmatrix}\begin{pmatrix} S_{22} & S_{23} \\ 0 & S_{33}\end{pmatrix}' = \begin{pmatrix} S_{22}S'_{22}+S_{23}S'_{23} & S_{23}S'_{33} \\ S_{33}S'_{23} & S_{33}S'_{33}\end{pmatrix}$$

From (B. 2) with $V_{33}=S_{33}S'_{33}$, for some $P_{33}\in\mathcal{O}(p_3)$

$$\text{(B. 4)} \qquad \begin{pmatrix} Z_{13} \\ Z_{23}\end{pmatrix} S'^{-1}_{33} P_{33} = \begin{pmatrix} Z^*_{13} \\ Z^*_{23}\end{pmatrix}(S^{*\prime}_{33})^{-1}.$$

Here let $A'_{33}=S'^{-1}_{33}P_{33}S^{*\prime}_{33}$. Then from (2. 2. 4)

$$\text{(B. 5)} \qquad A_{33}V_{33}A'_{33}=V^*_{33} \quad\text{and}\quad \begin{pmatrix} Z_{13} \\ Z_{23}\end{pmatrix} A'_{33} = \begin{pmatrix} Z^*_{13} \\ Z^*_{23}\end{pmatrix}.$$

Next from (B. 1)$\sim$(B. 3), (B. 1) is written as

$$\text{(B. 6)} \qquad (Z_{12}S'^{-1}_{22} - Z_{13}S'^{-1}_{33}S'_{23}S'^{-1}_{22})(Z_{12}S'^{-1}_{22} - Z_{13}S'^{-1}_{33}S'_{23}S'^{-1}_{22})' = (Z^*_{12}S^{*\prime-1}_{22} - Z^*_{13}S^{*\prime-1}_{33}S^{*\prime}_{23}S^{*\prime-1}_{22})(Z^*_{12}S^{*\prime-1}_{22} - Z^*_{13}S^{*\prime-1}_{33}S^{*\prime}_{23}S^{*\prime-1}_{22})'$$

Hence for some $P_{22}\in\mathcal{O}(p_2)$

$$\text{(B. 7)} \qquad (Z_{12}S'^{-1}_{22} - Z_{13}S'^{-1}_{33}S'_{23}S'^{-1}_{22})P_{22} = Z^*_{12}S^{*\prime-1}_{22} - Z^*_{13}S^{*\prime-1}_{33}S^{*\prime}_{23}S^{*\prime-1}_{22}.$$

From (B. 5), (B. 7) and $A'_{33}=S'^{-1}_{33}P_{33}S^{*\prime}_{33}$,

$$Z_{12}S'^{-1}_{22}P_{22}S^{*\prime}_{22} + Z_{13}[A_{33}S^{*\prime-1}_{33}S^{*\prime}_{23} - S'^{-1}_{33}S'_{23}S'^{-1}_{22}P_{22}S^{*\prime}_{22}] = Z^*_{12}.$$

Here let $A'_{22}=S'^{-1}_{22}P_{22}S^{*\prime}_{22}$, $A'_{32}=A'_{33}S^{*\prime-1}_{33}S^{*\prime}_{23} - S'^{-1}_{33}S'_{23}A'_{22}$, $A'_{11}=S'^{-1}_{11}S^{*\prime}_{11}$, $A'_{12}=S'^{-1}_{22}(P_{22}S^{*\prime}_{12} - S'_{12}S'^{-1}_{11}S^{*\prime}_{11})$, $\quad A'_{13}=S'^{-1}_{33}[P_{33}S^{*\prime}_{13} - S'_{13}S'^{-1}_{11}S^{*\prime}_{11} - S'_{23}A'_{12}]$ $F_{11}=Z^*_{11}-(Z_{11}A'_{11}+Z_{12}A'_{12}+Z_{13}A'_{13})$, $F_{22}=Z^*_{22}-(Z_{22}A'_{22}+Z_{23}A'_{23})$ and $F_{21}=Z^*_{21}-(Z_{21}A'_{11}+Z_{22}A'_{12}+Z_{23}A'_{13})$. Then it can be easily checked that $ZA+F=Z^*$ and $A'VA=V^*$, completing the proof of Proposition 2. 2.

To prove Proposition 2. 3, note $\Theta_{12}\Sigma^{-1}_{22.3}\Theta'_{12}=(\Theta_{12}, 0)\begin{pmatrix}\Sigma_{22} & \Sigma_{23} \\ \Sigma_{32} & \Sigma_{33}\end{pmatrix}^{-1}(\Theta_{12}, 0)'$. Setting $Z_{13}=0$ and $Z_{23}=0$ in the above proof and replacing Z_{11}, Z_{12}, Z_{21} and Z_{22} by $\Theta_{11}, \Theta_{12}, \Theta_{21}$ and Θ_{22} respectively and V by Σ, the above proof goes through for (1). To prove (2), note that if $\mathrm{Ch}(AA')=\mathrm{Ch}(BB')$ for $A: n\times m$ and $B: n\times m$, there exists $C\in\mathcal{O}(n)$ such that $CAA'C'=BB'$ where $\mathrm{Ch}(\cdot)$ denotes the vector of the ordered characteristic roots of $\cdot$. Hence applying this to $\mathrm{Ch}(\Theta_{12}\Sigma^{-1}_{22.3}\Theta'_{21})=\mathrm{Ch}(\Theta^*_{12}\Sigma^{*-1}_{22.3}\Theta^{*\prime}_{12})$ and then the above proof to $C\Theta_{12}\Sigma^{-1}_{22.3}\Theta'_{12}C'=$

$\Theta_{12}^{*}\Sigma_{22.3}^{*-1}\Theta_{12}^{*\prime}$ yields the result.

