ON TESTS FOR SELECTION OF VARIABLES AND INDEPENDENCE UNDER MULTIVARIATE R (U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS T KARTYA ET AL AUG 85 TR-85-33
ON TESTS FOR SELECTION OF VARIABLES AND INDEPENDENCE UNDER MULTIVARIATE REGRESSION MODEL*

T. Kariya
Hitotsubashi University

Y. Fujikoshi
Hiroshima University

P. R. Krishnaiah
University of Pittsburgh

Center for Multivariate Analysis
University of Pittsburgh

Approved for public release; distribution unlimited.
ON TESTS FOR SELECTION OF VARIABLES AND INDEPENDENCE UNDER MULTIVARIATE REGRESSION MODEL*

T. Kariya  
Hitotsubashi University  
Y. Fujikoshi  
Hiroshima University  
P. R. Krishnaiah  
University of Pittsburgh

August 1985  
Technical Report No. 85-33  
Center for Multivariate Analysis  
Fifth Floor, Thackeray Hall  
University of Pittsburgh  
Pittsburgh, PA 15260

*Research sponsored by the Air Force Office of Scientific Research (AFSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon. This work was done by the authors at the Center for Multivariate Analysis, University of Pittsburgh.
1. INTRODUCTION AND SUMMARY

Consider the classical MANOVA model

\[ Y = X\Theta + E \]  \hspace{1cm} (1.1)

where \( E \sim N(0, I_n \otimes \Sigma) \), \( E = (E_1, E_2) \), \( Y = (Y_1, Y_2) \), \( X = (X_1, X_2) \), and

\[ \Theta = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}. \]  \hspace{1cm} (1.2)

Also, \( \Sigma_{ij} \) and \( X_i \) are of order \( p_i \times p_j \) and \( n \times r_i \) respectively, \( p = p_1 + p_2 \) and \( r = r_1 + r_2 \). In addition, \( Y_i \) and \( E_i \) are of order \( n \times p_i \). Then, we are interested in testing the following hypotheses:

Problem [I]  \( H: \theta_{12} = 0 \) and \( \Sigma_{12} = 0 \) vs \( K: \) not \( H \)

Problem [II]  \( H: \theta_{12} = 0, \theta_{21} = 0 \) and \( \Sigma_{12} = 0 \) vs \( K: \) not \( H \)

Problem [III]  \( H: \Sigma_{12} = 0 \) under \( \theta_{12} = 0 \) and \( \theta_{21} = 0 \)

vs \( K: \Sigma_{12} \neq 0 \) under \( \theta_{12} = 0 \) and \( \theta_{21} = 0 \).

The motivation behind each problem is stated in Section 2 and some examples are also given there. In this section, some formal features of the problems are made clear and our results are briefly summarized together with some results in the literature. A basic feature in the problems treated here is that in each hypothesis the independence \( (\Sigma_{12} = 0) \) between \( Y_1 \) and \( Y_2 \) is included corresponding to the structure of the regression coefficient matrix \( \Theta \).
In Problem [I], the hypothesis will be regarded in Section 2 as a formulation of the hypothesis of no causality from $X_1$ to $Y_2$ where $X = [X_1, X_2]$ may be random but is fixed with full rank. Also $\theta_{12} = 0$ in the hypothesis may be viewed as a special case of the GMANOVA (general MANOVA) hypothesis

$$M_1 \Theta M_2 = 0$$

(1.3)

where $M_1$ and $M_2$ are fixed matrices of full rank. In fact, $M_1 = [I, 0]$ and $M_2 = \begin{pmatrix} 0 \\ I \end{pmatrix}$ implies $\theta_{12} = 0$. The problem of testing (1.3) in the GMANOVA model of the form $Y = Z_1 \Theta Z_2 + E$ is known as the GMANOVA problem and has been treated by Potthoff and Roy (1964), Rao (1965, 1966), Khatri (1967), Krishnaiah (1969), Gleser and Olkin (1970), Fujikoshi (1973), Kariya (1978), Marden (1982a) and others. Potthoff and Roy (1964) proposed the GMANOVA model and considered ad hoc procedures for testing the general linear hypothesis on the location parameters. Rao (1965, 1966) reduced the problem to testing the general linear hypothesis under a conditional model. Khatri (1966) derived the LRT for testing the general linear hypothesis under Potthoff and Roy model. Later, Gleser and Olkin (1970) gave a canonical reduction of the problem and discussed the LRT procedure using the canonical form. However, they have not treated Problem [I]. It should be noted that when (1.3) is dealt with, the presence of a general $M_2$ in (1.3) affects the covariance structure so that in the case of (1.3) $\Sigma_{12} = 0$ should be replaced by the hypothesis

$$\Sigma = M_2^T M_2 + P_2^T P_2$$

(1.4)

where $P_2$ is a matrix of full rank satisfying $P_2^T M_2 = 0$ (see Kariya (1985), pp 175-176). That is, Problem [I] is considered equivalent to the problem of testing (1.3) and (1.4) simultaneously since both problems give the same canonical form. Hence, solving the former problem implies solving the latter.
problem and vice versa. Now for Problem [I], applying the invariance principle, we first expand the power function of an invariant test in the neighborhood of the hypothesis and based on it propose a test, in Section 3, which maximizes the power in a slightly restricted neighborhood of the null hypothesis. Of course due to the local optimality, it is admissible. The test statistic there is a linear combination of the LBI test statistic \( T_1 \) for testing \( \theta_{12} = 0 \) (Schwartz (1967)) and the LBI test statistic \( T_2 \) for testing \( \Sigma_{12} = 0 \) under \( \theta_{12} = 0 \). The latter test is equivalent to the LBI test of independence with some data missing in Eaton and Kariya (1983). Because of the form of \( T_2 \), \( T_1 \) and \( T_2 \) are correlated and so our test is not equal to a test combining the two independent LBI test statistics \( T_1 \) and \( T_3 \) for the two separate hypotheses \( \theta_{12} = 0 \) and \( \Sigma_{12} = 0 \) (without \( \theta_{12} = 0 \)), though \( T_1 \) is the same. This is a feature of the joint treatment of the two hypotheses and it implies that a test combining the two independent statistics \( T_1 \) and \( T_3 \) which are LBI for each hypothesis does not maximize the local power in any direction except for the case that the test depends on \( T_1 \) only and that the alternative space is restricted to the space on which \( \Sigma_{12} = 0 \). The problem of how to combine independent tests is discussed in the literature (see e.g., Marden (1982b)), though we do not discuss it here.

But the LRT statistic for Problem [I] gives a natural combination for two separate hypotheses. In fact, it is the product of the two independent LRT statistics. This might support the idea that we separately treat the hypotheses and then combine the two tests. However, as has been observed in Eaton and Kariya (1983), even when \( \theta_{12} = 0 \), the LRT for testing \( \Sigma_{12} = 0 \) ignores the additional information (data) available through \( \theta_{12} = 0 \). In this sense, the above fact may not be seriously taken into account. The asymptotic null distributions of the test
based on $T_1$ and $T_2$ and the LRT are derived in Section 5 and the unbiasedness of the LRT is shown. It is noted that the group leaving the problem invariant is small so that the power function of an invariant test including the LRT depends on many parameters including the canonical correlations.

In Problem [II], the hypothesis will be regarded in Section 2 as the hypothesis of no additional information in canonical correlation analysis or a formulation of the hypothesis of no causality from $X_1$ to $Y_2$ and from $X_2$ to $Y_1$, where $X$ may be random but is fixed with full rank. Here the restrictions $\theta_{12} = 0$ and $\theta_{21} = 0$ are special cases of

$$M_1^0M_2 = 0 \text{ and } M_3^0M_4 = 0. \quad (1.5)$$

However, the two GMANOVA type restrictions in (1.5) cannot be expressed as a single GMANOVA hypothesis of the form $M_1^0M_2 = 0$. That is, the problem of testing (1.5) even in our MANOVA model $Y = X\theta + E$ is no longer the GMANOVA problem and difficult to treat unless $M_1$ and $M_3$ are nested relative to $X'X$ or orthogonal relative to $X'X$, i.e., $M_1'XM_3' = 0$ (see Kariya (1985) p 143). Since $M_1 = [I, 0]$ and $M_3 = [0, I]$ in our present case, $M_1'XM_3' = X_1'X_2$ and hence without $X_1'X_2 = 0$ the problem of testing.

Problem [IV]: $H: \theta_{12} = 0$ and $\theta_{21} = 0$ is difficult to treat. In fact, it is not only difficult to derive the LRT explicitly but it is also difficult to find a similar test detecting both $\theta_{12} = 0$ and $\theta_{21} = 0$ in a meaningful manner (see Section 7). On the other hand, the hypothesis on the covariance structure which corresponds consistently to the hypothesis (1.5) is expressed as

$$\Sigma = M_2^0M_2' + P_2\Delta_2P_2' \quad \text{and} \quad \Sigma = M_4^0M_4' + P_4\Delta_4P_4'. \quad (1.6)$$
where $P_2 M_2 = 0$ and $P_4 M_4 = 0$. Since in Problem [II] $M_2 M_4 = 0$, we can take $P_2 = M_4$ and $P_4 = M_2$ so that the two covariances in (1.6) become the same.

In Section 4, first in the case of $X_1 X_2 \neq 0$ we analyze Problem [II] via invariance, but because the group leaving the problem invariant is quite small, no sufficient reduction is obtained and the space of a maximal invariant parameter is of high dimension. Hence in the case of $X_1 X_2 \neq 0$ we simply show the unbiasedness of the LRT derived by Fujikoshi (1982). The LRT statistic here is the product of the three LRT statistics for the three separate hypotheses $\theta_{12} = 0$, $\theta_{21} = 0$ and $\Sigma_{12} = 0$ but the three are dependent. It is noted that it is difficult to consider the monotonicity of the power function of the LRT because of the high dimensional parameter space. Next in the case of $X_1 X_2 = 0$, we expand the power function of an invariant test in the neighborhood of the null hypothesis and based on it propose a test, which is a linear combination of three statistics $R_1$, $R_2$ and $R_3$. Here similar to Problem [II], $R_1$ and $R_2$ are respectively the LBI test statistics for the hypotheses $\theta_{12} = 0$ and $\theta_{21} = 0$ and $R_3$ is the LBI test statistic for $\Sigma_{12} = 0$ under $\theta_{12} = 0$ and $\theta_{21} = 0$. Hence $R_1$, $R_2$ and $R_3$ are dependent. This is a feature different from the separate treatment of the three hypotheses. The asymptotic null distributions of this test as well as the LRT are given in Section 5.

Problem [III] was treated by Kariya, Fujikoshi and Krishnaiah (1984) (abbreviated as KFK henceforth). In this model $X$ is fixed but it may not be of full rank. The model (1.1) with $\theta_{12} = 0$ and $\theta_{21} = 0$ is regarded as a combined expression of two correlated multivariate regression models with different design matrices. When $p_1 = p_2 = 1$, Zellner (1962, 1963) called it a seemingly unrelated regression (SUR) equation model, while KFK called it a correlated regression equations (CRE) model. As has
been discussed above, the information \( \theta_{12} = 0 \) and \( \theta_{21} = 0 \) on \( \Theta \) cannot be expressed as a single GMANOVA restriction and when it is expressed as the form in (1.5), \( M_1 \) and \( M_3 \) are neither nested nor orthogonal relative to \( X'X \) (see Kariya (1985) p 143, 159-163, 207). Therefore, the model itself is not put in the framework of the GMANOVA model and even the LRT for Problem [II] is difficult to derive (in the case that \( X_1 X_2 \neq 0 \) or \( X_1 \) is not nested to \( X_2 \)).

