A STUDY OF REDUNDANCY OF SOME VARIABLES IN COVARIATE DISCRIMINANT ANALYSIS

YASUNORI FUJIKOSHI¹ AND CHINUBAL G. KHATRI²

¹Department of Mathematics, Hiroshima University, Hiroshima 730, Japan
²Center for Multivariate Analysis, Pennsylvania State University, University Park, PA 16802, U.S.A.

(Received March 1, 1989; revised August 15, 1989)


Key words and phrases: Covariate and discriminate variables, likelihood ratio test, redundancy of variables, test of redundancy, Wilks lambda distribution.

1. Introduction

Rao (1946, 1948, 1970) discussed the redundancy of a given set of variables, for purposes of discrimination between two groups with the same covariance matrix. Rao (1948, 1970), McKay (1977) and Fujikoshi (1982) have extended this idea to the case of multiple groups. Such a question arises in the signal detection theory due to the inclusion of more variables than necessary, and the techniques developed here will be useful in this area as well.

In this paper we consider the redundancy of a given set of variables in covariate discriminant analysis, i.e., in the situation where there are covariate variables as well as discriminate variables. Two-group covariate discriminant analysis was considered by Cochran and Bliss (1948) and Cochran (1964). They proposed using the classification statistic \( W^* \). Memon and Okamoto (1970) obtained an asymptotic expansion of the distribution of \( W^* \) and discussed the information gained by assuming that some variables
have equal means. Fujikoshi and Kanazawa (1976) considered the ML rule in the covariate case and introduced the classification statistic \( Z^* \). For a summary, see, e.g., Siotani (1982). In Section 2 we give several equivalent statements for the redundancy in the case of two groups. These statements are extended to the case of multiple groups in Section 3. In Section 4 we consider the tests for redundancy. We give a canonical form for the testing problem. It is shown that the LR test can be decomposed as two independent LR tests; some lemmas useful in obtaining the LR test are given in the Appendix. An alternative test is also suggested.

2. The case of two groups

Cochran and Bliss (1948) and Cochran (1964) discussed the classification problem when some variables are known to have the same means in two groups \( \Pi_1 \) and \( \Pi_2 \). Let \( x \) and \( z \) be the vector variates of \( p \) discriminators and \( q \) covariates, respectively. It is assumed that

\[
E \left( \begin{pmatrix} x \\ z \end{pmatrix} \bigg| \Pi_i \right) = \begin{pmatrix} \mu^{(i)} \\ \xi \end{pmatrix},
\]

(2.1)

\[
\text{Var} \left( \begin{pmatrix} x \\ z \end{pmatrix} \bigg| \Pi_i \right) = \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xz} \\ \Sigma_{zx} & \Sigma_{zz} \end{pmatrix}.
\]

Letting \( \Sigma_{xx,z} = \Sigma_{xx} - \Sigma_{xz} \Sigma_{zz}^{-1} \Sigma_{zx} \),

\[
y = \begin{pmatrix} x \\ z \end{pmatrix}, \quad \eta^{(i)} = \begin{pmatrix} \mu^{(i)} \\ \xi \end{pmatrix} \quad \text{and} \quad \delta = \eta^{(1)} - \eta^{(2)} = \begin{pmatrix} \delta_x \\ 0 \end{pmatrix},
\]

we can write the population linear discriminant function as

\[
(\eta^{(1)} - \eta^{(2)})' \Sigma^{-1} \left\{ y - \frac{1}{2} (\eta^{(1)} + \eta^{(2)}) \right\}
\]

\[
= \delta' \Sigma_{xx,z}^{-1} \left\{ x - \Sigma_{xz} \Sigma_{zz}^{-1} (z - \xi) - \frac{1}{2} (\mu^{(1)} + \mu^{(2)}) \right\}.
\]

The last expression shows that the discriminant function is one of the random variates \( x^* = x - \Sigma_{xz} \Sigma_{zz}^{-1} z \) obtained by subtracting from \( x \) its regression on \( z \). The vector of coefficients in \( y \) is

(2.2)

\[
\gamma = \Sigma^{-1} (\eta^{(1)} - \eta^{(2)})
\]

\[
= \left( \begin{pmatrix} I \\ -\Sigma_{zz}^{-1} \Sigma_{xz} \end{pmatrix} \Sigma_{xx,z}^{-1} \right) \delta_x.
\]
In order to formulate the redundancy of subvectors of \( \mathbf{x} \) and \( \mathbf{z} \), we partition \( \mathbf{x} \) and \( \mathbf{z} \) as

\[
\mathbf{x}' = (\mathbf{x}', \mathbf{z}'), \quad \mathbf{x}'_i : p_i \times 1, \quad \mathbf{z}' = (\mathbf{z}', \mathbf{z}'), \quad \mathbf{z}'_i : q_i \times 1
\]

and \( \mathbf{y}, \gamma, \Sigma \) conformably as follows:

\[
(2.3) \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}, \quad \gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \Sigma_{34} \\ \Sigma_{41} & \Sigma_{42} & \Sigma_{43} & \Sigma_{44} \end{pmatrix}.
\]