C Section 4 of Chapter 3

A second approach to the essential complete class theorem. Here we directly show that for any $\mathcal{G}$-invariant test ϕ_0 based on $T=(T_1,T_2,T_3,T_4)$, there exists a $\mathcal{G}$-invariant test ϕ_1 based on (T_1,T_2) only such that the power of ϕ_1 is as good as ϕ_0. Let $\phi_0 \in \mathcal{D}(\mathcal{G})$. By a remark at the end of Section 2, without loss of generality, ϕ_0 can be expressed as a function of T on the space $\mathcal{T}_1\times\mathcal{T}_2\times\mathcal{T}_3\times\mathcal{T}_4$. Since ϕ_0 is $\mathcal{G}$-invariant, it satisfies (2.26). Let $Q(dt_3,dt_4\,|\,T_2=t_2)$ be the conditional distribution of (T_3,T_4) given $T_2=t_2$ and define

$$\text{(C.1)}\qquad \phi_2(t_1,t_2)=\frac{\int_{\mathcal{T}_3\times\mathcal{T}_4}\phi_0(t_1,t_2,t_3,t_4)Q(dt_3,dt_4\,|\,T_2=t_2)}{\int_{\mathcal{T}_3\times\mathcal{T}_4}Q(dt_3,dt_4\,|\,T_2=t_2)}$$

where on the set on which the denominator of (4.3) is zero, we define $\phi_2(t_1,t_2)=0$. Then, noting that by Lemma 4.1, the joint distribution of (T_1,T_2,T_3,T_4), say $P_\Upsilon(dt_1,dt_2,dt_3,dt_4)$ is decomposed as

$$P_\Upsilon(dt_1,dt_2,dt_3,dt_4)=P_\Upsilon(dt_1\,|\,T_2=t_2)Q(dt_3,dt_4\,|\,T_2=t_2)P(dt_2),$$

the power function of ϕ_2 is equal to that of ϕ_0 :

$$\text{(C.2)}\qquad \pi(\phi_2,\Upsilon)=\pi(\phi_0,\Upsilon)\equiv\pi(\phi_0,\delta)$$

where $\pi(\phi_i,\Upsilon)=E_\Upsilon[\phi_i(T_1,T_2,T_3,T_4)]$ $(i=0,2)$, $\Upsilon=\Theta_{12}\Sigma_{22.3}^{-1}\Theta_{12}'$ and $\delta=(\delta_1,\dots,\delta_{n_1})$ is the ordered latent roots of Υ (see Section 2). Next, using the invariant probability measure ν on $\mathcal{O}(n_1)$, define

$$\phi_1(t_1,t_2)=\int_{\mathcal{O}(n_1)}\phi_2(ht_1h',\ ht_2h')\nu(dh).$$

Then ϕ_1 is a $\mathcal{G}$-invariant test based on (t_1,t_2). In fact, $\phi_1(kt_1k',kt_2k')=\phi_1(t_1,t_2)$ for any $k\in\mathcal{O}(n_1)$. Further,

$$\begin{aligned}\pi(\phi_1,\delta)&=\int_{\mathcal{T}_1\times\mathcal{T}_2\times\mathcal{T}_3\times\mathcal{T}_4}\phi_1(t_1,t_2)P_\Upsilon(dt_1,dt_2,dt_3,dt_4)\\&=\int_{\mathcal{O}(n_1)}\Big[\int_{\mathcal{T}_1\times\mathcal{T}_2}\phi_2(ht_1h',ht_2h')P_\Upsilon(dt_1,dt_2)\Big]\nu(dh)\\&=\int_{\mathcal{O}(n_1)}\Big[\int_{\mathcal{T}_1\times\mathcal{T}_2}\phi_2(t_1,t_2)P_{h\Upsilon h'}(dt_1,dt_2)\Big]\nu(dh)\\&=\int_{\mathcal{O}(n_1)}\pi(\phi_2,h\Upsilon h')\nu(dh)\\&=\int_{\mathcal{O}(n_1)}\pi(\phi_0,\delta)\nu(dh)=\pi(\phi_0,\delta)\end{aligned}$$

where $P_\Upsilon(dt_1,dt_2)=\int_{\mathcal{T}_3\times\mathcal{T}_4}P_\Upsilon(dt_1,dt_2,dt_3,dt_4)$. This implies that ϕ_1 is as good as ϕ_0.

D Section 8 of Chapter 3

Proof of Lemma 8.1 The following argument is based on that of linear representation of a compact group.