In Section 7, based on the idea of Mukherjee and Chandra (1984), we compare the power of the LRT-like test, the Pillai type test and the LBI test proposed in KFK. The comparison is made asymptotically in \( n \) with contiguous alternatives and asymptotically in small canonical correlations \( \rho_1 \geq \cdots \geq \rho_p \) (as \( \rho_1 \to 0 \)). There it is shown that the Pillai type test is asymptotically (in the same sense) equivalent to the LBI test and that for those small \( \rho_i \)'s such that \( \tilde{\tau} = \varepsilon_1^4 / (\varepsilon_1^2)^2 > \) some constant \( c \), the LRT-like test is asymptotically better than the LBI test, while for those small \( \rho_i \)'s such that \( \tilde{\tau} < c \), the LBI test is better. Further, because in the case of \( X_1 = X_2 = X_0 \), the model in (1.1) is reduced to

\[
[Y_1, Y_2] = X_0 [\theta_1, \theta_2] + [E_1, E_2],
\]

which is nothing but the MANOVA model. Therefore the comparison in Section 7 holds as it stands for the LRT and the Pillai (LBI) test of independence in the MANOVA model. This is a corollary of our result.
2. MOTIVATION OF THE WORK

The motivation behind Problem [I]: $\theta_{12} = 0$ and $\theta_{21} = 0$ is associated with the problem of no causality from $X_1$ to $Y_2$ and total exogeneity of $X_1$ for $Y_2$. We will first write the model (1.1) as two correlated classical multivariate regression models.

\[
Y_1 = X_1\theta_{11} + X_2\theta_{21} + E_1
\]

\[
Y_2 = X_1\theta_{12} + X_2\theta_{22} + E_2.
\]

Then the hypothesis $\theta_{12} = 0$ is equivalent to no effect of $X_1$ on $Y_2$ as in the usual case. However, since $E_1$ and $E_2$ are correlated, the regression equation of $Y_2$ under $\theta_{12} = 0$, conditional on $Y_1$ is expressed as

\[
Y_2 = X_2\theta_{22} + (Y_1 - X_1\theta_{11} - X_2\theta_{21})\Sigma^{-1}_{11}\Sigma_{12} + E_3.
\]

In this sense the effect of $X_1$ on $Y_2$ still remains unless $\Sigma_{12} = 0$. Therefore the hypothesis $\theta_{12} = 0$ and $\Sigma_{12} = 0$ in Problem [I] is considered as a formulation of no causality from $X_1$ to $Y_2$ or total exogeneity of $X_1$ for $Y_2$. An example for Problem [I] is found in a problem of economic policy evaluation. Suppose the model (2.1) is a reduced form of an econometric simultaneous equations model which describes the interaction of economic variables, and $X_1$ is a matrix of policy variables (tax rate, government investment etc.). Then the hypothesis in Problem [I] is interpreted as no effect of the policy on some economic variables such as inflation rate, sales, consumption etc.

The motivation behind Problem [II]: $\theta_{12} = 0$, $\theta_{21} = 0$ and $\Sigma_{12} = 0$ is similarly given in association with no causality from $X_1$ to $Y_2$ and no causality from $X_2$ to $Y_1$. In addition, the problem is also considered as a formulation
of no additional information hypothesis in canonical correlation analysis, which was given by Fujikoshi (1982) based on Mckay (1977). To see this, suppose there are two groups of measurements (variables), say $x_1$ and $x_2$, where $x_1$'s are $r_1 \times 1$ random vectors with means $\mu_i$ and joint covariance matrix $\Sigma = (\Sigma_{ij}) : (r_1 + r_2) \times (r_1 + r_2)$ with $\Sigma_{ij} : r_1 \times r_j$ ($i, j = 1, 2$). Let $\delta^2(x_1, x_2)$ denote the sum of squares of the canonical correlations between $x_1$ and $x_2$

$$\delta^2(x_1, x_2) = \text{tr} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21},$$

(2.2)

which is regarded as a measure of total correlation between $x_1$ and $x_2$. Sometimes for each group, there are some other measurements available, say $y_1$ and $y_2$, which appear to be of some relevance for the correlation between the two groups where $y_i : p_i \times 1$ ($i = 1, 2$). Then adding these variables to $x_1$ and $x_2$, the total correlation is measured by the sum of the canonical correlations between $z_1 = (x_1', y_1')'$ and $z_2 = (x_2', y_2')'$, say $\delta^2(z_1, z_2)$ as in the case of (2.2). But the real or significant relevance of including the additional variables $y_1$ and $y_2$ may be in question relative to the original variables. This question gives the following testing problem

$$H: \delta^2(z_1, z_2) = \delta^2(x_1, x_2) \text{ vs } K: \delta^2(z_1, z_2) > \delta^2(x_1, x_2).$$

(2.3)

Using a conditional argument, Fujikoshi (1982) showed that this problem is equivalent to Problem [II] and he derived the LRT where $X_i$'s and $Y_j$'s in (1.1) are the sample matrices of $x_i$'s and $y_j$'s. Mckay (1977a) treated the hypothesis $\delta^2(z_1, x_2) = \delta^2(x_1, x_2)$ (i.e., no additional information in $y_1$ relative to $(x_1, x_2)$) and showed that this hypothesis is equivalent to $\Sigma_{12}$ in the model (1.1). Some related topics are also found in Mckay (1977b) and Rao (1970).

The motivation for Problem [III] is stated in KFK (1984), Zellner (1962, 1963) and the articles therein.
3. TESTS FOR \( \theta_{12} = 0 \) AND \( \Sigma_{12} = 0 \)

Based on the motivation stated in Section 2, we here consider the problem of testing the hypothesis \( H \) against \( K \) where

\[
H: \theta_{12} = 0, \Sigma_{12} = 0, \quad K: \text{not } H.
\]  

(3.1)

First we make an invariance consideration into the problem and obtain an expression for the local behavior of the power function of an invariant test in the neighborhood of the null hypothesis. Based on the expression, an invariant test together with the LRT will be proposed and then the null distributions of these test statistics will be given in Section 5. To begin with, a canonical reduction of the problem is performed. Write

\[ X = PA[0] \text{ with } P \in O(n) \text{ and } A = (X'X)^{1/2} \in G \mathcal{L}(r) \]

and express \( \theta_{12} = 0 \) as

\[ M_1 \Theta M_2 = 0 \text{ with } M_1 = [I_{p_1}, 0] \text{ and } M_2 = \begin{pmatrix} 0 \\ I_{p_2} \end{pmatrix}, \]

(3.2)

where \( O(n) \) denotes the group of \( n \times n \) orthogonal matrices and \( G \mathcal{L}(r) \) the group of \( r \times r \) nonsingular matrices. Further let

\[ M_1 A^{-1} = F(I, 0) \Psi \text{ with } F \in G \mathcal{L}(r), \quad \text{and } \Psi \in O(r) \]

and

\[
Z = \begin{pmatrix} \Psi & 0 \\ 0 & I_{n-r} \end{pmatrix} \quad \text{and} \quad \eta = \begin{pmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{pmatrix} \quad \text{with} \quad \eta_{11} = \Psi \Theta A \Psi.
\]

(3.3)

Then the problem is to test \( (I, 0) \eta_{12} = 0 \) and \( \Sigma_{12} = 0 \), i.e.,

\[
H: \eta_{12} = 0, \Sigma_{12} = 0.
\]  

(3.4)
in the canonical model

\[
Z = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \\ Z_{31} & Z_{32} \end{pmatrix}
\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}
\sim N\left(\begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \\ 0 & 0 \end{pmatrix}, I \otimes \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)
\]

where \( r_3 = n - r \). This problem is clearly left invariant under the group \( G = O(n) \times B(p) \times F \) acting on \( Z \) and \( (n, \Sigma) \) by

\[
g(Z) = PZB + F \\
g(n, \Sigma) = (P_1 \ 0) nB + \begin{pmatrix} F_{11} \ 0 \\ F_{21} \ F_{22} \end{pmatrix}, B'E'B) \]

where

\[
g = (P, B, F) \in G
\]

\[
\tilde{O}(n) = \{ P = \begin{pmatrix} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ 0 & 0 & P_3 \end{pmatrix} | P_i \in O(r_i) \},
\]

\[
B(p) = \{ B = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} | B_i \in GL(p_i) \}
\]

\[
F = \{ F = \begin{pmatrix} F_{11} & 0 \\ F_{21} & F_{22} \\ 0 & 0 \end{pmatrix} | F_{ij} \in R \}
\]

It follows from (3.7) that the power function of an invariant test is a function of \((\xi_{12}, \xi_{12}, \Omega)\) where
\[ \xi_{12} = \eta_{12} \Sigma_{22}^{-1} \phi, \quad \Omega \equiv \Omega(\rho) = (\Delta' \Delta) \] and

\[ \Delta \equiv \Delta(\rho) = \text{diag}\{\rho_1, \ldots, \rho_t\} : p_1^T p_2 \]

with \( t = \min(p_1, p_2) \). Here \( \rho_1^2 \geq \ldots \geq \rho_t^2 \) are the characteristic roots of \( \Sigma_{11} \Sigma_{12} \Sigma_{21} \Sigma_{22} \) and \( \phi \) is an orthogonal matrix which diagonalizes \( \Sigma_{22} \Sigma_{21} \Sigma_{12} \Sigma_{11} \).

Hence without loss of generality we assume \( \eta_{12} = \xi_{12} \) and \( \Sigma = \Omega \). Further, it also follows from (3.6) that any invariant test is a function of \( (Z_{12}, U) \) with \( U = (Z_{31}, Z_{32}) \), on which \( G \) acts by

\[ g(Z_{12}, U) = (P_1 Z_{12} B_2, P_3 U B) \text{ for } g = (P, B, F) \in G. \] (3.10)

Now to state one of our main results in this section, let \( D^I_\alpha \) be the set of all invariant tests of size \( \bar{\alpha} \) (i.e., \( \phi \in D^I_\alpha \Leftrightarrow \phi(\xi) = \phi(\zeta) \)),

\[ S = U'U \equiv \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \quad S_{ij} = Z_{3i}^T Z_{3j} \] (3.11)

and let \( \delta = \delta_1 + \delta_2 \) with

\[ \begin{cases} \delta_1 = \text{tr} \xi_{12} \xi_{12}' = \text{tr} \eta_{12} \Sigma_{22}^{-1} \xi_{12}' \\ \delta_2 = \sum_{i=1}^t \rho_i^2 = \text{tr} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22} \Sigma_{21}^{-1}. \end{cases} \] (3.12)

Clearly \( \eta_{12} = 0 \) and \( \Sigma_{12} = 0 \) if and only if \( \delta = 0 \), or \( \delta_1 = 0 \) and \( \delta_2 = 0 \).

**Theorem 3.1.** There is \( \epsilon > 0 \) such that on the set \( \{(n, \xi) | \delta < \epsilon\} \), the power function of any test \( \phi \) in \( D^I_\alpha \) is evaluated as

\[ \pi(\phi, (n, \xi)) = a + \delta_1 C_1(\phi) + \delta_2 C_2(\phi) + o(\delta) \] (3.13)
where

\[ C_1(\phi) = \frac{r_1 + r_3}{2p_1p_2} \left\{ \frac{p_1}{r_1} E_0[\phi \text{tr} Z_{12}' Z_{12} (Z_{12}' Z_{12} + S_{22})^{-1}] \right\}, \quad (3.14) \]

\[ C_2(\phi) = \frac{r_1 + r_3}{2p_1p_2} E_0[\phi [r_3 \text{tr} S_{11}' S_{12} (S_{22} + Z_{12}' Z_{12})^{-1} S_{21}]
- p_1 \text{tr} S_{22} (S_{22} + Z_{12}' Z_{12})^{-1} \] - \frac{r_3}{2}, \quad (3.15) \]

\[ \lim_{\delta \to 0} \sup \phi |o(\delta) / \delta| = 0. \]

Here \( E_0(\cdot) \) denotes the expectation of \( \cdot \) under the null hypothesis. The proof is given at the end of this section. The expression (3.13) shows the local behavior of the power function of an invariant test \( \phi \), according to which the power function is approximated by \( \alpha + \delta_1 C_1(\phi) + \delta_2 C_2(\phi) \) in the neighborhood

\[ N_\varepsilon = \{(n, Z) | \text{tr} n_{12} Z_{12}' Z_{12}^{-1} + \text{tr} Z_{11}' Z_{11} Z_{12}^{-1} \varepsilon < \varepsilon \} \]

and the approximation is uniform in \( \phi \). Since a test maximizing the local power \( \alpha + \delta_1 C_1(\phi) + \delta_2 C_2(\phi) \) depends on at least the ratio of \( \delta_1 \) and \( \delta_2 \), no LBI test exists. This is natural in the sense that the quantities \( \delta_1 \) and \( \delta_2 \) indicate are different and the deviation of \( \delta_1 \) (or \( n_{12} \)) from 0 is independent of the deviation of \( \delta_2 \) (or \( \Sigma_{12} \)) from 0. On the contrary, the form of the power function in (3.13) reflects how an invariant test can detect each local deviation from each null case. That is, \( C_1(\phi) \) basically measures the local power of \( \phi \) against the local deviation of \( \delta_1 \) from 0, while \( C_2(\phi) \) the local power of \( \phi \) against the deviation of \( \delta_2 \) from 0.