The Mahalanobis distance between \( \Pi_1 \) and \( \Pi_2 \) based on \( \mathbf{y} \) is \( d_{p+q} = (\delta' \Sigma^{-1} \delta)^{1/2} \), while the one based on \( \mathbf{y}' = (y'_1, y'_2) \) is \( d'_{p+q,\Pi_1} = \delta_a'^{-1} \delta_a \), where \( y'_b = (y'_2, y'_4) \),

\[
(2.4) \quad E \begin{pmatrix} y_a \\ y_b \end{pmatrix} = \begin{pmatrix} \delta_a \\ \delta_b \end{pmatrix}, \quad \text{Var} \begin{pmatrix} y_a \\ y_b \end{pmatrix} = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}.
\]

We note that \( y_b \) involves both the subvectors of \( \mathbf{x} \) and \( \mathbf{z} \). We consider the following five statements on the redundancy of \( y_b \), which are shown to be equivalent.

1. The random variate \( y_a - \Sigma_{ba} \Sigma^{-1} y_a \) obtained by subtracting from \( y_b \) its regression on \( y_a \) has the same expected value for both groups.
2. The Mahalanobis distances between \( \Pi_1 \) and \( \Pi_2 \) based on \( y \) and \( y_a \) are the same, i.e.,

\[
(2.5) \quad d_{p+q} = d_{p+q,\Pi_1}.
\]

3. The coefficients of \( y_2 \) and \( y_4 \) in the linear discriminant function based on \( y \) are all zero, i.e., \( \gamma_2 = 0, \gamma_4 = 0 \).
4. The Mahalanobis distances between \( \Pi_1 \) and \( \Pi_2 \) based on the random variates \( y_{x,z} = \mathbf{x} - \Sigma_{xx} \Sigma^{-1} \mathbf{z} \) and \( y_{1,3} = \mathbf{x} - \Sigma_{11} \Sigma^{-1} \mathbf{z} \) are the same, i.e.,

\[
(2.6) \quad \delta_x' \Sigma_{xx}^{-1} \delta_x = \delta_i' \Sigma_{11,3}^{-1} \delta_i,
\]

where \( \Sigma_{xx} = \Sigma_{xx} - \Sigma_{xx} \Sigma_{zz}^{-1} \Sigma_{zx} \) and \( \Sigma_{11,3} = \Sigma_{11} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31} \).

5. \( \delta_2 - \Sigma_{21,2} \Sigma_{11,2}^{-1} \delta_1 = 0 \) and \( \Sigma_{41,2} \Sigma_{11,2}^{-1} \delta_1 = 0 \).
available were given by Rao (1970). His proof for the equivalence of (1)–(3) is applicable to our case. The equivalence (1)–(5) will be shown as a special case of multiple groups.

3. The case of multiple groups

We consider the formulation for redundancy in the case of \( k + 1 \) groups. Suppose that \( y' = (x', z') \) has the means

\[
E(y'|\Pi_i) = \eta^{(i)} = \begin{pmatrix} \mu^{(i)} \\ \xi \end{pmatrix}, \quad i = 1, \ldots, k + 1
\]

and the same covariance matrix \( \Sigma \) as in (2.1). Following McKay (1977) and Fujikoshi (1982), we extend the statements (1)–(5) to the case of multiple groups. The statement (1) can be extended as

\[
\eta_{b}^{(1)} - \Sigma_{ba} \Sigma_{aa}^{-1} \eta_{a}^{(1)} = \cdots = \eta_{b}^{(k+1)} - \Sigma_{ba} \Sigma_{aa}^{-1} \eta_{a}^{(k+1)}.
\]

Let \( \Omega \) be the population between-groups covariance matrix defined by

\[
\Omega = \Theta G \Theta'
\]

\[
= \sum_{i=1}^{k+1} g_i (\eta^{(i)} - \bar{\eta}) (\eta^{(i)} - \bar{\eta})' = \begin{pmatrix} \Omega_{xx} & 0 \\ 0 & 0 \end{pmatrix},
\]

where \( g_i \)'s are positive constants such that \( \sum g_i = 1, \bar{\eta} = \sum g_i \eta^{(i)} \), \( \Theta = (\eta^{(i)} - \bar{\eta}, \ldots, \eta^{(k+1)} - \bar{\eta}) \) and \( G = \text{diag} (g_1, \ldots, g_{k+1}) \). A choice of \( \{g_i\} \) may be done, based on the sample sizes from \( \Pi_i \). The population Fisher’s discriminant functions are defined by using the characteristic vectors of \( \Omega \) with respect to \( \Sigma \). Let \( \gamma_j \) be the solutions of

\[
\Omega \gamma_j = l_j \Sigma \gamma_j, \quad \gamma_j' \Sigma \gamma_j = 1, \quad j = 1, \ldots, m,
\]

