A GENERALIZATION OF TESTING INDEPENDENCE OF SETS OF VARIATES

Akio Suzukawa* and Yoshiharu Sato*

Assuming that $X = (X^{(1)'}, X^{(2)'}, \dots, X^{(q)'})'$ is distributed according to $N_p(\mu, \Sigma)$ and each $X^{(j)}$ has p_j components, where $p = p_1 + \dots + p_q$ and $p_1 \geq p_2 \geq \dots \geq p_q$, we consider the testing of the hypothesis that for $a = 2, \dots, q$, all the smallest $p_a - k_a$ canonical correlations between $(X^{(1)'}, X^{(2)'}, \dots, X^{(a-1)'})'$ and $X^{(a)'}$ are zero. In the case of $k_2 = k_3 = \dots = k_q = 0$, this hypothesis is equivalent to the hypothesis of independence of $X^{(1)}, X^{(2)}, \dots, X^{(q)}$. This testing problem can then be considered to be a generalization of the problem of testing independence. In this paper we derive the likelihood ratio statistic (LR statistic) and obtain the asymptotic expansion of its null distribution. We also obtain the asymptotic expansions of the null distributions of the other two statistics, which are the statistics corresponding to Lawley-Hotelling's trace and the Bartlett-Nanda-Pillai's trace criterion.

1. Introduction

Let $\mathbf{X}$ be p -variate random vector distributed according to $N_p(\mu, \Sigma)$. We partition $\mathbf{X}$ into q subvectors with $p_1, p_2, \dots, p_q$ components, respectively, as follows;

$$\mathbf{X} = (\mathbf{X}^{(1)'}, \mathbf{X}^{(2)'}, \dots, \mathbf{X}^{(q)'})',$$

where $p_1 \geq p_2 \geq \dots \geq p_q$. The testing problem where $H: \mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(q)}$ are mutually independently distributed against $K: \text{not } H$ is well-known as a test of independence of sets of variates (see, Anderson [1], Murhead [6]). When the covariance matrix Σ is partitioned similarly:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1q} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2q} \\ \cdots & \cdots & \cdots & \cdots \\ \Sigma_{q1} & \Sigma_{q2} & \cdots & \Sigma_{qq} \end{bmatrix},$$

the null hypothesis H can be expressed as $H: \Sigma_{ij} = \mathbf{0}, i \neq j$.

Let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ be a random sample of size $N (\geq p+1)$ drawn from $N_p(\mu, \Sigma)$ and let $\bar{\mathbf{X}} = 1/N \sum_{\alpha=1}^N \mathbf{X}_\alpha$ and

$$\begin{aligned} \mathbf{W} &= \sum_{\alpha=1}^N (\mathbf{X}_\alpha - \bar{\mathbf{X}})(\mathbf{X}_\alpha - \bar{\mathbf{X}})' \\ &= \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} & \cdots & \mathbf{W}_{1q} \\ \mathbf{W}_{21} & \mathbf{W}_{22} & \cdots & \mathbf{W}_{2q} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{W}_{q1} & \mathbf{W}_{q2} & \cdots & \mathbf{W}_{qq} \end{bmatrix}. \end{aligned}$$

Received February, 1994. Revised July, 1994. Accepted August, 1994.

* Division of Information Engineering, Graduate School of Hokkaido University; Kita 13, Nishi 8, Kita-ku, Sapporo, 060, Japan

The likelihood ratio criterion Λ_0 for testing H against K is given by $\Lambda_0^{2/N} = |\mathbf{W}| / \prod_{i=1}^q |\mathbf{W}_{ii}|$. Now we wish to test

$$H_{k_2, \dots, k_q}: \text{rank}(\boldsymbol{\Sigma}_{j(j-1)}) = k_j, j=2, \dots, q, \\ \text{against } K_{k_2, \dots, k_q}: \text{rank}(\boldsymbol{\Sigma}_{j(j-1)}) > k_j, \text{ for some } j,$$