Lemma 1. Let B be an $n\times n$ real matrix. Suppose $\Gamma B\Gamma'=B$ for all $\Gamma\in\mathcal{O}(n)$. Then $B=\alpha I_n$ where α is a complex number.

Proof. This is a direct corollary to Schur's lemma. (See Pontriyagin (1966).)

Lemma 2. Let ν be the invariant probability measure on $\mathcal{O}(n)$. Then $\int_{\mathcal{O}(n)}\gamma_{ij}\gamma_{kl}d\nu(\Gamma)=\frac{1}{n}\delta_{ik}\delta_{jl}$, where γ_{ij} and γ_{kl} are the (i,j) and (k,l) elements of Γ respectively and δ_{ij} is Kronecker's δ.

Proof. For any $n\times n$ real matrix C, let $B=\int_{\mathcal{O}(n)}\phi C\phi' d\nu(\phi)$. Then it is easy to see $\Gamma B\Gamma'=B$ for any $\Gamma\in\mathcal{O}(n)$ since ν is left and right invariant. Hence by Lemma 1, $B=\alpha I_n$. To determine α, we take trace of B ; $\mathrm{tr}B=n\alpha=\int_{\mathcal{O}(n)}\mathrm{tr}\phi C\phi' d\nu(\phi)=\mathrm{tr}C$. Thus $\alpha=\mathrm{tr}C/n$. Take $c_{kl}=0$ for all (k,l) except the (i,j) element and let $c_{ij}=1$. Then $\alpha=\frac{1}{n}\mathrm{tr}C=\frac{1}{n}\delta_{ij}$. Comparing the both sides of $B=\int_{\mathcal{O}(n)}\phi C\phi'$

$d\nu(\phi)$ with $\alpha=\frac{1}{n}\delta_{ij}$, we obtain the result.

Lemma. 3. Let A, B be $n\times n$ real matrices. Then

(a) $$\int \mathrm{tr} A\Gamma B\Gamma' d\nu(\Gamma)=\frac{1}{n}\mathrm{tr} A \mathrm{tr} B,$$

(b) $$\int \mathrm{tr} A\Gamma \ \mathrm{tr} B\Gamma \ d\nu(\Gamma)=\frac{1}{n}\mathrm{tr} AB'.$$

Proof. (a) By Lemma. 2,

$$\int \mathrm{tr} A\Gamma B\Gamma' \,d\nu(\Gamma)=\sum_i\sum_l\sum_k\sum_j a_{ij}b_{kl}\int \gamma_{jk}\gamma_{il}d\nu(\Gamma)$$
$$=\sum_i\sum_l\sum_k\sum_j a_{ij}b_{kl}\delta_{ij}\delta_{kl}/n=\frac{1}{n}\sum_i\sum_k a_{ii}b_{kk}=\frac{1}{n}\mathrm{tr} A \ \mathrm{tr} B.$$

(b) $$\int (\mathrm{tr} A\Gamma)(\mathrm{tr} B\Gamma)d\nu(\Gamma)=\sum_i\sum_j\sum_k\sum_l a_{ij}b_{kl}\int \gamma_{jl}\gamma_{lk} \ d\nu(\Gamma)$$
$$=\sum_i\sum_j\sum_k\sum_l a_{ij}b_{kl}\,\delta_{jl}\delta_{ik}/n=\frac{1}{n}\sum_i\sum_j a_{ij}b_{ij}=\frac{1}{n}\mathrm{tr} AB'.$$

By a different argument these results have been obtained by James (1964).

E Section 8 of Chapter 3.

Proof of (1) and (2) in Lemma 8.4 (1) Note $U_1=(Z_{12}, Z_{13})$, $U_3=(Z_{32}, Z_{33})$, $(W_2, W_3)=U_1C'=(Z_{12}C'_{22}+Z_{13}C'_{23}, Z_{13}C'_{33})$ and $C(U'_1U_1+U'_3U_3)C'=I$ where $C=(C_{ij})\in\mathcal{A}$. Write $U'_1U_1+U'_3U_3=C^{-1}C'^{-1}\equiv(H_{ij})$ $(i,j=2,3)$. Then $C_{22}=(H_{22}-H_{23}H_{33}^{-1}H_{32})^{-1/2}=H_{22.3}^{-1/2}$ and $C'_{23}=-H_{33}^{-1}H_{32}C'_{22}$, while $H_{22}=Z'_{12}Z_{12}+V_{22}$, $H_{23}=H'_{32}=Z'_{12}Z_{13}+V_{23}$ and $H_{33}=Z'_{13}Z_{13}+V_{33}$. Hence

$$W_2=Z_{12}C'_{22}+Z_{13}C'_{23}=(Z_{12}-Z_{13}H_{33}^{-1}H_{32})H_{22.3}^{-1/2}$$
$$=[Z_{12}-Z_{13}(Z'_{13}Z_{13}+V_{33})^{-1}(Z'_{13}Z_{12}+V_{32})]H_{22.3}^{-1/2}$$
$$=[Z_{12}-\tilde{Z}_{13}(\tilde{Z}'_{13}\tilde{Z}_{13}+I)^{-1}(\tilde{Z}'_{13}Z_{12}+V_{33}^{-1/2}V_{32})]\ H_{22.3}^{-1/2}$$

where $\tilde{Z}_{13}=Z_{13}V_{33}^{-1/2}$. Here using $(\tilde{Z}'_{13}\tilde{Z}_{13}+I)^{-1}=I-\tilde{Z}'_{13}(I+\tilde{Z}_{13}\tilde{Z}'_{13})^{-1}$