where \( m = \text{rank } \Omega \leq \min (p, k) \). Let \( \Omega \) and \( \gamma_j' = (\gamma_1', \gamma_2', \gamma_3', \gamma_4') \) partition as in (2.3) and (2.4). Similarly we partition \( \Theta \) as

\[
\Theta = \begin{pmatrix} \Theta_1 \\ \Theta_2 \\ \Theta_3 \\ \Theta_4 \end{pmatrix}, \quad \Theta_2 = \begin{pmatrix} \Theta_1 \\ \Theta_3 \end{pmatrix}, \quad \Theta_b = \begin{pmatrix} \Theta_2 \\ \Theta_4 \end{pmatrix}.
\]

The statements (2)–(5) can be extended as follows:
(3.6) \[ \text{tr } \Omega \Sigma^{-1} = \text{tr } \Omega_{aa} \Sigma_{aa}^{-1}, \]

(3.7) \[ \gamma_{aj} = 0, \quad \gamma_{bj} = 0, \quad j = 1, \ldots, m, \]

(3.8) \[ \text{tr } \Omega_{xx} \Sigma_{xx,z}^{-1} = \text{tr } \Omega_{111} \Sigma_{11,3}^{-1}, \]

(3.9) \[ \Theta_2 - \Sigma_{21,z} \Sigma_{11,z}^{-1} \Theta_1 = 0 \quad \text{and} \quad \Sigma_{41,3} \Sigma_{11,3}^{-1} \Theta_1 = 0. \]

Statements (3.6) and (3.7) in the case when the covariate is not available have been considered by McKay (1977) and Fujikoshi (1982), respectively.

**Theorem 3.1.** The statements (3.2), (3.6)–(3.9) are equivalent.

**Proof.** Though the equivalence of (3.2), (3.6) and (3.7) follows from the proof as in McKay (1977) and Fujikoshi (1982), we give a complete proof. Noting that

\[
\left( \begin{array}{cc}
\Sigma_{aa} & \Sigma_{ab} \\
\Sigma_{ba} & \Sigma_{bb}
\end{array} \right)^{-1} = \left( \begin{array}{cc}
\Sigma_{aa}^{-1} & 0 \\
0 & 0
\end{array} \right) + \left( -\Sigma_{aa}^{-1} \Sigma_{ab} \right) \Sigma_{bb,a} \left( -\Sigma_{aa}^{-1} \Sigma_{ab} \right),
\]

we obtain

\[
\text{tr } \Omega \Sigma^{-1} = \text{tr } \Omega_{aa} \Sigma_{aa}^{-1}
\]

\[+ \text{tr } \Sigma_{bb,a}(\Theta_b - \Sigma_{ba} \Sigma_{aa}^{-1} \Theta_a)G(\Theta_b - \Sigma_{ba} \Sigma_{aa}^{-1} \Theta_a)'.
\]

Therefore, (3.6) is equivalent to

\[
\Theta_b - \Sigma_{ba} \Sigma_{aa}^{-1} \Theta_a = 0,
\]

which is equivalent to (3.2). We can write (3.4) as

\[
\left( \begin{array}{cc}
\Omega_{aa} & \Omega_{ab} \\
\Omega_{ba} & \Omega_{bb}
\end{array} \right) \left( \begin{array}{cc}
\gamma_{aj} \\
\gamma_{bj}
\end{array} \right) = l_j \left( \begin{array}{cc}
\Sigma_{aa} & \Sigma_{ab} \\
\Sigma_{ba} & \Sigma_{bb}
\end{array} \right) \left( \begin{array}{cc}
\gamma_{aj} \\
\gamma_{bj}
\end{array} \right), \quad j = 1, \ldots, m.
\]

Suppose that (3.12) is true. Then, premultiplying both sides of (3.13) by \((-\Sigma_{ba} \Sigma_{aa}^{-1} , I)\) we have

\[0 = l_j(0, \Sigma_{bb,a}) \left( \begin{array}{cc}
\gamma_{aj} \\
\gamma_{bj}
\end{array} \right).
\]

Hence, \(\gamma_{bj} = 0, \quad j = 1, \ldots, m\). Conversely, suppose that (3.7) is true, i.e., \(\gamma_{bj} = 0\). Then from (3.13) we have

\[\Omega_{aa} \gamma_{aj} = l_j \Sigma_{aa} \gamma_{aj}, \quad j = 1, \ldots, m\]
which implies (3.6), i.e., \( \text{tr } \Omega_{aa} \Sigma_{aa}^{-1} = \sum_{j=1}^{m} l_j = \text{tr } \Omega \Sigma^{-1} \). The equivalence of (3.8) and (3.9) follows from