where $\sum_{j=2}^q k_j < \sum_{j=2}^q p_j$ and $\boldsymbol{\Sigma}_{j(j-1)} = [\boldsymbol{\Sigma}_{j1}, \boldsymbol{\Sigma}_{j2}, \dots, \boldsymbol{\Sigma}_{j, j-1}]$. The null hypothesis $H_{k_2, \dots, k_q}$ is equivalent to the hypothesis that the smallest $p_2 - k_2$ canonical correlations between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, the smallest $p_3 - k_3$ canonical correlations between $(\mathbf{X}^{(1)'}, \mathbf{X}^{(2)'})'$ and $\mathbf{X}^{(3)}$, $\dots$, the smallest $p_q - k_q$ canonical correlations between $(\mathbf{X}^{(1)'}, \mathbf{X}^{(2)'}, \dots, \mathbf{X}^{(q-1)'})'$ and $\mathbf{X}^{(q)}$ are all zero. Testing H against K is a special case of the problem of testing $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$, because in the case of $k_2 = k_3 = \dots = k_q = 0$, the hypothesis $H_{k_2, \dots, k_q}$ and $K_{k_2, \dots, k_q}$ are equivalent to H and K , respectively. In the case of $q=2$, this problem has been considered as a test of dimensionality in canonical correlation analysis (see Fujikoshi [3], Fujikoshi [4], Glynn and Muirhead [5]).

In this paper, the likelihood ratio statistic for testing $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$, and the asymptotic expansion of its null distribution are discussed. We also propose two other statistics for testing $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$, and obtain the asymptotic expansions of their null distributions.

2. LR test for $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$

We prepare the following lemma (proved by Fujikoshi [2]) to derive the LR statistic for testing of $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$.

LEMMA 1 (Fujikoshi [2]). Let $\mathbf{L}$ be $u \times v$ matrix, $\mathbf{Q}$ be $v \times v$ positive definite matrix, $\hat{\mathbf{S}} = \mathbf{Q}^{-1/2} \mathbf{L}' \mathbf{L} \mathbf{Q}^{-1/2}$, $\text{rank}(\hat{\mathbf{S}}) \geq r$, and $\hat{\mathbf{S}}(\mathbf{M}) = \mathbf{Q}^{-1/2} (\mathbf{L} - \mathbf{M})' (\mathbf{L} - \mathbf{M}) \mathbf{Q}^{-1/2}$. Then

$$\min_{\text{rank}(\mathbf{M})=r} |\mathbf{I}_v + \mathbf{S}(\mathbf{M})| = \prod_{i=r+1}^v (1 + d_i^2),$$

where $d_1^2 \geq d_2^2 \geq \dots \geq d_v^2$ are the latent root of $\hat{\mathbf{S}}$.

Applying this lemma, we obtain the following theorem.

THEOREM 1. Based on a random sample $\mathbf{X}_1, \dots, \mathbf{X}_N$ drawn from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the LR criterion Λ for testing $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$ is given by

$$(2.1) \quad \Lambda^{2/N} \equiv V = \prod_{j=2}^q \prod_{i=k_j+1}^{p_j} (1 - l_{ji}^2),$$

where l_{ji} is the i -th sample canonical correlation between $\mathbf{X}^{(j)}$ and $(\mathbf{X}^{(1)'}, \dots, \mathbf{X}^{(j-1)'})'$.

PROOF. The likelihood function is

$$L(\boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}|^{-N/2} \exp \left\{ -\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{W}) \right\}.$$

The maximum of $L(\boldsymbol{\Sigma})$ with respect to $\boldsymbol{\Sigma} > 0$ (positive definite) is

$$(2.2) \quad \max_{\boldsymbol{\Sigma} > 0} \{L(\boldsymbol{\Sigma})\} = |\mathbf{W}|^{-N/2} N^{pN/2} \exp \left(-\frac{pN}{2} \right).$$