However, the statistics which define the expectations of \( C_1(\phi) \) and \( C_2(\phi) \)
in (3.14) and (3.15), are not independent even under the null hypothesis. In fact, the statistics
\[
\begin{align*}
T_1 &= \frac{p_1}{r_1} \text{tr} Z_{12} Z_{12}' (Z_{12} Z_{12}' + S_{22})^{-1} \\
T_2 &= r_3 \text{tr} S_{12}^{-1} S_{12}' (S_{22} + Z_{12} Z_{12}')^{-1} S_{21}' - p_1 \text{tr} S_{22}' (S_{22} + Z_{12} Z_{12}')^{-1}
\end{align*}
\]
are dependent on each other. This is a feature of the simultaneous treatment of the two separate hypotheses. To investigate this point further, observe that the test, say \( \psi_1 \), which maximizes \( C_1(\phi) \) is given by the critical region
\[ T_1 > c \]
and it is the LBI test for testing \( \eta_{12} = 0 \) without \( \Sigma_{12} = 0 \). More specifically it is LBI for testing the General MANOVA hypothesis
\[ H': \tilde{M}_1 \otimes \tilde{M}_2 = 0 \tag{3.17} \]
in the MANOVA model \( Y = \mathbf{X}\Theta + \mathbf{E} \) with \( \mathbf{E} \sim N(0, I \otimes \Sigma) \) where \( \mathbf{X}: n \times r, \tilde{M}_1: r_1 \times r \) and \( \tilde{M}_2: p \times p_2 \) are arbitrarily fixed matrices of full rank. This test is even UMPI (uniformly most powerful invariant) when \( r_1 = 1 \) or \( p_2 = 1 \) and the power function of an invariant test \( \psi \) for the hypothesis in (3.17) is locally expressed as
\[ \pi_1(\psi, (\eta, \Sigma)) = \alpha + \delta_1 C_1(\psi) + o(\delta_1) \tag{3.18} \]
where \( (\eta, \Sigma) \) is the parameter of a canonical form corresponding to \((\Theta, \Sigma)\) (see Kariya (1985), p 109). On the other hand, the test, say \( \psi_2 \), which maximizes \( C_2(\phi) \) is given by the critical region
\[ T_2 > c \]
with $T_2$ in (3.16) and it is the LBI test for testing $\Sigma_{12} = 0$ in the case of $\eta_{12} = 0$. More specifically it is LBI for testing independence $\Sigma_{12} = 0$ in the missing data model

$$Z_{12} \sim N(\eta_{12}'I_{r_1} \otimes \Sigma_{22}) \text{ with } \eta_{12} = 0$$

(3.19)

$U \sim N(0,I_{r_3} \otimes \Sigma)$ and $Z_{12}$ and $U$ are independent

where the counterpart of $Z_{12}$ is missing (see Eaton and Kariya (1983)) and the power function of any invariant test in the model (3.19) is expressed as

$$\pi_2(\psi, (0, \Sigma)) = \alpha + \delta_2 C_2(\psi) + o(\delta_2) \quad (3.20)$$

However, when $\eta_{12} \neq 0$, the test based on the critical region $T_2 > c$ is not easy to interpret as a test for testing the single hypothesis of the independence $\Sigma_{12} = 0$ alone because in the case of $\eta_{12} \neq 0$, $Z_{12}$ would not be involved in a test statistic. In other words, this is a difference between treating the two hypotheses simultaneously and treating them separately.

Now by taking this point into account, we propose the test maximizing a linear combination of $C_1(\phi)$ and $C_2(\phi)$;

$$C_B(\phi) = \beta r_3 C_1(\phi) + C_2(\phi) \quad (0 < \beta < \infty) \quad (3.21)$$

where $\beta$ is a constant independent of $n$. Using the Generalized Neyman-Pearson Lemma, the critical region is given by

$$T(\beta) = \beta r_3 T_1 + T_2 > k. \quad (3.22)$$

Here $T_1$ is multiplied by $r_3 = n-r$ because from (3.3) $n = O(n^{\frac{1}{2}})$ so that $\delta_1 = O(n)$ provided $X'X = O(n)$. The test $\phi_B$ with critical region $T(\beta) > c$ maximizes the power $\pi(\phi, (\eta, \Sigma))$ locally in the neighborhood
\[ N_{eB} = \{(n, \Sigma) \mid \delta_1 = \beta r_3 \delta_2\} \cap N_e \]

since it maximizes \((\delta_1 / r_3) \beta r_3 C_1(\phi) + \delta_2 C_2(\phi) = \delta_2 [\beta r_3 C_1(\phi) + C_2(\phi)] \) on \(N_{eB}\).

The constant \(\beta\) may be regarded as a weight for the importance of the hypothesis \(\eta_{12} = 0\) relative to the hypothesis \(\Sigma_{12} = 0\), and it is chosen in advance. It is noted that the test based on \(T(\beta)\) is not a linear combination of the two LBI tests \(\Psi_1\) and \(\Psi_2\) stated above.

We remark that for a given \(\phi \in D_\alpha^I\), the local sensitivity of \(\phi\) against \((\delta_1, \delta_2)\) is measured by the two coordinate \((C_1(\phi), C_2(\phi))\). Second, when the information \(\eta_{12} = 0\) is ignored in the missing data model (3.22), the LBI test for independence \(\Sigma_{12} = 0\) is given by the critical region

\[ T_3 = \text{tr} S_{11}^{-1} S_{12} S_{22}^{-1} S_{21} > k \quad (3.23) \]

(see Schwartz (1967)). That is, this is the LBI test for independence in the MANOVA model without \(\eta_{12} = 0\). Since the test with critical region \(T_1 > k\) is LBI for testing \(\eta_{12} = 0\) without \(\Sigma_{12} = 0\), we may combine these two tests. Here while the simultaneous treatment of the hypotheses yielded the dependent statistics \(T_1\) and \(T_2\), the separate treatment yields the independent test statistics \(T_1\) and \(T_3\).

The problem of how to combine independent test statistics is discussed in the literature (see, e.g., Marden (1982)). In this paper we do not get involved in this problem. But as is shown next, the LRT for the simultaneous hypotheses gives a natural combination of the LRT statistics for two separate hypotheses.

The LRT for our problem is easily obtained by using the canonical model in (3.5) as follows:

\[ \frac{|S|}{|S_{11}| S_{22} + Z_{12} Z_{12}^* |} = \frac{|S|}{|S_{11}| S_{22}|} \cdot \frac{|S_{22}|}{|S_{22} + Z_{12}^* Z_{12}|} = L_1 L_2 \quad (3.24) \]
Of course, when $L_1 L_2$ is small, the joint hypothesis is rejected. As is well known, $L_1$ is the LRT statistic for $\Sigma_{12} = 0$ without $\eta_{12} = 0$ while $L_2$ is the LRT statistic for $\eta_{12} = 0$ without $\Sigma_{12} = 0$. Further $L_1$ and $L_2$ are independent so that the two independent LRT statistics are combined in (3.24). It is noted that even when $\eta_{12} = 0$, the LRT statistic for $\Sigma_{12} = 0$ is given by $L_1$ so that the LRT ignores the additional data $Z_{12}$ (see Eaton and Kariya (1984)). The unbiasedness property of the power function is considered as a special case of Problem (II) in the next section.

Proof of Theorem 3.1. The proof is similar to KFK (1984). To derive the distribution of a maximal invariant $T = T = T(Z_{12}, U)$ under the action (3.10) of $G$ on $(Z_{12}, U)$, we apply the Wijsman's representation theorem. Let $p_T^{T}(\xi, \Omega)$ be the distribution of a maximal invariant $T$. Then the density of $T$ with respect to $p_T^{(0, I)}$ evaluated at $T = T(Z_{12}, U)$ is given by

$$R = \frac{dP_T^{T}(\xi_{12}, \Omega)}{dP_T^{(0, I)}} = H(Z_{12}, U|\xi_{12}, \Omega)/H(Z_{12}, U|0, I) \quad (3.25)$$

where

$$H(Z_{12}, U|\xi_{12}, \Omega) = \int_H f(p_1 Z_{12}^B 22, p_3 U|\xi_{12}, \Omega) X_B(\nu_1(dP_1)\nu_3(dP_3)\mu_1(dB_1)\mu_2(dB_2))$$

$$X_B = \begin{vmatrix} r_3/2 & (r_1 + r_3)/2 \\ B_1^T B_1 & B_2^T B_2 \end{vmatrix}, \quad \nu_1(dB_j) = |B_j B_j|^{-p_j/2} dB_j$$

$H = O(r_1) \times O(r_3) \times GL(p_1) \times GL(p_2)$, $f(Z_{12}, U|\xi_{12}, \Omega)$ is the density of $Z_{12}$ and $U$ with $Z_{12} \sim N(\xi_{12}, I \otimes I)$ and $U \sim N(0, I \otimes \Omega)$, and $\nu_1(dP_i)$ is the invariant probability measure on $O(r_i) \ (i=1, 3; j=1, 2)$. The condition for which (3.25) holds is satisfied (see Wijsman (1967)). In order to obtain the local behavior of the power function of an invariant test in a neighborhood of $\xi_{12} = 0$ and $\Sigma_{12} = 0$, we evaluate $R$ in (3.25) locally. After cancellation of some constants the
numerator in (3.25) is expressed as

\[ H(Z_{12}, U | \xi_{12} , \Omega ) = \int_{0(r_1) \times GL(p_1) \times GL(p_2)} |\Omega|^{-3/2} \exp[-\frac{1}{2}K] \chi(B) \nu_1(dP_1) \nu_1(dB_1) \nu_2(dB_2) \]

where

\[ K = \text{tr}B_1B_2Z_{12}^\prime B_2' - 2\text{tr}Z_{12}B_2^\prime \xi_{12} + \text{tr}Z_{12}^\prime \xi_{12} \]

\[ + \text{tr}B_1S_{11}B_1F_{11} + 2\text{tr}B_1S_{12}^\prime B_2'F_{12} + \text{tr}B_2S_{22}B_2F_{22} \]

with

\[ \Omega^{-1} = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \]

(3.27)