\[
\begin{align*}
\text{tr } \Omega_{xx} \Sigma_{xx}^{-1} - z &= \text{tr } \Omega_{11} \Sigma_{11}^{-1} \\
&+ \text{tr } (\Sigma_{22} - \Sigma_{21.3} \Sigma_{11.3}^{-1} \Sigma_{14.3})^{-1} (\Theta_2 - \Sigma_{21.3} \Sigma_{11.3}^{-1} \Theta_1) \\
&\cdot G(\Theta_2 - \Sigma_{11.3} \Sigma_{11.3}^{-1} \Theta_1)' \\
&+ \text{tr } (\Sigma_{44.3} - \Sigma_{41.3} \Sigma_{11.3}^{-1} \Sigma_{14.3})^{-1} (\Sigma_{41.3} \Sigma_{11.3}^{-1} \Theta_1) \\
&\cdot G(\Sigma_{41.3} \Sigma_{11.3}^{-1} \Theta_1)'.
\end{align*}
\]

which is obtained in the same way as in (3.11). Substituting

\[
\Sigma_{aa}^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{33}^{-1} \end{pmatrix} + \begin{pmatrix} -1 & \Sigma_{31}^{-1} \\ -\Sigma_{33}^{-1} \Sigma_{31} \end{pmatrix} \Sigma_{11.3}^{-1} \begin{pmatrix} -1 & \Sigma_{33}^{-1} \\ -\Sigma_{33}^{-1} \Sigma_{31} \end{pmatrix}'
\]

into (3.12), we can see that (3.12) is equivalent to

\[
(3.14) \quad \Theta_2 - \Sigma_{21.3} \Sigma_{11.3}^{-1} \Theta_1 = 0 \quad \text{and} \quad \Sigma_{41.3} \Sigma_{11.3}^{-1} \Theta_1 = 0.
\]

The equivalence of (3.12) and (3.14) is shown by using

\[
\Sigma_{11.3}^{-1} = \Sigma_{11.3}^{-1} + \Sigma_{11.3}^{-1} \Sigma_{14.3} (\Sigma_{44.3} - \Sigma_{41.3} \Sigma_{11.3}^{-1} \Sigma_{14.3})^{-1} \Sigma_{41.3} \Sigma_{11.3}^{-1}.
\]

This completes the proof.

4. Tests for redundancy

4.1 A canonical form

We consider the problem of testing the hypothesis that \( y_b = (y_i' \ y_i')' \) is redundant, based on random samples \( y_{1i}, \ldots, y_{ki} \) of sizes \( N_i \) from \( \Pi_i, \ i = 1, \ldots, k + 1 \). It is assumed that the samples are multivariate normal. Let \( W, B \) and \( T = W + B \) be the matrices of sums of squares and products due to within-groups, between-groups and the total variation, respectively, i.e.,

\[
\begin{align*}
B &= \sum_{i=1}^{k+1} N_i (\overline{y}_{(i)} - \overline{y})(\overline{y}_{(i)} - \overline{y})', \\
W &= \sum_{i=1}^{k+1} \sum_{j=1}^{N_i} (y_{ji} - \overline{y}_{(i)}) (y_{ji} - \overline{y}_{(i)})',
\end{align*}
\]

where \( \overline{y}_{(i)} = (1/N_i) \sum_{j=1}^{N_i} y_{ji} \) and \( \overline{y} = (1/N) \sum_{i=1}^{k+1} N_i \overline{y}_{(i)} \) with \( N = N_1 + \cdots + N_{k+1} \). Then \( W \) and \( B \) are independently distributed as a central Wishart distribu-
tion $W_{p+q}(\Sigma, n)$ and a noncentral Wishart distribution $W_{p+q}(\Sigma, k; N\Omega)$, where $n = N - k - 1$ and $\Omega$ is defined by (3.3) with $g_i = N_i / N$. We can write

$$B = XX', \quad W = UU',$$

where the columns of $(X \quad U)$: $(p + q) \times (k + n)$ are independently distributed as $N_{p+q}(\cdot, \Sigma)$ with

$$E(X \quad U) = (v \quad 0).$$

Here $v$ satisfies $vv' = N\Omega$. Using the same partitions for $X$, $U$ and $v$ as in (3.5), we can write the hypothesis on the redundancy of $y_b$ as

$$H: v_b - \Sigma_{ba} \Sigma_{aa}^{-1} v_a = 0.$$

Here we note that $v_3 = 0$ and $v_4 = 0$. From Theorem 3.1 the hypothesis (4.1) can be decomposed as

$$H_1: v_2 - \Sigma_{21.2} \Sigma_{11.2}^{-1} v_1 = 0 \quad \text{and} \quad H_2: \Sigma_{41.3} \Sigma_{11.3}^{-1} v_1 = 0.$$