To find the maximum of $L(\mathbf{Z})$ under $H_{k_2, \dots, k_q}$, we write $L(\mathbf{Z})$ as

$$L(\mathbf{Z}) = \{ |\mathbf{Z}_{11}| |\mathbf{Z}_{22,1}| |\mathbf{Z}_{33,(2)}| |\mathbf{Z}_{44,(3)}| \cdots |\mathbf{Z}_{qq,(q-1)}| \}^{-N/2} \\ \times \exp \left[-\frac{1}{2} \{ \text{tr}(\mathbf{Z}_{11}^{-1} \mathbf{W}_{11}) + \text{tr}(\mathbf{Z}_{22,1}^{-1} \tilde{\mathbf{W}}_{22,1}) + \cdots + \text{tr}(\mathbf{Z}_{qq,(q-1)}^{-1} \tilde{\mathbf{W}}_{qq,(q-1)}) \} \right],$$

where

$$\mathbf{Z}_{jj,(j-1)} = \mathbf{Z}_{jj} - \mathbf{Z}_{j(j-1)} \mathbf{Z}_{(j-1)(j-1)}^{-1} \mathbf{Z}_{(j-1)j}, \\ \mathbf{Z}_{(j-1)(j-1)} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} & \cdots & \mathbf{Z}_{1,j-1} \\ \mathbf{Z}_{21} & \mathbf{Z}_{22} & \cdots & \mathbf{Z}_{2,j-1} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{Z}_{j-1,1} & \mathbf{Z}_{j-1,2} & \cdots & \mathbf{Z}_{j-1,j-1} \end{bmatrix}, \\ \mathbf{Z}_{j(j-1)} = [\mathbf{Z}_{j1}, \mathbf{Z}_{j2}, \cdots, \mathbf{Z}_{j,j-1}], \quad \mathbf{Z}_{(j-1)j} = \mathbf{Z}'_{j(j-1)}, \\ \mathbf{B}_{(j-1)j} = \mathbf{Z}_{(j-1)(j-1)}^{-1} \mathbf{Z}_{(j-1)j}, \\ \mathbf{W}_{jj,(j-1)} = \mathbf{W}_{jj} - \mathbf{W}_{j(j-1)} \mathbf{W}_{(j-1)(j-1)}^{-1} \mathbf{W}_{(j-1)j}, \\ \mathbf{W}_{(j-1)(j-1)} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} & \cdots & \mathbf{W}_{1,j-1} \\ \mathbf{W}_{21} & \mathbf{W}_{22} & \cdots & \mathbf{W}_{2,j-1} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{W}_{j-1,1} & \mathbf{W}_{j-1,2} & \cdots & \mathbf{W}_{j-1,j-1} \end{bmatrix}, \\ \mathbf{W}_{j(j-1)} = [\mathbf{W}_{j1}, \mathbf{W}_{j2}, \cdots, \mathbf{W}_{j,j-1}], \quad \mathbf{W}_{(j-1)j} = \mathbf{W}'_{j(j-1)}, \\ \tilde{\mathbf{W}}_{jj,(j-1)} = \mathbf{W}_{jj,(j-1)} + \{ \mathbf{W}_{(j-1)(j-1)}^{-1} \mathbf{W}_{(j-1)j} - \mathbf{B}_{(j-1)j} \}' \\ \times \mathbf{W}_{(j-1)(j-1)} \{ \mathbf{W}_{(j-1)(j-1)}^{-1} \mathbf{W}_{(j-1)j} - \mathbf{B}_{(j-1)j} \}.$$

The maximization of

$$L(\mathbf{Z}) \equiv L(\mathbf{Z}_{11}, \mathbf{Z}_{22,1}, \cdots, \mathbf{Z}_{qq,(q-1)}, \mathbf{B}_{12}, \mathbf{B}_{(2)3}, \cdots, \mathbf{B}_{(q-1)q})$$

with respect to $\mathbf{Z}_{11} > 0, \mathbf{Z}_{22,1} > 0, \mathbf{Z}_{qq,(q-1)} > 0$ yields

$$\hat{\mathbf{Z}}_{11} = \frac{1}{N} \mathbf{W}_{11}, \quad \hat{\mathbf{Z}}_{jj,(j-1)} = \frac{1}{N} \tilde{\mathbf{W}}_{jj,(j-1)}, \quad j = 2, 3, \cdots, q,$$

so that

$$L^*(\mathbf{B}_{12}, \mathbf{B}_{(2)3}, \cdots, \mathbf{B}_{(q-1)q}) \equiv \max_{\mathbf{Z}_{11} > 0, \mathbf{Z}_{22,1} > 0, \cdots, \mathbf{Z}_{qq,(q-1)} > 0} \{L(\mathbf{Z})\} \\ \equiv \{ |\mathbf{W}_{11}| |\tilde{\mathbf{W}}_{22,1}| \cdots |\tilde{\mathbf{W}}_{qq,(q-1)}| \}^{-N/2} N^{pN/2} \exp \left(-\frac{pN}{2} \right).$$