Here replacing \( B_1 \) by \( S_{11}^\prime B_1 \) and \( B_2 \) by \( (Z_{12}^\prime Z_{12} + S_{22})^{-1} B_2 \) leaves the ratio \( R \) remain the same since the Jacobin factors coming out are cancelled out with those of the denominator. Further writing \( F_{11} \) as

\[ F_{11} = (I - \Delta_{11})^{-1} = I + \Delta_{11} + \Delta_{11}^2 (I - \Delta_{11})^{-1} \]

with \( \Delta_{11} = \Delta_{11}^\prime \) and \( \Delta_{22} = \Delta_2^\prime \Delta \). Then \( K \) in (3.26) becomes

\[ K = \text{tr}B_1B_1' + \text{tr}B_2B_2' - 2\text{tr}Z_{12}B_2^\prime \xi_{12} + \text{tr}Z_{12}^\prime \xi_{12} + 2\text{tr}B_1B_2S_{11}F_{12} + K_1 + K_2 \]

where

\[ K_1 = \text{tr}B_1B_1[\Delta_1 + \Delta_1^2 (I - \Delta_1)^{-1}]; \ K_2 = \text{tr}B_2S_{22}B_2[\Delta_2 + \Delta_2^2 (I - \Delta_2)^{-1}] \]

\[ \tilde{Z}_{12} = Z_{12}(Z_{12}^\prime Z_{12} + S_{22})^{-1/2}, \ \tilde{S}_{12} = S_{11}^{-1/2} (Z_{12}^\prime Z_{12} + S_{22})^{-1/2} \] and
Now expand the components of $\exp[-\frac{1}{2}K]$ as

$$\exp[\text{tr}P_1^*Z_{12}B_{12}F'] = 1 + \text{tr}P_1^*Z_{12}B_{12}F' + \frac{1}{2}(\text{tr}P_1^*Z_{12}B_{12}F')^2 + o_1$$

$$\exp[-\text{tr}B_1^*S_{12}B_{12}F'] = 1 - \text{tr}B_1^*S_{12}B_{12}F' + \frac{1}{2}(\text{tr}B_1^*S_{12}B_{12}F')^2 + o_2$$

$$\exp[-\frac{1}{2}K_1] = 1 - \frac{1}{2}\text{tr}B_1^*B_{11} + o_3 \quad \text{and}$$

$$\exp[-\frac{1}{2}K_2] = 1 - \frac{1}{2}\text{tr}B_2^*S_{22}B_{22} + o_4.$$ 

Then in the same way as in KFK(1984), the remainder terms $o_i$'s are shown to be $o(\delta)$ with $\delta$ in (3.12), which is uniform in $(Z_{12}, S)$ because of the boundedness of $Z_{12}$ and $S_{ij}$'s. Further when $o_i$'s are integrated over $P_1$ and $B_i$'s, they are shown to be $o(\delta)$ uniformly in $(Z_{12}, S)$. Now taking the product of these terms and ignoring the odd functions of either $P_1$ or $B_1$ or $B_2$ because they are zero when integrated, we obtain

$$R = \int S\Omega^{-3/2} \nu_1(dP_1)h_1(B_1)h_2(B_2)dB_1dB_2$$

where

$$h_1(B_1) = \exp(-\frac{1}{2}\text{tr}B_1^*B_1)|B_1B_1|^M_1/2dB_1/D_1$$

$$J = 1 + \frac{1}{4}[(\text{tr}P_1^*Z_{12}B_{12}F')^2 + (\text{tr}B_1^*S_{12}B_{12}F')^2 - \text{tr}B_1^*B_{11}^2]$$

$$- \text{tr}B_2^*S_{22}B_{22}^2 + o_5 = 1 + \frac{1}{4}[I + II + III + IV] + o_5,$$ say

$$D_1 = \int \exp(-\frac{1}{2}\text{tr}B_1^*B_1)|B_1B_1|^M_1/2dB_1,$$
$M_1 = r_3 - p_1$ and $M_2 = r_1 + r_3 - p_2$. Here $H(Z_k, U(0, I)) = D_1D_2$ and $|\Omega|^{-k_3/2} = 1 + o(\delta)$ were used. The remainder $O_5$ is shown to be $o(\delta)$ for $\delta$ small uniformly in $(Z_{12}, S)$, for $F_{12} = (I - \Delta')^{-1}\Delta$. We now need to integrate $J$. But arguing as in KFK (1984) pp 391-392, we obtain

$I = \int (\text{tr}P_1\tilde{z}_{12}Bz_{12})^2v_1(dP_1)h_2(B_2)dB_2 = \frac{r_1 + r_3}{r_1p_2} \text{tr} \tilde{z}_{12}^2 \text{tr}z_{12}^2$,

$II = \int (\text{tr}B_1S_{12}B_2F_{12})^2h_1(B_1)h_2(B_2)dB_1dB_2 = \frac{r_3(r_1 + r_3)}{p_1p_2} \text{tr}S_{12}^2 \text{tr}S_{12}^2 \text{tr}\Delta' + o(\delta)$,

$III = \int (\text{tr}B_1\Delta' B_1)h_2(B_1)dB_1 = \frac{r_3p_1}{p_1} \text{tr}\Delta'$,

$IV = \int (\text{tr}B_2\Delta' B_2)h_2(B_2)dB_2 = \frac{p_1 + r_3}{p_2} \text{tr}S_{22} \text{tr}\Delta'$,

and $\int O_5 1(dP_1)h_1(B_1)h_2(B_2)dB_1dB_2 = o(\delta)$. Thus, observing that the power function of an invariant test $\phi$ is given by

$$\pi(\phi, (n, \Sigma)) = \int \phi dP(\xi_{12}) = \int \phi RdP(0, I),$$

the result in (3.13) is finally obtained, completing the proof.
4. TESTS FOR THE HYPOTHESIS $\theta_{12}=0$, $\theta_{21}=0$, AND $\Sigma_{12}=0$

We shall first consider via invariance the Problem (II):

$$H: \theta_{12} = 0, \theta_{21} = 0, \Sigma_{12} = 0 \text{ vs } K: \text{not } H.$$  \hspace{1cm} (4.1)

Since $\theta_{12} = 0$ and $\theta_{21} = 0$ are expressed as

$$M_1 \theta M_2 = 0 \text{ and } M_3 \theta M_4 = 0$$

and since $M_1 X' X M_3 = X_1 X_2 \neq 0 \text{ in general}$, the group leaving this problem invariant is smaller than the group leaving the problem (I) invariant, where $M_1$ and $M_2$ are defined by (3.2), $M_3 = (0, I)$ and $M_4 = (I)$. The special case $X_1' X_2 = 0$ will be briefly treated later. Here we use the following canonical form

$$\begin{align*}
W &= K(X' X)^{-1} X' Y - N(n, A \otimes \Sigma) \text{ with } n = K \theta \\
S &= Y' (I - P_0) Y - W(\Sigma, r_3) \text{ with } r_3 = n - r \\
W \text{ and } S \text{ are independent}
\end{align*}$$  \hspace{1cm} (4.2)

where $P_0 = X(X' X)^{-1} X'$,

$$K = \begin{pmatrix} Q_{-1} & 0 \\ 0 & Q_{-2} \end{pmatrix} \text{ with } (X' X)^{-1} = Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$  \hspace{1cm} (4.3)

and

$$A = \begin{pmatrix} I & \bar{Q}_{12} \\ \bar{Q}_{12} & I \end{pmatrix} \text{ with } \bar{Q}_{12} = Q_{-2} Q_{12} Q_{-1}.$$  \hspace{1cm} (4.4)

Partition $W$ and $n$ as
respectively. Then in the model (4.2) the problem is to test

\[ H: \eta_{12} = 0, \eta_{21} = 0, \Sigma_{12} = 0. \]

The problem is clearly left invariant under the group \( G = GL(p_1) \times GL(p_2) \times \mathbb{R}^r_1 \times \mathbb{R}^r_2 \)
acting on \((W,S)\) by

\[ g(W,S) = (WB+F_B'SB) \text{ with } B = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} \text{ and } F = \begin{pmatrix} F_1 & 0 \\ 0 & F_2 \end{pmatrix} \] (4.5)

where \( g = (B_1,B_2,F_1,F_2) \in G \). From this action of \( G \), it is easy to see that an
invariant test is a function of \((W_{12},W_{21},S)\) only and that the power function of
an invariant test is a function of

\[ \Omega = \begin{pmatrix} I & \Delta' \\ \Delta & I \end{pmatrix}, \quad \xi_{12} = Q_{11}^{-\frac{1}{2}} \eta_{12} \Sigma_{22}^{-\frac{1}{2}} p_2 \text{ and } \xi_{21} = Q_{22}^{-\frac{1}{2}} \eta_{21} \Sigma_{11}^{-\frac{1}{2}} p_1 \] (4.6)

where \( p_i \in \alpha(p_i)'s \) \((i=1,2)\) satisfy

\[ p_1 \Sigma_{11}^{-\frac{1}{2}} 12 r_2 \Sigma_{22}^{-\frac{1}{2}} p_2 = \Delta = \text{diag}(\rho_1,\ldots,\rho_t,0,\ldots,0): p_1 \times p_2 \text{ with } t = \min(p_1,p_2) \] (4.7)

and \( \rho_1 \geq \rho_t \) are the canonical correlations. So we can assume \( \eta_{12} = \xi_{12}, \)
\( \eta_{21} = \xi_{21} \) and \( \Sigma = \Omega \) without loss of generality. In this set-up, we may proceed
in the same way as we did in Section 3 to obtain an expression for the locally
approximate power of an invariant test. However by doing so we end up with a
very complicated expression of local power which depends on many parameters in an
intractable way. This implies that a further invariance consideration into the
problem will not help to propose a new test which possibly takes into account the
simultaneous occurrence of the three hypotheses \( \eta_{12} = 0, \eta_{21} = 0, \) and \( \Sigma_{12} = 0. \) Of course no LBI test exists.

On the other hand, the LRT is given by the critical region \( L < c, \) which was originally derived by Fujikoshi (1982), where

\[
L = \frac{|S|}{|S_{22} + W_{12} W_{12}||S_{11} + W_{21} W_{21}|} = \frac{|S|}{|S_{22}|} \cdot \frac{|S_{11}|}{|S_{11} + W_{21} W_{21}|} = L_1 L_2 L_3, \text{ say.}
\]

(4.8)

It is also directly obtained from the distribution of \((W_{12}, W_{21}, S)\) since the LRT is always invariant (under a very mild condition). The statistics \( L_1, L_2 \) and \( L_3 \) in (4.8) are respectively the LRT statistics for the three separate hypotheses \( \Sigma_{12} = 0, \theta_{12} = 0, \) and \( \theta_{21} = 0, \) and the LRT statistic \( L_1 L_2 L_3 \) for our simultaneous hypothesis in (4.8) may be viewed as showing how to combine the three LRT statistics for the three separate hypotheses. But here it is noted that \( L_2 \) and \( L_3 \) are correlated because of the correlation of \( W_{12} \) and \( W_{21} \) unless \( X_1^T X_2 = 0 \) or \( \Sigma_{12} = 0. \) In fact, if \( u_i^T \) and \( v_j^T \) are respectively the \( i-th \) and \( j-th \) rows of \( W_{12} \) and \( W_{21} \), the covariance matrix of \( u_i \) and \( v_j \) is shown to be

\[
\text{Cov}(u_i, v_j) = q_{ij} \Delta^2: p_2 \times p_1
\]

where \( q_{ij} \) is the \((i,j)th \) element of \( \bar{\Omega}_{12} \) in (4.4) and \( \bar{\Omega}_{12} \approx 0 \) and \( \Delta = 0 \) are respectively equivalent to \( X_1^T X_2 = 0 \) and \( \Sigma_{12} = 0. \) This correlation between \( L_2 \) and \( L_3 \) makes it difficult to investigate optimality properties of the LRT. Now to show the unbiasedness of the LRT, note that from \( W_{12} \sim N(\xi_{12}, I \otimes I) \) and \( W_{21} \sim N(\xi_{21}, I \otimes I) \)

\[
\begin{align*}
V_{22} &= W_{12} W_{12} - W_{p_2}(I, \tilde{r}_2; \tau_2) \text{ with } \tilde{r}_2 = r_1 \text{ and } \tau_2 = \xi_{12} \xi_{12} \\
V_{11} &= W_{21} W_{21} - W_{p_1}(I, \tilde{r}_1; \tau_1) \text{ with } \tilde{r}_1 = r_2 \text{ and } \tau_1 = \xi_{21} \xi_{21}
\end{align*}
\]

(4.9)
where \( W(m,v) \) denotes the noncentral Wishart distribution with mean \( m \), degrees of freedom \( m \) and noncentrality parameter \( v \). In the original term, \( V_{11} \) is easily shown to be equal to

\[
V_{11} = Y_1' [P_0 - P_1] Y_1 \quad \text{with} \quad P_1 = X_1' (X_1 X_1)^{-1} X_1'.
\]  

(4.10)

Further, from (4.2), \( S \) and \( \{V_{11}, V_{22}\} \) are independent, and when \( \Delta = 0 \), \( V_{11} \) and \( V_{22} \) are independent. Here we use a conditional argument. First write \( L_1 L_2 \) in (4.8) as

\[
L_1 L_2 = |S_{22.1}| / |S_{22.1} + V_{22} + S_{21} S_{11}^{-1} S_{12}| = |\tilde{S}_{22.1}| / |\tilde{S}_{22.1} + \tilde{U}_{22}|.
\]  