We shall obtain the LR test for $H$ by using a conditional approach. It is easily seen that

$$E \left[ \begin{pmatrix} X_2 & U_2 \end{pmatrix} \begin{pmatrix} X_1 & U_1 \\ X_2 & U_2 \end{pmatrix} \right] = (v_2^* \quad 0) + (\Gamma_1 \quad \Gamma_2) \begin{pmatrix} X_1 & U_1 \\ X_2 & U_2 \end{pmatrix},$$

and

$$E \left[ \begin{pmatrix} X_4 & U_4 \end{pmatrix} \begin{pmatrix} X_1 & U_1 \\ X_3 & U_3 \end{pmatrix} \right] = (\beta_1 \quad \beta_3) \begin{pmatrix} X_1 - v_1 & U_1 \\ X_3 & U_3 \end{pmatrix},$$

where $v_2^* = v_2 - \Gamma_1 v_1$,

$$(\Gamma_1 \quad \Gamma_2) = (\Sigma_{21} \quad \Sigma_{22}) \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1},$$

$$(\beta_1 \quad \beta_3) = (\Sigma_{41} \quad \Sigma_{43}) \begin{pmatrix} \Sigma_{11} & \Sigma_{13} \\ \Sigma_{31} & \Sigma_{33} \end{pmatrix}^{-1}.$$  

Noting that $\Gamma_1 = \Sigma_{21.2} \Sigma_{11.2}^{-1}$ and $\beta_1 = \Sigma_{41.3} \Sigma_{11.3}^{-1}$, we can write (4.2) as

$$H_1: v_2^* = 0 \quad \text{and} \quad H_2: \beta_1 v_1 = 0.$$
Therefore, the LR criterion $\lambda$ for $H$ can be decomposed as

\begin{equation}
\lambda = \lambda_1 \lambda_2 ,
\end{equation}

where

$\lambda_1 = $ the LR criterion for $H_1$ based on the conditional density

$ f \left( X_2, U_2 \left| \begin{pmatrix} X_1 \\ X_2 \\ U_1 \\ U_2 \end{pmatrix} \right. \right) $

and

$\lambda_2 = $ the LR criterion for $H_2$ based on the conditional density

$ f \left( \begin{pmatrix} X_1 \\ X_2 \\ U_4 \end{pmatrix} \mid X_3, U_3 \right) = f \left( X_4, U_4 \left| \begin{pmatrix} X_1 \\ X_3 \\ U_1 \\ U_3 \end{pmatrix} \right. \right) f(X_1, U_1 \mid X_3, U_3) . $

The explicit formulas for $\lambda_1$ and $\lambda_2$ are given in Sections 4.2 and 4.3, respectively.

### 4.2 The LR Criterion $\lambda_1$

The problem of testing $H_1: \nu^* = 0$ in the conditional model of $\begin{pmatrix} X_2 \\ U_2 \end{pmatrix}$ given $\begin{pmatrix} X_1 \\ X_2 \\ U_4 \end{pmatrix}$ is the one of testing a linear hypothesis in a multivariate linear model. Therefore, we obtain

\begin{equation}
A_1 = \lambda_1^{2/N} = \frac{|W_{22, 12}|}{|T_{22, 12}|}
\end{equation}

where $W_{22, 12} = W_{22} - W_{2(12)}W_{1(12)}^{-1}W_{12}$ and $T_{22, 12} = T_{22} - T_{2(12)}T_{1(12)}^{-1}T_{1(12)}$. Here we use the same partitions for $W$ and $T$ as in (2.1) and (2.3). The conditional null distribution of $A_1$ is the Wilks lambda distribution $A_p(k, n - p_1 - q)$ which does not depend on $(X_1 \ U_1)$ and $(X_2 \ U_2)$. Therefore, the null distribution of $A_1$ is $A_p(k, n - p_1 - q)$ and is independent of $\lambda_2$. The limiting null distribution of $- \log A_1$ is a chi-squared distribution with $kp_2$ degrees of freedom.

### 4.3 The LR criterion $\lambda_2$

First we consider the maximum of $f \left( \begin{pmatrix} X_1 \\ X_4 \\ U_1 \\ U_4 \end{pmatrix} \mid X_3, U_3 \right)$ when the parameters are unrestricted. The conditional density can be written as
\[
(4.8) \quad f(X_1, U_1 | X_2, U_2) f(X_4, U_4 | X_3, U_3) \\
= (2\pi)^{-\frac{(p_1+p_2)N^2}{2}} |\Sigma_{11,2}|^{-N/2} \\
\times \text{etr} \left\{ -\frac{1}{2} \Sigma_{11,2}^{-1} [(X_1 - \nu_1) U_1] \\
- \Sigma_{12} \Sigma_{22}^{-1} (X_2 U_2) [(X_1 - \nu_1) U_1] \\
- \Sigma_{12} \Sigma_{22}^{-1} (X_2 U_2)' \right\} \\
\times |\Sigma_{44,3}|^{-N/2} \text{etr} \left\{ -\frac{1}{2} \Sigma_{44,3}^{-1} [(X_4 U_4) - \Sigma_{43} \Sigma_{33}^{-1} (X_3 U_3)] \\
\times [(X_4 U_4) - \Sigma_{43} \Sigma_{33}^{-1} (X_3 U_3)'] \right\}.
\]

This expression implies that

\[
(4.9) \quad \max f \left( \begin{pmatrix} X_1 \\ X_4 \end{pmatrix} \begin{pmatrix} U_1 \\ U_4 \end{pmatrix} \right) | X_3, U_3 = \text{const.} \times \{ |W_{11,2}| | T_{44,3}| \}^{-N/2}.
\]

Next we consider the maximum of \( f \left( \begin{pmatrix} X_1 \\ X_4 \end{pmatrix} \begin{pmatrix} U_1 \\ U_4 \end{pmatrix} \right) | X_3, U_3 \) when \( H_2 \) is true.