$\tilde{Z}_{13}$ and $(I+\tilde{Z}_{13}\tilde{Z}'_{13})^{-1}=I-(I+\tilde{Z}_{13}\tilde{Z}'_{13})^{-1}\tilde{Z}_{13}\tilde{Z}'_{13}$,

$$W_2=[Z_{12}-(I+\tilde{Z}_{13}\tilde{Z}'_{13})^{-1}(\tilde{Z}_{13}\tilde{Z}'_{13}Z_{12}+\tilde{Z}_{13}V_{33}^{-1/2}V_{32})]H_{22.3}^{-1/2}$$
$$=[(I+T_2)^{-1}(Z_{12}-Z_{13}V_{33}^{-1}V_{32})]H_{22.3}^{-1/2},$$

where $T_2=Z_{13}V_{33}^{-1}Z'_{13}$. On the other hand, using the same relations

$$H_{22.3}=Z'_{12}Z_{12}+V_{22}-(Z'_{12}Z_{13}+V_{23})(Z'_{13}Z_{13}+V_{33})^{-1}(Z'_{13}Z_{12}+V_{32})$$
$$=Z'_{12}Z_{12}+V_{22}-(Z'_{12}\tilde{Z}_{13}+V_{23}V_{33}^{-1/2})(\tilde{Z}'_{13}\tilde{Z}_{13}+I)^{-1}(\tilde{Z}'_{13}Z_{12}+V_{33}^{-1/2}V_{32})$$
$$=V_{22.3}+X'X \quad \text{with} \quad X=(I+T_2)^{-1/2}(Z_{12}-Z_{13}V_{33}^{-1}V_{32}).$$

Therefore

$$\mathrm{tr}\, W'_2W_2=\mathrm{tr}(I+T_2)^{-1}X(X'X+V_{22.3})^{-1}X'$$

as is to be proved.

(2) $$\mathrm{tr}\, W'_3W_3=\mathrm{tr} Z_{13}(Z'_{13}Z_{13}+V_{33})^{-1}Z'_{13}=\mathrm{tr}\tilde{Z}_{13}(\tilde{Z}'_{13}\tilde{Z}_{13}+I)^{-1}\tilde{Z}'_{13}$$
$$=\mathrm{tr}\tilde{Z}_{13}\tilde{Z}'_{13}-\mathrm{tr}\tilde{Z}_{13}\tilde{Z}'_{13}(I+\tilde{Z}_{13}\tilde{Z}'_{13})^{-1}\tilde{Z}_{13}\tilde{Z}'_{13}$$
$$=\mathrm{tr} T_2(I+T_2)^{-1}=n_1-\mathrm{tr}(I+T_2)^{-1}$$

REFERENCES

[1] Abramowitz, M. and Stegun, L.A. (1964). *Handbook of Mathematical Functions.* Dover, New York.

[2] Anderson, T.W. (1948). On the theory of testing serial correlation. *Skand, Aktuarie Tidskrt, 31,* 88–116.

[3] Anderson, T.W. (1956). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc. 6,* 170–176.

[4] Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[5] Andersson, S.A. (1982). Distributions of maximal invariants using quotient measures. *Ann. Statist., 10,* 955–961.

[6] Andersson, S.A., Brons, H.K. and Jensen, S.T. (1983). Distribution of eigenvalues in multvariate statistical analysis. *Ann. Statist., 11,* 392–415.

[7] Banken, L. (1983). Maximality groups of multivariate models. unpublished manuscript.

[8] Banken, L. (1984). On the reduction of the general MANOVA model. Tech. Rep. Universitat Trier, Trier.

[9] Bhargava, R.P. (1975). Some one-sample hypothesis testing problems when there is a monotone sample from a multivariate normal population. *Ann. Inst. Statist. Math., 27,* 325–339.

[10] Bhattacharya, R.N. and Ghosh, J.K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist., 6,* 434–451.

[11] Bhoj, D.S. (1979). Testing equality of variances of correlated variates with incomplete data on both responses. *Biometrika, 66,* 681–683.

[12] Bloch, D.A. and Watson, G.S. (1967). A Bayesian study of the multinomial distribution. *Ann. Math. Statist., 38,* 1423–1435.

[13] Bloomfield, P. and Watson, G.S. (1975). The inefficiency of least squares. *Biometrika, 62,* 121–128.

[14] Bondar, J.V. (1976). Borel cross-sections and maximal invariants. *Ann. Statist., 4,* 866–877.