(4.11)

where

\[
S_{22.1} = S_{22} - S_{21} S_{11}^{-1} S_{12}, \quad \tilde{S}_{22.1} = \gamma^{-k} S_{22.1} \gamma^{-k}
\]

\[
\tilde{U}_{22} = \gamma^{-k} (V_{22} + S_{21} S_{11}^{-1} S_{12}) \gamma^{-k} \quad \text{and} \quad \gamma = I - \Delta' \Delta.
\]

Lemma 4.1

(1) \( S_{22.1} \sim W_p(I, n-r-p_1) \)

(2) \( \tilde{S}_{22.1} \) and \( \tilde{U}_{22} \) are independent

(3) Conditional on \( Y_1 \), \( \tilde{U}_{22} \sim W_p(I, r_1 + p_1 : \psi) \) where

\[
\psi = \gamma^{-k} [\tilde{\Sigma} + \Delta' S_{11} \Delta] \gamma^{-k}
\]

\[
\tilde{\Sigma} = [X_1 \tilde{Y}_{12} + Y_1 \Delta]' (P_0 - P_2) [X_1 \tilde{Y}_{12} + Y_1 \Delta] \quad \text{with} \quad \tilde{\theta}_{12} = \theta_{12} \tilde{\Sigma}^{-1/2} P_2.
\]

Proof. Our original model may be viewed as

\[
[Y_1, Y_2] \sim N([X_1, X_2] \begin{bmatrix} 0 & \tilde{\theta}_{12} \\ \tilde{\theta}_{21} & 0 \end{bmatrix}, I_n \otimes \Omega)
\]

with \( \Omega \) in (4.6), where \( \tilde{\theta}_{21} = \theta_{21} \tilde{\Sigma}^{-1/2} P_1 \). Hence conditional on \( Y_1 \),
\[ Y_2 \sim N(X_1 \bar{Y}_{12} + (Y_1 - X_2 \bar{Y}_{21}) \Delta, I \otimes \Gamma), \]

from which conditional on \( Y_1 \), we obtain

\[ S_{21} S_{12}^{-1} \sim Y_2^T p_2 Y_2 \sim \text{W}_p (\gamma, p_1 : \Delta, S_{11} \Delta), \]

\[ V_{22} = Y_2^T [p_0 - p_1] Y_1 \sim \text{W}_p (\gamma, r_1 : \Xi) \text{ and} \]

\[ S_{22,1} \sim \text{W}_p (\gamma, n - r - p_1). \]

Further conditional on \( Y_1 \), \( S_{22,1}, S_{21} S_{12}^{-1} \) and \( V_{22} \) are independent and \( S_{22,1} \) does not depend on \( Y_1 \). Thus all the result follow.

**Theorem 4.1.** The LRT with critical region \( L < c \) is unbiased.

**Proof.** Let

\[ C = \{(S_{22,1}, V_{22} + S_{21} S_{12}^{-1} S_{11}, V_{11}, S_{11}) : L_1 L_2 L_3 < c) \}

be the critical region of size \( \alpha \) in the space of \((S_{22,1}, V_{22} + S_{21} S_{12}^{-1} S_{11}, V_{11}, S_{11})\). Then since \( S_{11} \) and \( V_{11} \) are functions of \( Y_1 \) only, from (4.10) the power function of the LRT \( \phi_L \) is expressed as

\[ \pi(\phi_L, (\bar{Y}_{12}, \bar{Y}_{21}, \Delta)) = P(C | \bar{Y}_{12}, \bar{Y}_{21}, \Delta) \]

\[ = E_{Y_1} [P((S_{22,1} \bar{Y}_{22}) | S_{22,1}, \bar{Y}_{22} < c L_3^{-1}) | (S_{11}, V_{11}, \bar{Y}_{12}, \bar{Y}_{21}, \Delta)] \]

(4.13)

where \( E_{Y_1} \) denotes the expectation with respect to \( Y_1 \). Since

\[ |\bar{S}_{22,1}/|\bar{S}_{22,1} + \bar{Y}_{22}| < c \]
is regarded as the LRT for testing $\psi = 0$ in the MANOVA set-up with $(\tilde{U}_{22}, \tilde{S}_{22}, \Gamma)$, as is shown in Anderson et al. (1964), it is an increasing function of each characteristic root of $\psi$. Since $\tilde{e}_{12} = 0$ and $\Delta = 0$ imply $\psi = 0$, from (4.13)

$$
\pi(\phi_L, (\tilde{e}_{12}, \tilde{e}_{21}, \Delta)) \geq \pi(\phi_L, (0, \tilde{e}_{21}, 0)) = P(C|0, \tilde{e}_{21}, 0). \quad (4.14)
$$

But under $\Delta = 0$, $L_1L_2$ and $L_3$ are independent because $V_{11}$ and $V_{22}$ are independent. Further $L_3 = |S_{11}|/(S_{11} + V_{11})$ is regarded as the LRT statistic for testing $\tilde{e}_{21} = 0$ in the MANOVA set-up with $(V_{11}, S_{11})$. Hence the inside of the conditional expectation

$$
P(C|0, \tilde{e}_{21}, 0) = E_{L_1L_2} [P((S_{11}, V_{11})|L_3 < c(L_1L_2)^{-1})|0, \tilde{e}_{21}, 0)]
$$

is an increasing function of each characteristic root of $\tilde{e}_{21}$ implies

$$
P(C|0, \tilde{e}_{21}, 0) \geq P(C|0, 0) \geq \alpha.
$$

Combining this with (4.13) yields the result.

Under $\Delta \neq 0$ or equivalently $\Sigma_{12} \neq 0$, because $L_1L_2$ and $L_3$ are correlated unless $X_1X_2 = 0$, it seems difficult not only to establish a monotonicity property of the power function of the LRT but also to find a natural parameter space on which the monotonicity is considered.

We remark that the above result holds even when the model $Y = X\Theta + \epsilon$ is defined for $X$ conditioned and the marginal distribution of $X$ does not depend on $(\Theta, \Sigma)$. Hence the LRT for testing no additional information hypothesis in canonical correlation model, which is nothing but our LRT though $X$ is random, is unbiased (see Section 2).

In the case of $X_1^T X_2 \neq 0$, we consider two special cases. First consider
the case $\Delta = 0$ or equivalently $\Sigma_{12} = 0$. In this case the problem of testing joint hypothesis $\theta_{12} = 0$ and $\theta_{21} = 0$ is simply split into two independent problems in the two independent models $Y_i = X_i \theta_{11} + X_i \theta_{21} + E_i (i=1,2)$. However, the LRT for the joint hypothesis is given by $\phi_L = X(L_2 L_3 < c)$ and it is not the product of the two LRT $X(L_2 < c_2) X(L_3 < c_3)$ where $X(A)$ denotes the indicator function of a set $A$. The power function of $\phi_L$ for $\theta_{12} = 0$ and $\theta_{21} = 0$ is a function of the characteristic roots.

$$\lambda_i = ch_i (\Sigma_{22}^{-1} X_j [P_0 - P_1] X_i \theta_{12} \Sigma_{22}^{-1} ) (i=1,...,p_2)$$

and the characteristic roots

$$\gamma_j = ch_j (\Sigma_{11}^{-1} X_j [P_0 - P_1] X_i \theta_{21} \Sigma_{11}^{-1} ) (j=1,...,p_1),$$

and by similar argument as in the proof of Theorem 4.1, we obtain

**Corollary 4.1.** Under $\Sigma_{12} = 0$, the power function of the LRT with critical region $L_2 L_3 < c$ is increasing in each $\lambda_i$ or $\gamma_j$.

Next we consider the unbiasedness of the LRT with critical region $L_1 L_2 < c$. Since the distribution property of $L_1 L_2$ is simply obtained in the proof of Theorem 4.1 by setting $V_{11} = 0$, we obtain

**Corollary 4.2.** The LRT with critical region $L_1 L_2 < c$ for Problem [I] is unbiased.

In fact, from (4.10) and Lemma 4.1, the power function of the LRT is easily seen to be an increasing function of each characteristic root of $\psi$ conditional on $Y$. Hence the unbiasedness immediately follows from $\psi > 0$ and the fact that $\theta_{12} = 0$ and $\Sigma_{12} = 0$ imply $\psi = 0$.

Finally we consider Problem [II] under $X_1^t X_2 = 0$. An example for this case
is found in time series models (see Anderson (1971, p. 92)). Under \( X_1'X_2 = 0 \), we can take \( A = I \) in the canonical form in (4.2) as well as \( Q_{11} = (X_1'X_1)^{-1} \). Then the group \( \tilde{G} = O(r) \times O(r_2) \times G \) leaves the problem invariant by the action

\[
\tilde{g}(W, S) = (\Gamma WB + F, S)B'
\]

where \( \tilde{g} = (P_1, P_2, g) \in G \). Hence in the same way as above, the problem is reduced to the problem of testing \( \xi_{12} = 0 \), \( \xi_{21} = 0 \) and \( \Delta = 0 \) in the canonical form

\[
\begin{align*}
W_{12} &\sim N(\xi_{12}, I \otimes I) \text{ with } \xi_{12} = Q_{11}^{-1} \Sigma_{12}^{-1} \Sigma_{22}^{-1} \\
W_{21} &\sim N(\xi_{21}, I \otimes I) \text{ with } \xi_{21} = Q_{22}^{-1} \Sigma_{21}^{-1} \Sigma_{11}^{-1} \\
S &\sim W(\omega, r_3) \text{ with } \omega \text{ in (4.6)}
\end{align*}
\]

where \( W_{12}, W_{21} \) and \( S \) here are independent. Also \( \tilde{G} \) acts on \( (W_{12}, W_{21}, S) \) by

\[
\tilde{g}(W_{12}, W_{21}, S) = (P_1W_{12}B_2, P_2W_{21}B_1, S)B'.
\]

Since the group \( \tilde{G} \) is bigger than \( G \), a result corresponding to Theorem 4.1 can be derived. To see this, let \( \nu = \nu_1 + \nu_2 + \nu_3 \) with

\[
\begin{align*}
\nu_1 &= \text{tr} \xi_{12}' \Sigma_{12}^{-1} = \text{tr} X_1'X_1 \theta_{12}^{-1} \\
\nu_2 &= \text{tr} \xi_{21}' \Sigma_{21}^{-1} = \text{tr} X_2'X_2 \theta_{21}^{-1} \\
\nu_3 &= \text{tr} \Delta = \text{tr} \Delta' = \text{tr} \Sigma_{11}^{-1} \Sigma_{12}^{-1} \Sigma_{22}^{-1}
\end{align*}
\]

**Theorem 4.2.** There is an \( \epsilon > 0 \) such that on the set \( \{(\theta, \Sigma | \nu < \epsilon)\} \) the power function of an invariant test \( \phi \) of size \( \alpha \) under \( \tilde{G} \) is evaluated as...
\[ \pi(\phi, (\Theta, \Sigma)) = \alpha + \frac{1}{2}(v_1D_1(\phi) + v_2D_2(\phi) + v_3D_3(\phi)) + o(v) \]

where

\[ D_i(\phi) = E_0[\phi R_i] \]

\[ R_1 = \frac{r_1^* + r_3^*}{p_1 p_2} \frac{p_1}{r_1^*} \text{tr} W_{12} W_{12} (W_{12} W_{12} + S_{22})^{-1} \]

\[ R_2 = \frac{r_2^* + r_3^*}{p_1 p_2} \frac{p_2}{r_2^*} \text{tr} W_{21} W_{21} (W_{21} W_{21} + S_{11})^{-1} \]

\[ R_3 = \frac{(r_1^* + r_3^*) (r_2^* + r_3^*)}{p_1 p_2} \text{tr} (W_{21} W_{21} + S_{11})^{-1} S_{12} (W_{12} W_{12} + S_{22})^{-1} S_{21} \]

\[ \quad - \frac{r_1^* + r_3^*}{p_1} \text{tr} S_{22} (W_{12} W_{12} + S_{22})^{-1} \quad - \frac{r_2^* + r_3^*}{p_2} \text{tr} S_{11} (W_{21} W_{21} + S_{11})^{-1} \]

and \( \lim \sup_{v \to 0} |o(v)/v| = \ldots \)

**Proof.** The proof is completely similar to that of Theorem 4.1. Regarding \( Z_{12} \) as \( W_{12} \), replace \( B_1 \) by \( W_{21} (W_{21} W_{21} + S_{11})^{-1} B_1 \). Then every step goes through and the result is obtained.