Under \( H_2 \) we can write the conditional density as

\[
(4.10) \quad f \left( X_4, U_4 \left| \begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \begin{pmatrix} U_1 \\ U_3 \end{pmatrix} \right. \right) f(X_1, U_1 | X_3, U_3) \\
= (2\pi)^{-\frac{(p_1+p_2)N^2}{2}} |\Sigma_{44,13}|^{-N/2} \\
\times \text{etr} \left\{ -\frac{1}{2} \Sigma_{44,13}^{-1} \left[ (X_4 U_4) - (\beta_1 \beta_3) \begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \begin{pmatrix} U_1 \\ U_3 \end{pmatrix} \right] \right. \\
\times \left. \left[ (X_4 U_4) - (\beta_1 \beta_3) \begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \begin{pmatrix} U_1 \\ U_3 \end{pmatrix} \right]' \right\} \\
\times |\Sigma_{11,3}|^{-N/2} \text{etr} \left\{ -\frac{1}{2} \Sigma_{11,3}^{-1} [(X_1 - \nu_1) U_1] \\
- \Sigma_{13} \Sigma_{33}^{-1} (X_3 U_3) \\
\times [(X_4 - \nu_1 U_1) - \Sigma_{13} \Sigma_{33}^{-1} (X_3 U_3)] \right\}.
\]

Considering the maximization of (4.10) with respect to \( \Sigma_{44,13} \) and \( \Sigma_{11,3} \), it is easily seen that
\[
A_2 = \frac{\lambda_2^{2/N}}{\text{Min}_{\beta, \nu, \zeta} g(\beta_1, \beta_3, \nu_1, \zeta)}
\]

where \(\zeta = \Sigma_{13} \Sigma_{33}^{-1}\),

\[
\begin{align*}
\text{(4.12)} \quad g(\beta_1, \beta_3, \nu_1, \zeta) &= |\{(X_4 - U_4 - \beta_1(X_1 + U_1) - \beta_3(X_3 + U_3)\} | \\
&\quad \times |(X_4 - U_4 - \beta_1(X_1 + U_1) - \beta_3(X_3 + U_3))' | \\
&\quad \times |(X_1 - \nu_1 - U_1 - \zeta(X_3 + U_3)) | \\
&\quad \times |(X_1 - \nu_1 - U_1 - \zeta(X_3 + U_3))' | .
\end{align*}
\]

Considering the minimization of (4.12) with respect to \(\beta_3\) and \(\zeta\), we have

\[
\text{(4.13)} \quad \text{Min}_{\beta, \nu, \zeta} g(\beta_1, \beta_3, \nu_1, \zeta)
\]

\[
= \text{Min}_{\beta, \nu, \zeta} | T_{44}.3 - \beta_1 T_{14}.3 - T_{41}.3 \beta_1' + \beta_1 T_{11}.3 \beta_1' | \\
\quad \times | T_{11}.3 - \nu_1 (X_1 - T_{13} T_{33}^{-1} X_3)' - (X_1 - T_{13} T_{33}^{-1} X_3) \nu_1' | \\
\quad + \nu_1 (I - X_3' T_{33}^{-1} X_3) \nu_1' .
\]

The problem of minimizing (4.13) is discussed in the Appendix. Using Lemma A.3 in the Appendix and noting that \(W_{11,3} = T_{11,3} - (X_1 - T_{13} T_{33}^{-1} X_3)(I - X_3' T_{33}^{-1} X_3)^{-1}(I - X_3' T_{33}^{-1} X_3)'\), we obtain

\[
\text{(4.14)} \quad \text{Min}_{\beta, \nu, \zeta} g(\beta_1, \beta_3, \nu_1, \zeta) = | W_{11,3} || T_{44,3} | \left( \prod_{i=p_1 - s_0 + 1}^{p_1} l_i \right),
\]

where \(l_1 \geq \cdots \geq l_{p_1 - s_0} > 1 \geq l_{p_1 - s_0 + 1} \geq \cdots \geq l_{p_1} > 0\) are the eigenvalues of \(T_{11,3} W_{11,3}^{-1}\). Hence

\[
\text{(4.15)} \quad A_2 = \frac{| W_{11,3} || T_{44,3} |}{| W_{11,3} || T_{44,3} | \prod_{i=p_1 - s_0 + 1}^{p_1} l_i} = \frac{\prod_{i=1}^{p_1} (1 - \rho_i^2)}{\prod_{i=p_1 - s_0 + 1}^{p_1} l_i},
\]

where \(\rho_1^2 \geq \cdots \geq \rho_{p_1}^2\) are the eigenvalues of \(W_{14,3} W_{44,3}^{-1} W_{41,3} W_{11,3}^{-1}\). Unfortunately, the distribution of \(A_2\) appears to be extremely complicated even under the null hypothesis \(H_2\).