[15] Bondar, J.V. and Milones, P. (1981). Amenability: A survey for statistical applications of Hunt-Stein and related conditions on groups. *Z. Wahrcheinlichkeitstheorie verw., 57,* 103–128.

[16] Bourbaki, N. (1963). *Elements de Mathematique, 29,* Livre VI, Chap. 7 Hermann Cie., Paris.

[17] Bourbaki, N. (1966). *Elements of Mathematics,* General Topology. Part I. Hermann, Paris, and Addison-Wesley, Reading, Massachusetts.

[18] Chmielewski, M.A. (1980). Invariant scale matrix hypothesis tests under elliptical symmetry. *Jour. Multivariate Anal., 10,* 343–350.

[19] Cohen, A. (1977). A result on hypothesis testing for a multivariate normal distribution when some observations are missing. *J. Amer. Statist. Assoc., 66,* 93–98.

[20] Constantine, A.G. (1963). Some noncentral distribution problems in multivariate analysis. *Ann. Math. Statist., 34,* 1270–1284.

[21] Chou, A.D. and Lo, C.D. (1981) The local minimaxity of the LBI test under missing data. Unpublished manuscript.

[22] Das Gupta, S., Anderson, T.W. and Mudholkar, G.S. (1964). Monotonicity of the power functions of some tests of the multivariate linear hypothesis. *Ann. Math. Statist., 35,* 200–205.

[23] Das Gupta, S. and Perlman, M. (1972). On the power of the noncentral F-and Wilks' U-tests: Choosing variates for increasing the power of Hotelling's T^2-test. *J. Amer. Statist. Assoc. 69,* 174–180.

[24] Dawid, A.P. (1977). Spherical matrix distributions and a multivariate model. *Jour. Roy. Statist. Soc., B 39,* 254–261.

[25] Dempster, A.P. (1969). *Elements of Continuous Multivariate Analysis.* Addison Wesley, Reading, Mass.

[26] Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression I. *Biometrika, 37,* 409–428.

[27] Eaton, M.L. (1970). Gauss-Markov estimation for multivariate linear models. A coordinate-free approach. *Ann. Math. Statist., 41,* 528–538.

[28] Eaton, M.L. (1983). *Multivariate Statistics.* Wiley, New York.

[29] Eaton, M.L. and Kariya, T. (1980). Testing problems with additional or missing data. Tech. Rep. No.238, Univ. of Minnesota, Minneapolis.

[30] Ekbohm. G. (1976). On comparing means in the paired case with incomplete data on both responses. *Biometrika, 63,* 299–304.

[31] Farrell, R.H. (1962). Representation of invariant measures. *Illinois J. Math. 7,* 447–467.

[32] Farrell, R. (1976). *Techniques of Multivariate Calculation.* Springer Verlag., New York.

[33] Fearn, T. (1975). A Bayesian approach to growth curves. *Biometrika, 62,* 89–100.

[34] Feller, W. (1966). *An Introducion to Probability and its Applications, 2.* Wiley, New York.

[35] Ferguson, T. S. (1967). *Mathematical Statistics–A Decision Theoretic Approach.* Academic Press, New York.

[36] Fraser, D.A.S. and Ng, K.W. (1980). Multivariate regression analysis with spherical error. *Multivariate Analysis, V,* 369–386. (ed. Krishnaiah, P.R.) North-Holland, New York.

[37] Fujikoshi, Y. (1970). Asymptotic expansions of the distributions of test statistics in multivariate analysis. *Jour. Science Hiroshima Univ.,* A-I, *34,* 73–144.

[38] Fujikoshi, Y. (1973a). Monotonicity of the power functions of some tests in general MANOVA models. *Ann. Statist., 1,* 388–391.

[39] Fujikoshi, Y. (1973b). Asymptotic expansions of the nonnull distributions of three statistics in GMANOVA. *Ann. Inst. Statist. Math., 25,* 289–297.

[40] Fujikoshi, Y. (1977). Asymptotic expansions for the distributions of some multivariate tests. *Multivariate Analysis IV.* (ed. Krishnaiah, P.R.). North-Holland.

[41] Fujikoshi, Y. (1984). Error bounds for asymptotic approximations to the distribution function of an estimate in a multivariate analysis. Hiroshima Univ. Tech. Rep.

[42] Geisser, S. (1970). Bayesian analysis of growth curve. *Sankhyā, A32,* 53–64.

[43] Geisser, S. (1980). Growth Curve Analysis. *Handbook of Statistics, 1,* 89–115 (ed. Krishnaiah, P.R.) North-Holland, New York.

[44] Ghosh, J.K. (1969). Only linear transformations preserve normality. *Sankhyā, A,31,* 309–312.

[45] Giri, N. (1962). On a multivariate testing problem. *Calcutta Statist. Assoc. Bull., 11,* 55–60.

[46] Giri, N. (1964). On the likelihood ratio test of a normal multivariate testing problem. *Ann. Math. Statist., 35,* 181–190.

[47] Giri, N. (1977). *Multivariate Statistical Inference.* Academic Press, New York.