Thus L^* can be written as

$$L^* = |\mathbf{W}|^{-N/2} N^{pN/2} \exp \left(-\frac{pN}{2} \right) \\ \times \prod_{j=2}^q [|\mathbf{I}_{p_j} + \mathbf{W}_{jj,(j-1)}^{-1/2} \{ \mathbf{W}_{(j-1)(j-1)}^{-1} \mathbf{W}_{(j-1)j} - \mathbf{B}_{(j-1)j} \}' \\ \times \mathbf{W}_{(j-1)(j-1)} \{ \mathbf{W}_{(j-1)(j-1)}^{-1} \mathbf{W}_{(j-1)j} - \mathbf{B}_{(j-1)j} \} \mathbf{W}_{jj,(j-1)}^{-1/2} |]^{-N/2}.$$

The null hypothesis $H_{k_2, \dots, k_q}$ is equivalent to

$$H_{k_2, \dots, k_q}: \text{rank}(\mathbf{B}_{12}) = k_2, \text{rank}(\mathbf{B}_{(2)3}) = k_3; \cdots, \text{rank}(\mathbf{B}_{(q-1)q}) = k_q.$$

Applying LEMMA 1, we have

$$\max_{H_{k_2, \dots, k_q}} \{L(\Sigma)\} = |\mathbf{W}|^{-N/2} N^{pN/2} \exp\left(-\frac{pN}{2}\right) \left\{ \prod_{j=2}^q \prod_{i=k_j+1}^{p_j} (1+d_{ji}) \right\}^{-N/2},$$

where d_{ji} is the i -th latent root of $\mathbf{W}_{jj}^{-1} \mathbf{W}_{j(j-1)} \mathbf{W}_{(j-1)}^{-1} \mathbf{W}_{(j-1)j}$. Since l_{ji}^2 is the i -th latent root of $\mathbf{W}_{jj}^{-1} \mathbf{W}_{j(j-1)} \mathbf{W}_{(j-1)}^{-1} \mathbf{W}_{(j-1)j}$, $d_{ji} = l_{ji}^2 / (1 - l_{ji}^2)$. Thus we have

$$(2.3) \quad \max_{H_{k_2, \dots, k_q}} \{L(\Sigma)\} = |\mathbf{W}|^{-N/2} N^{pN/2} \exp\left(-\frac{pN}{2}\right) \left\{ \prod_{j=2}^q \prod_{i=k_j+1}^{p_j} (1-l_{ji}^2) \right\}^{N/2}.$$

From (2.2) and (2.3), we obtain (2.1). $\square$

3. Asymptotic expansion of the null distribution of the LR statistic

Let

$$V_j = \prod_{i=k_j+1}^{p_j} (1-l_{ji}^2), \quad j=2, 3, \dots, q,$$

Then $V = V_2 V_3 \cdots V_q$, and for $j=2, 3, \dots, q$, V_j is the LR criterion for testing

$$H_{(j)}: \text{rank}(\Sigma_{j(j-1)}) = k_j \quad \text{against} \quad K_{(j)}: \text{rank}(\Sigma_{j(j-1)}) > k_j.$$

The null hypothesis $H_{k_2, \dots, k_q}$ is the intersection of $H_{(2)}, \dots, H_{(q)}$ and the likelihood ratio statistic V can be written as $V = V_2 V_3 \cdots V_q$. By using this property, we obtain the asymptotic expansion of the null distribution of the LR statistic for testing $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$.

Let $T = -\log V$ and $T_j = -\log V_j$. Then the following theorem concerning with the null distribution of T holds.