In this expression, a symmetry which is lacking in Theorem 4.1 is secured, and the statistics \( R_1, R_2 \) and \( R_3 \) are respectively the LBI tests for testing the separate hypotheses \( H_1: \xi_{12} = 0, H_2: \xi_{21} = 0 \) and \( H_3: \Sigma_{12} = 0 \) under \( \xi_{12} = 0 \) and \( \xi_{21} = 0 \) (see Eaton and Kariya (1983) for \( H_3 \)). Also for each hypothesis \( H_i \), the power function of an invariant test is expressed as \( \alpha + \frac{1}{2} v_i D_i(\psi) + o(v_i)(i=1,2,3) \).

Here again it is noted that the statistics \( R_1, R_2 \) and \( R_3 \) are not independent under the null hypothesis. \( R_1 \) and \( R_2 \) are independent under the null hypothesis. Here following the discussion in Section 3, we may propose a combined test of these statistics with critical region.
\[ R(\theta_1, \theta_2) = \beta_1 r_3 R_1 + \beta_2 r_3 R_2 + R_3 > c \]

(4.15)

where \( r_3 \) is put on \( R_1 \) 's because \( \nu_i = O(n) \) (i=1,2). Here the constants \( \beta_i \)'s may be considered indicating relative weights for the three hypotheses \( \theta_{12} = 0, \theta_{21} = 0 \) and \( \Sigma_{12} = 0 \). It will be often the case \( \beta_1 = \beta_2 \). This test clearly maximizes the local power \( \alpha + \sum_{i=1}^{3} \nu_i D_i(\phi) \) in the neighborhood

\[ \{(\theta, \Sigma) | \nu_i = r_3 \beta_i \nu_3 (i=1,2) \} \cap \{(\theta, \Sigma) | \nu < \epsilon \} \]

Also the local sensitivity of an invariant test \( \phi \) against \( (\nu_1, \nu_2, \nu_3) \) is described by the coordinates \( (D_1(\phi), D_2(\phi), D_3(\phi)) \). Further, from the observations above, we may use the test based on \( R(\theta_1, \theta_2) \) in (4.15) even if \( X_1 X_2 \neq 0 \).

It is remarked that the LRT statistic in this case is of course the same as \( L = L_1 L_2 L_3 \). But here because \( X_1 X_2 = 0 \), in addition to the independence of \( L_1 \) and \( L_i (i=2,3) \), \( L_2 \) and \( L_3 \) are independent though \( L_1 \), \( L_2 \) and \( L_3 \) are jointly dependent.
5. ASYMPTOTIC NULL DISTRIBUTIONS OF THE TEST STATISTICS

The asymptotic null distribution of the LRT for Problem [II]: \( \theta_{12} = 0, \theta_{21} = 0 \) and \( \Sigma_{12} = 0 \) has been derived by Fujikoshi (1982) in the context of the problem of testing no additional information hypothesis in canonical correlation analysis. From a more general viewpoint, we here briefly treat it in a systematic way and then consider the asymptotic null distribution of the LRT for Problem [I]: \( \theta_{12} = 0 \) and \( \Sigma_{12} = 0 \). The notation

\[ \lambda \sim \Lambda_p(q,n) \]

denotes that for independent Wishart matrices \( A \) and \( B \),

\[ \lambda = |A|/|A + B| \] with \( A \sim W_p(\Sigma,n) \), \( B \sim W_p(\Sigma,q) \)

**Lemma 5.1** Let \( \lambda_i \sim \Lambda_{p_i}(q_i,n-d_i) \) and \( \lambda_i 's \) be independent \( i=1,2 \). Then

\[ P(-m \log \lambda_1 \lambda_2 \leq x) = G_f(x) + \frac{1}{m^2} \left[ G_{f+4}(x) - G_f(x) \right] + O(m^{-3}), \quad (5.1) \]

where \( G_f(x) \) is the cdf of \( \chi^2(f) \),

\[ f_1 = p_1q_1, \quad f_2 = p_2q_2, \quad f = f_1 + f_2, \quad m = n-p \]

\[ \rho = \frac{1}{f} \{ f_1[d_1 - \frac{1}{2}(q_1-p_1-1)] + f_2[d_2 + \frac{1}{2}(q_2-p_2-1)] \} \]

(5.2)

\[ u = s + \frac{f_1 f_2}{4f} \left[ d_1 - d_2 + \frac{1}{2}(p_1-p_2-q_1+q_2)^2 \right] \]

\[ s = s_1 + s_2, \quad s_1 = \frac{f_1}{48} (p_1^2+q_1^2-5) \quad \text{and} \quad s_2 = \frac{f_2}{48} (p_2^2+q_2^2-5). \]

**Proof.** The result follows directly by the usual method based on characteristic function.

Now for Problem [II], the LRT statistic is given by \( L_1 L_2 L_3 \) in (4.8) and from Lemma 4.1 and (4.11), it is easy to see that under the null hypothesis
\[ \lambda_2 \equiv L_1 L_2 \sim \Lambda_{p_2} (r_1+p_1, n-r-p_1) \quad (5.3) \]

\[ \lambda_1 = L_3 \sim \Lambda_{p_1} (r_2+p_2, n-r) \quad (5.4) \]

and \( \lambda_1 \) and \( \lambda_2 \) are independent. Therefore the following result follows from Lemma 5.1.

**Proposition 5.1.** For Problem [II]: \( \theta_{12} = 0, \theta_{21} = 0 \) and \( \varepsilon_{12} = 0 \), the asymptotic null distribution of LRT based on \(-m \log(L_1 L_2) L_3\) is given by (5.1) where in (5.2) \( q_1 = r_2 + p_2, q_2 = r_1 + p_1, d_1 = r \) and \( d_2 = r + p \).

On the other hand, for Problem [I], the LRT statistic is given by \( L_1 L_2 \) in (3.24), which is the same as the \( L_1 L_2 \) in Problem [II] and hence from (5.3) under the null hypothesis

\[ \lambda_2 = L_1 L_2 \sim \Lambda_{p_2} (r_1+p_1, n-r-p_1) \]

The asymptotic null distribution of \(-2m \log \lambda_2\) is well known.

Next we consider the asymptotic null distribution of \( R(B_1, B_2) \) in (4.15) for Problem [II] and as a special case, obtain the null distribution of \( T(B) \) in (3.22). But here for simplicity, the case \( \beta_1 p_1 r_1 = 1(i=1,2) \) will be treated. A more general case can be obtained in a similar manner. From (4.15), let the test statistic be

\[ \tilde{R} = \tilde{R}_1 + \tilde{R}_2 + \tilde{R}_3 + 2p_1 p_2 [1 + \frac{1}{2m} (r_1 r_2)] \quad (5.5) \]

where \( m = n-r \)

\[ \tilde{R}_1 = (m+r_2) \text{tr} B_{11} (B_{11} + S_{11})^{-1} \quad \text{with} \quad B_{11} = W_{11} W_{21} \]

\[ \tilde{R}_2 = (m+r_1) \text{tr} B_{22} (B_{22} + S_{22})^{-1} \quad \text{with} \quad B_{22} = W_{12} W_{12} \]
\[
\tilde{R}_3 = \frac{1}{m} (m+r_1)(m+r_2)\text{tr}(B_{11}+S_{11})^{-1}S_{12}(B_{22}+S_{22})^{-1}S_{21}
- \frac{P_2}{m} (m+r_2)\text{tr}S_{11}(B_{11}+S_{11})^{-1} - \frac{P_1}{m} (m+r_1)\text{tr}S_{22}(B_{22}+S_{22})^{-1}.
\]

**Theorem 5.1.** For Problem [II]: \(\theta_{12} = 0, \theta_{21} = 0,\) and \(\Sigma_{12} = 0,\) the asymptotic null distribution of \(\tilde{R}\) in (5.5) is given by

\[
P(\tilde{R} \leq x) = G_f(x) - \frac{1}{4m} (f_1 s_1 + f_2 s_2 + f_3 s_3)[G_f(x) - 2G_{f+2}(x) + G_{f+4}(x)]
- \frac{1}{2m} p_1 p_2 (r_1 + r_2)[2G_f(x) - 3G_{f+2}(x) + G_{f+4}(x)] + O(m^{-3/2})
\]  
(5.6)

where \(f = f_1 + f_2 + f_3, f_1 = p_1 r_2, f_2 = p_2 r_1, f_3 = p_1 p_2, s_1 = p_1 + r_2 + 1, s_2 = p_2 + r_1 + 1\) and \(s_3 = p_1 + p_2 + 1.\)

The case of Problem [I] follows directly from this theorem.

**Corollary 5.1.** For Problem [I]: \(\theta_{12} = 0\) and \(\Sigma_{12} = 0,\) the asymptotic null distribution of

\[
\tilde{T} = T(r_1/p_1) + 2p_1 p_2 [1 + \frac{1}{2m} r_1]
\]  
(5.7)

with \(T(\beta)\) in (3.22) is given by (5.6) with \(r_2 = 0.\)

**Outline of the Proof of Theorem 5.1.** Under the null hypothesis, assuming \(\Sigma = I\) without loss of generality, we have the three independent Wishart variates

\(B_{11} - W_p(I, r_2), B_{22} - W_p(I, r_1)\) and \(S - W_p(I, m).\) As usual let

\[
V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} = \sqrt{m} \left( \frac{1}{m} S - I \right)
\]  
(5.8)

and expand \(R_i's\) in terms of \(V, B_{11}\) and \(B_{22}\) as
\[ R_j = \text{tr} B_j + \frac{1}{\sqrt{m}} q_{j1} + \frac{1}{m} q_{j2} + O_p(m^{-3/2})(j=1,2) \]

\[ R_3 = \text{tr} V_{12} V_{21} - 2p_1 p_2 + \frac{1}{\sqrt{m}} q_{31} + \frac{1}{m} [q_{32} + \tilde{q}_{32} - p_1 p_2 (r_1 + r_2)] + O_p(m^{-3/2}) \]

where \( q_{j1} = -\text{tr} B_{jj} V_{jj} \), \( q_{j2} = \text{tr} B_{jj} V_{jj}^2 - \text{tr} B_{jj}^2 + \tilde{r}_j \text{tr} B_{jj} \) with \( \tilde{r}_1 = r_2 \) and \( \tilde{r}_2 = r_1 \) (\( j=1,2 \)), \( q_{31} = -\text{tr} V_{11} V_{12} V_{21} = -\text{tr} V_{12} V_{21} V_{21} + \text{tr} V_{12} V_{21}^2 \), \( q_{32} = \text{tr} V_{11} V_{12} V_{21} \) and \( \tilde{q}_{32} = (r_1 + r_2) \text{tr} V_{12} V_{21} - \text{tr} B_{11} V_{12} V_{21} - \text{tr} B_{22} V_{12} V_{21} + p_1 \text{tr} B_{11} + p_2 \text{tr} B_{22} \). Then \( \tilde{R} \) is expressed as

\[ \tilde{R} = \text{tr} B_{11} + \text{tr} B_{22} + \text{tr} V_{12} V_{21} + \frac{1}{\sqrt{m}} (q_{11} q_{21} + q_{31}) \]
\[ + \frac{1}{m} (q_{12} q_{22} + q_{32} + \tilde{q}_{32}) + 2p_1 p_2 \left[ 1 + \frac{1}{2m} (r_1 + r_2) \right] + O_p(m^{-3/2}) \] \hspace{1cm} (5.9)

Then the characteristic function \( C(t) \) of \( \tilde{R} \) is expanded as

\[ C(t) = E[H(t) \{ 1 + \frac{1}{\sqrt{m}} A_1 + \frac{1}{m} A_2 \}] + O(m^{-3/2}). \] \hspace{1cm} (5.10)

where \( H(t) = \exp[\text{tr} B_{11} + \text{tr} B_{22} + \text{tr} V_{12} V_{21}] \), and \( A_1 \) and \( A_2 \) are functions of \( B_i \) 's and \( V_{ij} \) 's. Since \( V_{ij} \) 's are not independent, in evaluation of the expectation in (5.10), we may use the following lemma.