[48] Giri, N. and Kiefer, J. (1964). Local and asymptotic minimax properties of a normal multivariate testing problem. *Ann. Math. Statist., 35,* 21–35.

[49] Giri, N., Kiefer, J. and Stein, C. (1963). Minimax character of Hotelling's T^2-test in the simplest case. *Ann. Math. Statist., 34,* 1524–1535.

[50] Gleser, L.J. and Olkin, I. (1966). A k-sample regression model with covariance. *Multivariate Analysis* (ed. Krishnaiah, P.R.), 59–72. Academic Press, New York.

[51] Gleser, L.J. and Olkin, I. (1970). Linear models in multivariate analysis. *Essays in Probability and statistics,* 267–292. Wiley, New York.

[52] Hall, W. J., Wijsman, R. A. and Ghosh, J. K. (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist., 36,* 575–614.

[53] Hartley, H.O. and Hocking, R. R. (1971). The analysis of incomplete data. *Biometrics, 27,* 783–823.

[54] Hooper, P.M. (1982). Invariant confidence sets with smallest expected mèasure. *Ann. Statist, 12,* 784.

[55] Hooper, P.M. (1983). Simultaneous interval estimation in the general multivariate analysis of variance model. *Ann. Statist., 11,* 666–673.

[56] James, A.T. (1954). Normal multivariate analysis and the orthogonal group. *Ann. Math. Statist., 25,* 40–75.

[57] James, A.T. (1961). The distribution of noncentral means with known covariance. *Ann. Math. Statist., 32,* 874–882.

[58] James, A.T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist., 35,* 475–501.

[59] Jensen, D.R. and Good, I.J. (1981). Invariant distributions associated with matrix laws under structural symmetry. *Jour. Roy. Statist. Soc, B 43,* 327–332.

[60] Johnson, N.L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Distributions II.* Wiley, New York.

[61] Johnson, N.L. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*. Wiley, New York.
[62] Kadiyala, K.R. (1970). Testing for the independence of regression disturbance. *Econometrica*, *38*, 97–117.
[63] Kakwani, N.C. (1967). The unbiasedness of Zellner's seemingly unrelated regression equations estimators. *Jour. Amer. Statist. Assoc.*, *62*, 141–142.
[64] Kamat, A.R. (1958). Hypergeometric expansions for incomplete moments of the bivariate normal distribution. *Sankhyā*, *20*, 317–320.
[65] Kariya, T. (1977). A robustness property of the tests for serial correlation. *Ann. Statist*, *5*, 1212–1220.
[66] Kariya, T. (1978). The general MANOVA problem. *Ann. Statist.*, *6*, 200–214.
[67] Kariya, T. (1981a). Robustness of multivariate tests. *Ann. Statist*, *9*, 1267–1275.
[68] Kariya, T. (1981b). Tests for the independence between two seemingly unrelated regression equations. *Ann. of Statist. 9*, 381–390.
[69] Kariya, T. (1981c). Bounds for the covariance matrices of Zellner's estimator in the SUR model and 2SAE in a heteroscedastic model. *Jour. Amer. Statist. Assoc.*, *76*, 975–979.
[70] Kariya, T. (1982). A test for similarity between two groups of growth curves of economic time series variables. *Advances in Econometrics*, *1*, 301–310. (ed. Basmann, R.L. and Rhodes, G.F.)
[71] Kariya, T. and Eaton, M.L. (1977). Robust tests for spherical symmetry. *Ann. Statist*, *5*, 206–215.
[72] Kariya. T., Fujikoshi, Y. and Krishnaiah, P.R. (1984). Tests for independence of two multivariate regression models with different design matrices. *Jour. Multivar. Anal. 15*, 383–407.
[73] Kariya, T. and Kanazawa, M. (1978). A locally best invariant test for the equality of means in the presence of covariates. *Jour. Multivar. Anal.*, *8*, 134–140.
[74] Kariya, T., Krishnaiah, P.R. and Rao, C.R. (1938). Statistical inferences from multivariate normal population when some data is missing. *Development in Statistics*, *4*, 137–148.
[75] Kariya, T. and Maekawa, K. (1982). A method for approximations to the pdf's of GLSE's and its application to the