THEOREM 2. *Let $n = N - 1$ and $m = n - 2\eta$, where η is a correction factor. The null distribution of T for testing $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$ can be asymptotically approximated by*

$$(3.1) \quad \Pr(mT \leq x) \\ = \Pr(\chi_f^2 \leq x) + \frac{f}{2m} \left\{ 2\eta - \frac{1}{2}(a+1) - b + c \right\} \{ \Pr(\chi_f^2 \leq x) - \Pr(\chi_{f+2}^2 \leq x) \} \\ + o(m^{-1}),$$

where

$$(3.2) \quad f = \sum_{j=2}^q f_j, \quad f_j = (p_j - k_j)(p_1 + \cdots + p_{j-1} - k_j),$$

$$(3.3) \quad a = f^{-1} \sum_{j=2}^q \sum_{i=1}^j f_j p_i, \quad b = f^{-1} \sum_{j=2}^q f_j k_j, \quad c = f^{-1} \sum_{j=2}^q \sum_{i=1}^{k_j} f_j \lambda_{ji}^{-2},$$

and λ_{ji} is the i -th canonical correlation between $\mathbf{X}^{(j)}$ and $(\mathbf{X}^{(1)'}, \dots, \mathbf{X}^{(j-1)'})'$. Further putting $m_1 = n - 1/2(a+1) - b + c$,

$$\Pr(m_1 T \leq x) = \Pr(\chi_f^2 \leq x) + o(m_1^{-1}).$$

PROOF. First we consider the null distribution of mT_j . We can assume without loss of generality that

$$\begin{bmatrix} \mathbf{W}_{(j-1)(j-1)} & \mathbf{W}_{(j-1)j} \\ \mathbf{W}_{j(j-1)} & \mathbf{W}_{jj} \end{bmatrix} \sim W_{p_1 + \cdots + p_j} \left(n, \begin{bmatrix} \mathbf{I} & \mathbf{D}_j \\ \mathbf{D}_j' & \mathbf{I} \end{bmatrix} \right),$$

where $\mathbf{D}_j = [\tilde{\mathbf{D}}_j; \mathbf{0}]'$ and $\tilde{\mathbf{D}}_j = \text{diag}(\lambda_{j1}, \dots, \lambda_{jk_j}, 0, \dots, 0)$. We put

$$\mathbf{A}_j = (\mathbf{I} - \tilde{\mathbf{D}}_j^2)^{-1/2} \mathbf{W}_{j(j-1)} (\mathbf{I} - \tilde{\mathbf{D}}_j^2)^{-1/2}$$

$$\mathbf{B}_j = (\mathbf{I} - \tilde{\mathbf{D}}_j^2)^{-1/2} \mathbf{W}_{j(j-1)} \mathbf{W}_{(j-1)(j-1)}^{-1} \mathbf{W}_{(j-1)j} (\mathbf{I} - \tilde{\mathbf{D}}_j^2)^{-1/2}.$$

For given $\mathbf{W}_{(j-1)(j-1)}$,

$$\mathbf{A}_j \sim W_{p_j}(n, \mathbf{I}),$$

$$\mathbf{B}_j \sim W_{p_j}(p_1 + \cdots + p_{j-1}, \mathbf{I}, (\mathbf{I} - \tilde{\mathbf{D}}_j^2)^{-1/2} \mathbf{D}_j' \mathbf{W}_{(j-1)(j-1)} \mathbf{D}_j (\mathbf{I} - \tilde{\mathbf{D}}_j^2)^{-1/2})$$

and $\mathbf{A}_j$ and $\mathbf{B}_j$ are independent. T_j can be expressed as $T_j = \sum_{i=k_j+1}^p \log(1 + d_{ji})$. From the above consideration, when $\mathbf{W}_{(j-1)(j-1)}$ is fixed, we can use the result of the dimensionality test in MANOVA to obtain the asymptotic expansions of the conditional null distribution of mT_j . From Fujikoshi [4], we see that the characteristic function of the conditional null distribution of mT_j for fixed $\mathbf{W}_{(j-1)(j-1)}$ is given by