**Lemma 5.1.** (1) The pdf of \( V \) is expanded as

\[ f(V) = c \exp(-\frac{1}{4} \text{tr} V^2) \left[ 1 + \frac{1}{\sqrt{m}} \left( \frac{1}{4} \text{tr} V + \frac{1}{6} \text{tr} V^3 \right) \right] + O(m^{-1}) \]

(2) The conditional pdf of \( V_{12} \) given \( V_{11} \) and \( V_{22} \) is expanded as

\[ f(V_{12} | V_{11}, V_{22}) = c \exp(-\frac{1}{2} \text{tr} V_{12} V_{21}) \left[ 1 + \frac{1}{2\sqrt{m}} \text{tr} V_{12} V_{21} - \text{tr} V_{11} V_{12} V_{21} + \text{tr} V_{11} V_{12} V_{21} \right] + O(m^{-1}). \]
Proof. (1) is well known. (2) is obtained by $f(V)/f_1(V_{11})f_2(V_{22})$, where the marginal pdf's of $V_{i1}$'s are first expanded as in (1).

Lemma 5.2. The characteristic function of $R$ is evaluated as

$$C(t) = (1-2it)^{-f/2} \left[ 1 - \frac{1}{4m} \left( f_1 s_1 + f_2 s_2 + f_3 s_3 \right) \left( (1-2it)^{-1} - 1 \right)^2 \right]$$

$$- \frac{1}{2m} \left( p_1 p_2 (r_1 + r_2) \left[ 2 - 3(1-2it)^{-1} + (1-2it)^{-2} \right] \right) + O(m^{-3/2}).$$

Proof. The proof is straightforward although it involves a lot of computation. The result is obtained by using Lemma 5.1 and the following well known results:

$$E[\exp(itm \tr V_{12} V_{21}) \{ 1 + \frac{it}{\sqrt{m}} q_{31} + \frac{1}{m} (it)q_{32} + \frac{(it)^2}{2} q_{31}^2 \}]$$

$$= E \exp(itm \tr \bar{S}_{11}^{-1} S_{12} \bar{S}_{22}^{-1} S_{21}) + O(m^{-3/2})$$

$$= (1-2it)^{-f_3/2} \left[ 1 - \frac{1}{4m} f_3 s_3 ((1-2it)^{-1} - 1)^2 \right]$$

$$+ O(m^{-3/2})$$

(Fujikoshi (1970), Muirhead (1970))

$$E[\exp(itm + r_2) \tr B_{11}(B_{11} + S_{11})^{-1} \left( f_1 s_1 + f_2 s_2 \right)/2]$$

$$+ it(m + r_1) \tr B_{22}(B_{22} + S_{22})^{-1}])$$

$$= (1-2it)^{-f_3/2} \left[ 1 - \frac{1}{4m} \left( f_1 s_1 + f_2 s_2 \right) + O(m^{-3/2}) \right]$$

(Fujikoshi (1970), Muirhead (1970))

Inverting $C(t)$ in Lemma 5.2, we obtain the result in Theorem 5.1.
6. TESTS FOR INDEPENDENCE WHEN \( \theta_{12} = 0 \) AND \( \theta_{21} = 0 \)

As has been stated in Section 2, for problem [III] \( H : \Sigma_{12} = 0 \) given \( \theta_{12} = 0 \) and \( \theta_{21} = 0 \) KFK (1984) proposed the LBI test, a LRT-like test and a trace test, and considered the asymptotic null and nonnull distributions of these tests. In this section, adopting what we call the method of small-small asymptotics due to Mukerjee and Chandra (1984), we compare the power functions of those tests. Since the model in Problem [III] is

\[
\begin{bmatrix}
Y_1, Y_2 \\
p_1, p_2 \\
r_1, r_2
\end{bmatrix} = \begin{bmatrix}
\theta_{11} & 0 \\
0 & \theta_{22}
\end{bmatrix} + \begin{bmatrix}
E_1, E_2
\end{bmatrix}
\begin{bmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22}
\end{bmatrix}
\]

(6.1)

(6.2)

it contains the MANOVA model as a special case with \( X_1 = X_2 \). The comparison we make here deals with the comparison of the Pillai test and the LRT for independence in the MANOVA model. In fact, as will be shown, when \( X_1 = X_2 \), the LBI test and the trace test in Problem [III] are both reduced to the Pillai test of independence in the MANOVA model, which is LBI, while the LRT-like test is reduced to the LRT of independence in the MANOVA model. Following KFK (1984), assume \( p_1 > p_2 \) without loss of generality and let

\[
\begin{cases}
Q_0 = I - X'(X'X)'^{-1}X' = Z_0Z_0' \text{ with } Z_0Z_0' = I_{n_0} \\
Q_i = I - X_i'(X_i'X_i)^{-1}X_i \quad (i=1,2)
\end{cases}
\]

(6.3)

where \( Z_0 : n \times n_0, n_0 = n-r_0 \) and \( r_0 = \text{rank } [X_1, X_2] \) and \( A^+ \) is Penrose inverse of \( A \). Further let

\[
X'(X'X)^{-1}X' - X_i'(X_i'X_i)^{-1} = \tilde{Z}_i\tilde{Z}_i' \text{ with } \tilde{Z}_i\tilde{Z}_i' = I_{r_0-r_1}
\]
where \( \tilde{Z}_i : n \times (r_0 - r_i) \), and let

\[
Z_i = \begin{bmatrix} \tilde{Z}_i \mid Z_0 \end{bmatrix} : n \times (n_0 + r_0 - r_i),
\]

\[ (6.5) \]

\[
\begin{bmatrix} M_i \mid (r_0 - r_i) \\ U_i \mid (n - r_0) \end{bmatrix} (n_0 + r_0 - r_i) = \begin{bmatrix} \tilde{Z}_i \mid Y_i \\ Z_0 \mid Y_i \end{bmatrix},
\]

\[ (6.6) \]

\[
P_i
\]

\[
S = G + B \quad \text{and} \quad R = S_{12} S_{22}^{-1} S_{21}^{-1}
\]

\[ (6.7) \]

where \( S = (S_{ij}) \) with \( S_{ij} : p_i \times p_j \), \( G = (G_{ij}) \) with \( G_{ij} = U_i U_j : p_i \times p_j \) and \( B = (B_{ij}) \) for \( i, j = 1, 2 \) with

\[
B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} M_{11} M_1 & M_{12} K M_2 \\ M_{21} K M_1 & M_{22} M_2 \end{pmatrix},
\]

\[ K = \tilde{Z}_i \tilde{Z}_2 \]

\[ (6.8) \]

Here we note that when \( X_1 = X_2, Q_0 = Q_i \) \( (i = 1, 2) \) so that \( B = 0 \) and \( S = G \) with \( G = W(E, n_0) \), which is nothing but the canonical form for the problem of testing \( \Sigma_{12} = 0 \) in the MANOVA model. Now based on the notation introduced above, our problem is to test \( \Sigma_{12} = 0 \) based on \( S \) in (5.7) and then the LRT-like test statistic, the tract test statistic and the LBI test statistic considered in KFK are respectively expressed as

\[
T_1 = - n_0 \log |I - R| \]

\[ (6.9) \]

\[
T_2 = n_0 \text{tr} R
\]

\[ (6.10) \]

\[
T_3 = \frac{1}{n_0} \left( n_1 n_2 \text{tr} R - n_1 p_2 \text{tr} s_{11}^{-1} Y_1' Q_1 Q_2 Y_1 \\
- n_2 p_1 \text{tr} s_{22}^{-1} Y_2' Q_2 Q_1 Y_2 + p_1 p_2 \left( 2 - \frac{1}{n_0} \right) t \right)
\]

\[ (6.11) \]
where
\[ n_i = n - r_i \ (i=0,1,2) \text{ and } t = r_1 + r_2 - 2r. \] (6.12)

As has been shown in KFK, the power functions of these invariant tests
depend on \((0,\Sigma)\) only through the canonical correlations \(\rho_1^2 \geq \ldots \geq \rho_{p_2}^2\) of
\(\Sigma\) or the characteristic roots of \(\Sigma^{-1}_2 \Sigma^{-1}_1\). Here to consider the small-small asymptotics for the power functions, first fix \(\omega_1 \geq \ldots \geq \omega_{p_2} \geq 0\) with
\(\omega_1 \neq 0\), which are chosen to be small later, and take
\[ \frac{1}{2} \rho^2_{j} = \frac{1}{n_0} \omega_j \ (j=1,\ldots,p_2) \] (6.13)
where \(\frac{1}{n_0}\) is supposedly small. This implies \(\rho_j = (2\omega_j/n_0)^{\frac{1}{2}}\) is eventually
small-small. Further let
\[ \delta = \omega_1 + \ldots + \omega_{p_2} = \text{tr} \phi \] (6.14)
\[ \phi = \text{diag}\{\omega_1,\ldots,\omega_{p_2}\} \] (6.15)
\[ f = p_1 p_2 \text{ and } s = p_1 + p_2 + 1. \] (6.16)

Then from the results in KFK, the asymptotic power functions of the tests
based on \(T_i\)'s in (5.9), (5.10) and (5.11) under the local alternative \(\{\rho_j\}\)
in (5.13) (as \(n_0 \rightarrow \infty\)) are given by
\[ P(T_i > x_i) = \delta_f(x_i,\delta) + \frac{1}{n_0} \sum_{j=0}^{s} a_{ij} \delta_{f+2j}(x_i,\delta) + O(n_0^{-2}), \] (6.17)
\((i=1,2,3)\) where \(G(x;\delta) = 1 - G_k(x;\delta), G_k(x;\delta)\) is the distribution function
of \(X^2\) distribution with degrees of freedom \(k\) and noncentrality parameter \(\delta\),
\[ a_{10} = -\frac{1}{4} f_s - \frac{1}{2} f(t+trKK') - (trKK')\delta - tr^2 \]

\[ a_{11} = \frac{1}{4} f_s + \frac{1}{2} f(t+trKK') - t\delta \]

\[ a_{12} = (t+trKK')\delta + 2tr^2, \quad a_{13} = -tr^2, \quad a_{14} = 0 \]

\[ a_{20} = -\frac{1}{4} f_s - \frac{1}{2} f(t+trKK') - (trKK')\delta - tr^2 \]

\[ a_{21} = \frac{1}{2} f_s + \frac{1}{2} f(t+trKK') - t\delta \]

\[ a_{22} = -\frac{1}{4} f_s + (s+t+trKK')\delta + 2tr^2 \]

\[ a_{23} = -s\delta, \quad a_{24} = -tr^2 \]

\[ a_{30} = -\frac{1}{4} f_s - \frac{1}{2} ftrKK' - (trKK')\delta - tr^2 \]

\[ a_{31} = \frac{1}{2} f_s + \frac{1}{2} ftrKK' \]

\[ a_{32} = -fs + (s+trKK')\delta + 2tr^2 \]

\[ a_{33} = -s\delta \text{ and } a_{34} = -tr^2. \]

Note that \( i^4_j = 0 \) \( (i=1,2,3) \). To evaluate the power functions in (6.17) further, let \( g_f(x:\delta) \) be the pdf of \( G_f(x:\delta) \) and let \( G_f(x) = G_f(x:0) \) and \( g_f(x) = g_f(x:0) \). Then it is easy to see that

\[ \bar{g}_{f+2}(x:\delta) = 2g_{f+2}(x:\delta) + \bar{g}_f(x:\delta) \]  (6.18)

\[ g_{f+2}(x) = xg_f(x)/f \]  (6.19)

\[ \bar{g}_f(x+\frac{c_m}{m}:\delta) = \bar{g}_f(x:\delta) - \frac{c}{m} g_f(x:\delta) + O(m^{-2}). \]  (6.20)