seemingly unrelated model. *Ann. Inst. Statist. Math. 34*, 281–297.
[76] Kariya, T. and Sinha, B.K. (1984). Nonnull and optimality robustness of some tests. Discussion paper No.99, Hitotsubashi Univ.
[77] Kariya, T. and Toyooka, Y. (1985). A nonlinear version of the Gauss-Markov Theorem in prediction. To appear from *Multivariate Analysis VI*.
[78] Kelker, D. (1970). Distribution theory of spherical distribution and a location scale parameter generation. *Sankhyā*, *A*, *43*, 419–430.
[79] Khatri, C.G. (1966). A note on a MANOVA model applied to problems in growth curves. *Ann. Inst. Statist. Math.*, *18*, 75–86.
[80] Khatri, C.G., Bhargava, R.P. and Shah, K.R. (1974). Distribution of regression estimate in double sampling. *Sankhyā*, *C*, *36*, 3–22.
[81] Kiefer, J. (1957). Invariance, minimax sequential estimation and continuous time process. *Ann. Math. Statist.*, *28*, 573–601.
[82] Kiefer, J. and Schwartz, R. (1965). Admissible Bayes character of T^2–, R^2–, and other fully invariant tests for classical multivariate normal problems. *Ann. Math. Statist.*, *36*, 747–770.
[83] Kmenta, J. and Gilbert, R.F. (1968). Small sample properties of alternative estimators of seemingly unrelated regression. *Jour. Amer. Statist. Assoc.*, *63*, 1180–1200.
[84] Knott, M. (1975). On the minimum efficiency of least squares. *Biometrika*, *62*, 129–132.
[85] Koehn, U. (1970). Global cross sections and the densities of maximal invariants. *Ann. Math. Statist.*, *41*, 2045–2056.
[86] Krishnaiah, P.R. (1966). Simultaneous test procedures for general MANOVA models. *Multivariate Analysis II*. 121–143. Academic Press.
[87] Krishnaiah, P.R. (1978). Some recent developments on real multivariate distributions. *Developments in Statistics*, *1*, 135–169. Academic Press, New York.
[88] Kruskal, W. (1968). When are Gauss-Markov and least squares estimators identical? A coordinate free approach. *Ann. Math. Statist.*, *39*, 70–75.
[89] Kshirsargar, A. M. (1972). *Multivariate Analysis*. Marcel

Decker, New York, N.Y.
[90] Kudo, H. (1955). On minimax invariant estimates of the transformation parameter. *Natural Science Report*, *6*, 31–73. Ochanomizu Univ.
[91] Lee, J.C. and Geisser, S. (1972). Growth curve prediction. *Sankhyā*, *A34*, 393–412.
[92] Lee, J.C. and Geisser, S. (1975). Applications of growth curve prediction. *Sankhyā*, *A 37*, 239–256.
[93] Lehmann, E.L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.
[94] Little, R.J.A. (1976). Inference about means from incomplete multivariate data. *Biometrika*, *63*, 593–604.
[95] Marden, J.I. (1981), Invariant tests on covariance matrices. *Ann. Statist.*, *9*, 1258–1266.
[96] Marden, J.I. (1982). Minimal complete classes of tests of hypotheses with multivariate one-sided alternatives. *Ann. Statist.*, *10*, 962–970.
[97] Marden, J.I. (1982). Admissibility of invariant tests in the general multivariate analysis of variance problem. Tech. Rep. Univ. of Illinois.
[98] Marden, J.I. (1983). Admissibility of invariant tests in the general multivariate analysis of variance problem. *Ann. Statist.*, *11*, 1086–1099.
[99] Marden, J. and Perlman, M.D. (1980). Invariant tests for means with covariates. *Ann. Statist.*, *8*, 25–63.
[100] Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.
[101] Matthes, T.K. and Truax, D.R. (1967). Tests of composite hypotheses for the multivariate exponential family. *Ann. Math. Statist.*, *38*, 681–697.
[102] Mehta, J.S. and Swamy, P.A.V.B. (1976). Further evidence on the relative efficiencies of Zellner's seemingly unrelated regressions estimator. *J. Amer. Statist. Assoc.*, *71*, 634–639.
[103] Memon, A.Z. and Okamoto, M. (1970). The classification statistic in covariate discriminant analysis. *Ann. Math. Statist.*, *41*, 1491–1499.
[104] Mitra, S.K. (1970), A density free approach to the matrix variate beta distribution. *Sankhyā*, *A32*, 81–88.

[105] Morrison, D.F. and Bhoj, D.S. (1973). Power of the likelihood ratio test on the mean vector of the normal distribution with missing observations. *Biometrika*, *60*, 365–368.
[106] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.
[107] Nabeya, S. (1951). Absolute moments in 2-dimensional normal distribution. *Ann. Ins. Statist. Math.* *3*, 2–6.
[108] Nabeya, S. and Kariya, T. (1984). Transformations preserving multivariate normal distributions. Unpublished manuscript.
[109] Nachbin, L. (1965). *The Haar Integral.* Van Nostrand, New York.
[110] Olkin, I. and Rubin, J. (1964). Multivariate beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.*, *35*, 261–269.
[111] Palais, R.S. (1961). On the existence of slices for actions of noncompact Lie groups. *Ann. Math.*, *73*, 295–323.
[112] Perlman, M. (1972). Monotonicity of the power function of Pillai's trace test. Tech. Rep. 194. Dept. of Statist., Univ. of Minnesota.
[113] Perlman, M.D. and Olkin, I. (1980). Unbiasedness of invariant tests for MANOVA and other multivariate problems. *Ann. Statist.*, *8*, 1326–1341.
[114] Pontryagin, L. (1966). *Topological groups*, 2nd ed., translated from the Russian. Gordon and Breach, New York.
[115] Potthoff, R.F. and Roy, S.N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve models. *Biometrika*, *51*, 313–326.
[116] Press, S.J. (1969). On serial correlation. *Ann. Math. Statist.*, *40*, 186–196.
[117] Rao, C.R. (1949). On some problems out of discrimination with multiple character. *Sankhyā*, *9*, 343–366.
[118] Rao, C.R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, *52*, 447–458.
[119] Rao, C.R. (1966). Covariance adjustment and related problems in multivariate analysis. *Multivariate Analysis I*, 87, (ed. Krishnaiah, P.R.), Academic Press, New York.
[120] Rao, C.R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of

signals. *Fifth Berk. Symp. Math. Statist. Prob.*, *1*, 355-372. Univ. of California, Berkeley.