For practical purposes we consider an approximate test. Starting the conditional model (4.4), we can use the test statistic

\[
\text{(4.16)} \quad A_{(2)} = | W_{44,13} | / | T_{44,13} |
\]
instead of (4.15). It is easily seen that the null distribution of $A(2)$ is the Wilks lambda distribution $A_q(k, n - p_1 - q_2)$. This test will not be efficient because it ignores the information on the mean value of $X_4$ and the conditional density of $[X_1, U_3]$ given $[X_3, U_3]$. An approximate test for $H$ is to use the statistic $-n \log A_1 A(2)$ whose asymptotic null distribution is a chi-squared distribution with $k(p_2 + q_2)$ degrees of freedom.

We shall rewrite (4.15) in an alternative form to point out the connection with (4.16). For this, we note that

$$ |T_{11.34}|/|W_{11.3}| = \prod_{i=1}^{p_1} l_i $$

$$ = \begin{vmatrix} T_{11.3} & T_{14.3} \\ T_{41.3} & T_{44.3} \end{vmatrix} / \{ |W_{11.3} T_{44.3}| \} = \frac{|T_{11.3}| |T_{44.3}|}{|T_{11.3}| |T_{44.3}|} $$

$$ = \left[ \prod_{i=1}^{p_1} (1 - \rho_i^2) \right] (|W_{44.3}| / |T_{44.3}|) $$

$$ \times \left( \begin{vmatrix} T_{11.3} & T_{14.3} \\ T_{41.3} & T_{44.3} \end{vmatrix} / \begin{vmatrix} W_{11.3} & W_{14.3} \\ W_{41.3} & W_{44.3} \end{vmatrix} \right). $$

Hence, we can rewrite (4.15) as

$$ A_2 = \left( \prod_{i=1}^{p_1} l_i \right) \left( \frac{|W_{44.3}|}{|W_{44.3}|} \right) \left( \frac{|W_{11.3}|}{|T_{11.3}|} \right) A(2) = \zeta A(2), \quad \text{say}. $$

Note that the quantity $\zeta$ depends on $(X_4, U_3)$ and hence $\zeta$ and $A(2)$ are dependent variables. It is not clear whether the distribution of $\zeta$ will depend on the unknown parameters under $H_2$ or not. For this, we require simulation study which will be done at a later stage.

Acknowledgements

The authors would like to express their deep thanks to Professor C. R. Rao, Pennsylvania State University for his encouragement and helpful discussion during this investigation. Thanks are also due to the referees for their useful comments.

Appendix

We consider the minimization problem which is related to (4.13). Let $Q$ and $T$ be the positive definite matrices of orders $m + p_1$ and $p_1 + p_2$, respectively, and let
\[ Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \quad Q_{11}: m \times m, \quad T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}, \quad T_{11}: p_1 \times p_1. \]

**Lemma A.1.** Let \( \beta \) and \( \nu \) be \( p_2 \times p_1 \) and \( p_1 \times m \) matrices such that \( \beta \nu = 0 \). The minimum value of \( |Q_{22} - \nu Q_{12} - Q_{21} \nu' + \nu Q_{11} \nu'| \) over \( \nu \) subject to \( \beta \nu = 0 \) for a given \( \beta \) is

\[
|Q_{22,1}||\beta_0 Q_{22} \beta_0' | / |\beta_0 Q_{22} \beta_0' |,
\]

where \( \beta = \beta_1 \beta_0 \), \( \beta_1 \) and \( \beta_0 \) are \( p_2 \times s \) and \( s \times p_1 \) matrices of ranks \( s \), \( s = \text{rank} \beta \), and \( Q_{22,1} = Q_{22} - Q_{21} Q_{11}^{-1} Q_{12} \).