Using (6.18), the power functions are evaluated as
\( P(T_i > x_i) = \tilde{g}_f(x_i; \delta) + \frac{1}{n_0} \sum_{j=1}^{4} b_{ij} \tilde{g}_{f+2j}(x_i; \delta) + o(n_0^{-2}) \) (6.21)

where

\[
\begin{align*}
    b_{11} &= \frac{1}{2} f_s + f(t+trKK') + 2(trKK')\delta + 2tr\phi^2 \\
    b_{12} &= 2(t+trKK')\delta + 2tr\phi^2 \\
    b_{13} &= -2tr\phi^2, \quad b_{14} = 0 \\
    b_{21} &= \frac{1}{2} f_s + f(t+trKK') + 2(trKK')\delta + 2tr\phi^2 \\
    b_{22} &= -\frac{1}{2} f_s + 2(t+trKK')\delta + 2tr\phi^2 \\
    b_{23} &= -2s\delta - 2tr\phi^2, \quad b_{24} = -2tr\phi^2 \\
    b_{31} &= \frac{1}{2} f_s + f trKK' + 2(trKK')\delta + 2tr\phi^2 \\
    b_{32} &= -\frac{1}{2} f_s + 2(trKK')\delta + 2tr\phi^2 \\
    b_{33} &= -2s\delta - 2tr\phi^2, \quad b_{34} = -2tr\phi^2.
\end{align*}
\]

Using (6.21) and (6.19), under the null hypothesis \( \delta = 0 \)

\[
\alpha = P(T_i > x_i | H_0) = G_f(x_i) - \frac{1}{n_0} \tilde{b}_{i0} g_f(x_i) + o(n_0^{-2}) \quad (6.22)
\]

\((i=1,2,3)\) where

\[
\begin{align*}
    \tilde{b}_{10} &= -x_1 \left( \frac{1}{2} s + t + trKK' \right) = -x_1 c_1 \\
    \tilde{b}_{20} &= -x_2 \left( \frac{1}{2} s + t + trKK' - \frac{s x_2}{2(f+2)} \right) = -x_2 c_2 \\
    \tilde{b}_{30} &= -x_3 \left( \frac{1}{2} s + trKK' - \frac{s x_3}{2(f+2)} \right) = -x_3 c_3
\end{align*}
\] (6.23)
From (6.20)(or Hill and Davis formula) we obtain with $u$ in $f(u) = x = u b_0 + O(n^{-2})$ (6.24)

where $b_0 = -u C_i$ with $C_i$ in (6.23). Hence using (6.20) and (6.24), we obtain the following theorem.

**Theorem 6.1.** Let $\pi(\phi_1, \phi)$ be the power function of the test $\phi_1$ of size $\alpha$ with critical region $T_i > x_i (i=1,2,3)$. Then for $u$ satisfying $\delta_f(u) = \alpha$ and for $\delta$ small, it is evaluated as

$$\pi(\phi_1, \phi) = \delta_f(u; \delta) + \frac{1}{n_0} H_i(\delta) g_f(u) + O(n_0^{-2})(i=1,2,3)$$

(6.25)

where

$$H_i(\delta) = c_{i1} \delta + c_{i2} \delta^2 + c_{i3} \text{tr}\phi^2 + O(n_0^{-2})(i=1,2,3)$$

(6.26)

$$c_{11} = c_{21} = c_{31} = \frac{2u}{f} \text{tr}\mathbf{K} - \frac{su^2}{f(f+2)}$$

$$c_{12} = -\frac{2u}{f} \text{tr}\mathbf{K} + (s + 2 \text{tr}\mathbf{K}) \frac{u^2}{f(f+2)} - \frac{su^3}{f(f+2)(f+4)}$$

$$c_{22} = c_{32} = c_{12} + \frac{2su^4}{f(f+2)(f+4)(f+6)}$$

$$c_{13} = \frac{2u}{f} + \frac{2u^2}{f(f+2)} - \frac{2u^3}{f(f+2)(f+4)}$$

and

$$c_{23} = c_{33} = c_{13} - \frac{2u^4}{f(f+2)(f+4)(f+6)}$$

**Proof.** Using (6.20) and (6.24), the power functions in (6.21) are directly shown to be equal to those in (6.25) after some algebra, where $g_{f+3j}(u; \delta) = e^{-\delta} \sum_{k=0}^{\infty} (\delta^k/k!) g_{f+2j+2k}(u)$ and $g_{f+2}(u) = ug_f(u)/f$ were used.
In the expression (6.25), $\delta = \sum_{i=1}^{p_2} \omega_i$ has been chosen to be small. That is, the power functions were first asymptotically expanded in the orders of $n_0^{-k}(k=0,1,2...)$, and then for $\delta$ small the terms of order $n_0^{-1}$ were asymptotically expanded in the orders of $\delta^k (k=0,1,...)$ since the terms of order $n_0^0$ are common to all the tests, i.e., $\delta_f(u;\delta)$. From this expression, the local behaviors of the power functions are compared as follows.

**Theorem 6.2.** For the tests $\phi_i$, it follows that with $\delta = \text{tr} \phi$

(1) $\lim_{\delta \to 0} \lim_{n_0 \to \infty} n_0 [{\pi}(\phi_j,\phi) - \pi(\phi_1,\phi)]/\delta = 0$ (j=2,3)

(2) $\lim_{\delta \to 0} \lim_{n_0 \to \infty} n_0 [{\pi}(\phi_j,\phi) - \pi(\phi_1,\phi)]/\delta^2 = \Delta(\phi)$ (j=2,3)

and

(3) $\lim_{\delta \to 0} \lim_{n_0 \to \infty} n_0 [{\pi}(\phi_3,\phi) - \pi(\phi_2,\phi)]/\delta^2 = 0$

where $\Delta(\phi) = [2u^4g_f(u)/f(f+2)(f+4)(f+6)]\gamma(\phi)$ with

$$\gamma(\phi) = (p_1+p_2+1) - (p_1p_2+2) \lim_{\delta \to 0} [\text{tr} \phi^2/(\text{tr} \phi)^2]$$

(6.27)

**Proof.** Immediate from Theorem 6.1.

Now (1) implies that in terms of power all the three tests are asymptotically equivalent up to $O(n_0^{-1})$ and $O(\delta)$. The asymptotic difference between the LBI test $\phi_3$ (or the trace test $\phi_2$) and the LRT-like test $\phi_1$ appears in the term of $O(n_0^{-1})$ and $O(\delta^2)$ as is shown in (2). Setting $\tau = \lim_{\delta \to 0} [\text{tr} \phi^2/(\text{tr} \phi)^2]$, from (2), if $\gamma(\phi) > 0$, or equivalently $(p_1+p_2+1)/(p_1p_2+2) > \tau$, the LBI test is asymptotically better up to $O(n_0^{-1})$ and $O(\delta^2)$ than the LRT-like test, while if $\gamma(\phi) < 0$, the LRT-like test is asymptotically better. Since $\text{tr} \phi^2 = \Sigma \omega_i^2$ and $\text{tr} \phi = \Sigma \omega_i = \delta$, the inequality

$$\Sigma \omega_i^2 \leq [\Sigma \omega_i]^2 \leq p_2 [\Sigma \omega_i^2]$$
follows from $\omega_1 \geq 0$ and Schwartz's inequality. This implies

$$\frac{1}{p_2} < \tau = \lim_{\delta \to 0} \frac{\text{tr}s^2}{[\text{tr}s]^2} < 1 \quad (6.28)$$

The equality in the first of (6.28) holds if and only if $\omega_1 = \omega_2 = \ldots = \omega_{p_2}$, while the equality in the second inequality holds if and only if $\omega_2 = \ldots = \omega_{p_2} = 0$ since $\omega_1 \geq \omega_2 \geq \ldots \geq \omega_{p_2} \geq 0$. On the other hand,

$$\frac{1}{p_2} < \tau_0 = (p_1+p_2+1)/(p_1p_2+1) < 1$$

since $p_1 > p_2 > 0$. Hence both the cases $1 \geq \tau \geq \tau_0$ and $\tau_0 > \tau \geq 1/p_2$ can occur. The above observation will show that the closer $\tau$ is to 1, the more concentrated $\omega_i$'s are around $\omega_1 = \ldots = \omega_{p_2}$, while the closer $\tau$ is to $1/p_2$, the more spread $\omega_i$'s are.

Next, we consider the small-small asymptotic comparison between the LRT and the Pillai test of independence in the MANOVA model. As has been pointed out, the problem of testing independence in the MANOVA model is included as a special case of our problem with $X_1 = X_2$. In case $X_1 = X_2$, $\bar{z}_i = 0$ in (6.4) so that $S = G$ in (6.7), $K = 0$ in (6.8), $n_1 = n_2 = n_0$ in (6.12), and $T_2 = T_3$ in (6.9) and (6.11). However this does not cause any changes in the results of Theorem 6.1 and 6.2 except for the slight changes of the coefficients $c_{ij}$'s in Theorem 6.1. That is, by setting $K = 0$ in $c_{11}$ and $c_{12}$, the results in Theorem 6.1 holds as it is, while Theorem 6.1 is effective whether or not $X_1 = X_2$.

Corollary 6.1. For testing independence in the MANOVA model with $X_1 = X_2$, all the results in Theorem 6.1 hold. If $\tau_0 > \tau(\geq 1/p_2)$, Pillai's test, which is LBI for fixed $n_0$, is asymptotically better up to $O(n_0^{-1})$ and $O(\delta^2)$ than the LRT while if $1 > \tau > \tau_0$, the LRT is asymptotically better.
7. REMARKS

In this section, we first consider Problem [IV]: $\theta_{12} = 0$ and $\theta_{12} = 0$. For this problem, it is easy to see that a canonical form of the model is also given by the model in (4.2) where the joint hypothesis $\theta_{12} = 0$, $\theta_{21} = 0$ and $\Sigma_{12} = 0$ in Problem [II] was tested, and the hypothesis here becomes $\eta_{12} = 0$ and $\eta_{21} = 0$. Further, the same group $\mathcal{G} = \mathcal{O}(p_1) \times \mathcal{O}(p_2) \times \mathbb{R}^1 \times \mathbb{R}^2$ acting on $(W,S)$ as in (4.5) leaves this problem. Hence the class of invariant tests in Problem [II] is exactly the same as the class of invariant test in Problem [IV] we are presently considering. This implies that the power function of an invariant test in Problem [IV] is also a function of $\Omega$, $\xi_{12}$ and $\xi_{21}$ in (4.6).

However, under the null hypothesis $\xi_{12} = 0$ and $\xi_{21} = 0$, it still depends on $\Omega$ because $\Omega$ may not be zero in Problem [IV]. Therefore in general an invariant test in the present problem is not similar. In fact, this follows from the fact that the group does not act transitively on the parameter space of the null hypothesis. This is true even in the case $X_1X_2 = 0$ where the group is enlarged to $\mathcal{G} = \mathcal{O}(r_1) \times \mathcal{O}(r_2) \times G$ as in (4.5). For example, suppose we construct such statistics as

$$L_2 = |S_{22}|/|S_{22} + W_{12}W_{12}| \quad \text{and} \quad L_3 = |S_{11}|/|S_{11} + W_{21}W_{21}|$$

or

$$L_2' = \text{tr} W_{12}S_{22}^{-1}W_{12}' \quad \text{and} \quad L_3' = \text{tr} W_{21}S_{11}^{-1}W_{21}'.$$

But here $L_2$ and $L_3$ (or $L_2'$ and $L_3'$) are correlated under the null hypothesis so that any test combining $L_2$ and $L_3$ (or $L_2'$ and $L_3'$) is not similar unless one of the two statistics is completely ignored. Because of the non-similarity
feature of the problem we leave it here. One might use a non-similar test by combining $L_2$ and $L_3$ in (7.1) in such a way as $L_2L_3$ or might test the two hypothesis separately. It is noted that an explicit form of the LRT for the present problem is difficult to derive.
REFERENCES


On tests for selection of variables and Independence under multivariate regression model

T. Kariya, Y. Fujikoshi, P. R. Krishnaiah

Center for Multivariate Analysis
515 Thackeray Hall
University of Pittsburgh, Pittsburgh, PA 15260

Air Force Office of Scientific Research
Department of the Air Force
Bolling Air Force Base, DC 20332

Approved for public release; distribution unlimited

Asymptotic distributions, Correlated multivariate regression equations (CMRE); Growth curve model, Multivariate distributions, Multivariate regression analysis, Optimum properties.

In this paper, the authors consider various procedures for testing the hypotheses of independence of two sets of variables and certain regression coefficients are zero under the classical multivariate regression model. Various properties of these procedures and the asymptotic distributions associated with these procedures are also considered.