[121] Rao, C.R. and Mitra, S.K. (1971). *Generalized Inverse of Matrices and its Application.* Wiley, New York.

[122] Revankar, N.S. (1974). Some finite sample results in the context of two seemingly unrelated regression equations. *J. Amer. Statist. Assoc.*, *69*, 187-190.

[123] Revankar, N.S. (1976). Use of restricted residuals in SUR systems: some finite sample results. *J. Amer. Statist. Assoc.*, *71*, 183-188.

[124] Rubin, D.B. (1974). Characterizing the estimation of parameters in incomplete data problems. *J. Amer. Statist. Assoc.* *69*, 467-474.

[125] Rubin, D.B. (1976). Inference and missing data. *Biometrika*, *63*, 581-592.

[126] Salaevskii, O.V. (1968). Minimax character of Hotelling's T^2-test. *Sov. Math. Dokl.*, *7*, 719.

[127] Sarkar, S.K. (1980). Inference from incomplete multivariate samples some testing and related problems. Ph.D. Thesis., University of Calcutta, Calcutta, India.

[128] Schwartz, R. (1967a). Local minimax tests. *Ann. Math. Statist.*, *38*, 340-360.

[129] Schwartz, R. (1967b). Admissible tests in multivariate analysis of variance. *Ann. Math. Statist.*, *38*, 698-710.

[130] Schwartz, R. (1969). Invariant proper Bayes tests for exponential families. *Ann. Math. Statist.*, *40*, 270-283.

[131] Sinha, B.K. (1983). Rejection of multivariate outliers. *Ann. Statist.*, *12*, 1558-1565.

[132] Srivastava, V.K. (1970). The efficiency of estimating seemingly unrelated regression equations. *Ann. Inst. Statist. Math.*, *22*, 483-493.

[133] Srivastava, V.K. and Dwivedi, T.D. (1979). Estimation of seemingly unrelated regression equations: A brief survey. *Jour. Econometrics*, *10*, 15-32.

[134] Stein, C. (1956a). The admissibility of Hotelling's T^2-test. *Ann. Math. Statist.*, *27*, 616-623.

[135] Stein, C. (1956b). Some problems in multivariate analysis. Part I. Tech. Rep. No.6, Dept. of Stat. Stanford Univ.

[136] Stein, C. (1966). Notes on Multivariate Analysis. Stanford

Univ. (Notes recorded by M.L. Eaton).

[137] Sugiura, N. (1969). Asymptotic Expansions of the distributions of the likelihood ratio criteria for covariance matrix. *Ann. Math. Statist.* *40*, 2051-2063.

[138] Sugiura, N. (1973). Asymptotic non-null distributions of the likelihood ratio criteria for covariance marrix under local alternatives. *Ann. Statist.* *1*, 718-719.

[139] Theil, H. (1971). *Principles of Econometrics.* Wiley, New York.

[140] Wald, A. (1950). *Statistical Decision Functions.* Wiley, New York.

[141] Ware, J.H. and Bowden, R.E. (1977). Circadian rhythm analysis when output is collected at intervals. *Biometrics*, *33*, 566-571.

[142] Wijsman, R.A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Fifth Berk. Symp. Math. Statist. Prob.*, *1*, 389-400. Univ. of California, Berkeley.

[143] Wijsman, R.A. (1969). General proof of termination with probability one of invariant sequential probability ratio tests based on multivariate observations. *Ann. Math. Statist.*, *38*, 8-24.

[144] Wijsman, R.A. (1978). Distribution of maximal invariants, using global and local cross-sections. Preprint, Univ. of Illinois at Urbana-Champaign.

[145] Wijsman, R.A. (1983). Proper action in steps, with application to density ratios of maximal invariants. Univ. of Illinois at Urbana-Champaign.

[146] Wijsman, R.A. (1984), Global cross sections as a tool for factorization of measures and distribution of maximal invariants. Univ. of Illinois at Urbana-Champaign.

[147] Wilks, S.S. (1962). *Mathematical Statistics.* Wiley, New York.

[148] Zellner, A. (1962). An efficient method of estimating seemingly unrelated regression and tests for aggregation bias. *J. Amer. Statist. Assoc.*, *57*, 348-368.

[149] Zellner, A. (1963). Estimators for seemingly unrelated regression equations: some exact sample results. *J. Amer. Statist Assoc.*, *58*, 977-992.

[150] Zerbe, G.O. and Jones, R.H. (1980). On application of growth

curve techniques to time series data. *J. Amer. Statist. Assoc.*, *75*, 507–509.
[151] Zykind, 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.

INDEX

Authors