**Proof.** Note that \( \beta \nu = 0 \iff \beta_0 \nu = 0 \). Using Lagrangian multipliers, we have to minimize

\[
\phi = \log |Q_{22} - \nu Q_{12} - Q_{21} \nu' + \nu Q_{11} \nu'| + 2 \text{ tr } \Lambda \beta_0 \nu .
\]

Differentiating \( \phi \) with respect to \( \nu \) and \( \Lambda \), we get

\[
(Q_{22} - \nu Q_{12} - Q_{21} \nu' + \nu Q_{11} \nu')^{-1} (Q_{21} - \nu Q_{11}) = \beta_0 A',
\]

\[
\beta_0 \nu = 0
\]

which are equivalent to

\[
(A.1) \quad Q_{21} - \nu Q_{11} = (Q_{22} - \nu Q_{12}) \beta_0 A', \quad \beta_0 \nu = 0.
\]

Premultiplying both sides of the first equation of (A.1) by \( \beta_0 \), we have \( A' = (\beta_0 Q_{22} \beta_0')^{-1} \beta_0 Q_{21} \). Hence (A.1) gives

\[
\nu = [I - Q_{22} \beta_0 (\beta_0 Q_{22} \beta_0')^{-1} \beta_0] Q_{21} [Q_{11} - Q_{12} \beta_0 (\beta_0 Q_{22} \beta_0')^{-1} \beta_0 Q_{21}]^{-1}
\]

\[
= [I - Q_{22,1} \beta_0 (\beta_0 Q_{22,1} \beta_0')^{-1} \beta_0] Q_{21} Q_{11}^{-1}
\]

because \( Q_{21} [Q_{11} - Q_{12} \beta_0 (\beta_0 Q_{22} \beta_0')^{-1} \beta_0 Q_{21}]^{-1} = Q_{21} Q_{11}^{-1} + Q_{21} Q_{11}^{-1} Q_{12} \beta_0 \cdot (\beta_0 Q_{22,1} \beta_0')^{-1} \beta_0 Q_{21} Q_{11}^{-1} \). Substituting this into \( |Q_{22} - \nu Q_{12} - Q_{21} \nu' + \nu Q_{11} \nu'| \), we have

\[
\min_{\beta_0 \nu = 0} |Q_{22} - \nu Q_{12} Q_{21} \nu' + \nu Q_{11} \nu' |
\]

\[
= |Q_{22,1} + Q_{22,1} \beta_0 (\beta_0 Q_{22,1} \beta_0')^{-1} \beta_0 (Q_{21} Q_{11}^{-1} Q_{12})
\times \beta_0 (\beta_0 Q_{22,1} \beta_0')^{-1} \beta_0 Q_{22,1} |
\]

\[
= |Q_{22,1} ||I + (\beta_0 Q_{22,1} \beta_0')^{-1} \beta_0 Q_{21} Q_{11}^{-1} Q_{12} \beta_0|
\]
which proves Lemma A.1.

**Lemma A.2.** Let \( \beta = \beta_1 \beta_0, \beta_1 \) and \( \beta_0 \) be matrices of orders \( p_2 \times p_1, \) \( p_2 \times s, s \times p_1 \) and of ranks \( s. \) Then, given \( \beta_0, \) the minimum of \( |T_{22} - T_{12} - T_{21} \beta' + \beta T_{11} \beta'| \) is \( |T_{22} - T_{21} \beta_0 (\beta_0 T_{11} \beta_0)^{-1} \beta_0 T_{12}| = |T_{22}| |\beta_0 T_{11}^{-1} \beta_0| / |\beta_0 T_{11} \beta_0| \) with \( T_{11}^{-1} = T_{11} - T_{12} T_{22}^{-1} T_{21}. \)

**Proof.** The result follows from the identity

\[
T_{22} - \beta T_{12} - T_{21} \beta' + \beta T_{11} \beta' = \{T_{21} \beta_0 (\beta_0 T_{11} \beta_0)^{-1} - \beta_1\} (\beta_0 T_{11} \beta_0)^{-1} \beta_1' + T_{22} - T_{21} \beta_0 (\beta_0 T_{11} \beta_0)^{-1} \beta_0 T_{12}.
\]

**Lemma A.3.** Suppose that \( Q_{22} = T_{11}. \) Then,

\[
(A.2) \quad \min_{\beta_1} |T_{22} - \beta T_{12} - T_{21} \beta' + \beta T_{11} \beta'| |Q_{22} - \nu Q_{12} - Q_{21} \nu' + \nu Q_{11} \nu| = |T_{22}| |Q_{22,1}| \prod_{i=p_1-s_0+1}^{p_1} l_i,
\]

where \( l_1 \geq \cdots \geq l_{p_1-s_0} > 1 \geq l_{p_1-s_0+1} \geq \cdots \geq l_{p_1} > 0 \) are the eigenvalues of \( T_{11}^{-1} Q_{22,1}. \)

**Proof.** Let \( s = \text{rank } \beta. \) Then, using Lemmas A.1 and A.2, we can write the left-hand side of (A.2) as

\[
\min_{s, \beta_0} \frac{|T_{22}| |\beta_0 T_{11}^{-1} \beta_0|}{|\beta_0 T_{11} \beta_0|} \cdot \frac{|Q_{22,1}| |\beta_0 Q_{22} \beta_0|}{|\beta_0 Q_{22,1} \beta_0|} = \min_s |T_{22}| |Q_{22,1}| \prod_{i=p_1-s_0+1}^{p_1} l_i
\]

which is equal to the right-hand side of (A.2).

**References**