$$(3.4) \quad \phi_j(t | \mathbf{W}_{(j-1)(j-1)}) \\ = (1 - 2it)^{-f_j/2} \left[1 + \frac{f_j}{2m} \left\{ 2\eta - \frac{1}{2} \left(\sum_{i=1}^j p_i + 1 \right) + \sum_{i=1}^{k_j} \theta_{ji}^{-1} \right\} \{ 1 - (1 - 2it)^{-1} \} \right. \\ \left. + o(m^{-1}) \right],$$

where $\theta_{ji} = \lambda_{ji}^2(1 - \lambda_{ji}^2)^{-1}$. Let $\phi(t)$ be characteristic function of the null distribution of mT . Using (3.4), $\phi(t)$ can be calculated as follows;

$$(3.5) \quad \phi(t) = E^{W(q)(q)} \{ \exp(itmT) \} \\ = E^{W(q-1)(q-1)} \left[\exp \left(it \sum_{j=2}^{q-1} mT_j \right) E^{W(q)(q) | W(q-1)(q-1)} \{ \exp(itmT_q) \} \right] \\ = E^{W(q-1)(q-1)} \left[\exp \left(it \sum_{j=2}^{q-1} mT_j \right) \phi_q(t | \mathbf{W}_{(q-1)(q-1)}) \right] \\ = (1 - 2it)^{-f_q/2} \left[\left[1 + \frac{f_q}{2m} \left\{ 2\eta - \frac{1}{2} (p + 1) + \sum_{i=1}^{k_q} \theta_{qi}^{-1} \right\} \{ 1 - (1 - 2it)^{-1} \} \right] \right. \\ \left. \times E^{W(q-1)(q-1)} \left\{ \exp \left(it \sum_{j=2}^{q-1} mT_j \right) \right\} + o(m^{-1}) \right] \\ \vdots \\ = (1 - 2it)^{-f/2} \left[1 + \frac{f}{2m} \left\{ 2\eta - \frac{a+1}{2} - b + c \right\} \{ 1 - (1 - 2it)^{-1} \} + o(m^{-1}) \right],$$

where $E^{W(j)(j)}$ and $E^{W(j)(j) | W(j-1)(j-1)}$ are denoted the expectations taken with respect to the distribution of $W_{(j)(j)}$ and the conditional distribution of $W_{(j)(j)}$ for given $W_{(j-1)(j-1)}$, respectively. Inverting (3.5), we obtain (3.1). m_1 is chosen so that the second term in (3.1) is eliminated. $\square$

4. Other test statistics

In the case of $q=2$, the following two statistics T_2^* and T_2^{**} are proposed for testing the rank($\mathbf{Z}_{12}$) = k_2 against rank($\mathbf{Z}_{12}$) > k_2 ;

$$T_2^* = \sum_{i=k_2+1}^{p_2} \frac{l_{2i}^2}{1 - l_{2i}^2} = \sum_{i=k_2+1}^{p_2} d_{2i}, \quad T_2^{**} = \sum_{i=k_2+1}^{p_2} l_{2i}^2 = \sum_{i=k_2+1}^{p_2} \frac{d_{2i}}{1 + d_{2i}}.$$

T_2^* and T_2^{**} are the statistics corresponding to Lawley-Hotelling's criterion and

Bartlett-Nanda-Pillai's trace criterion in MANOVA. The asymptotic expansions of the null and nonnull distributions of T_2^* and T_2^{**} were obtained by Fujikoshi [4]. We extend T_2^* and T_2^{**} for testing $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$, and obtain the asymptotic expansions of their null distributions.

As test statistics for $H_{k_2, \dots, k_q}$ against $K_{k_2, \dots, k_q}$, we consider the formulas

$$T^* = \sum_{j=2}^q T_j^* \quad \text{and} \quad T^{**} = \sum_{j=2}^q T_j^{**},$$

where

$$T_j^* = \sum_{i=k_j+1}^{p_j} \frac{l_{ji}^2}{1-l_{ji}^2} = \sum_{i=k_j+1}^{p_j} d_{ji} \quad \text{and} \quad T_j^{**} = \sum_{i=k_j+1}^{p_j} l_{ji}^2 = \sum_{i=k_j+1}^{p_j} \frac{d_{ji}}{1+d_{ji}}.$$

The null distributions of T^* and T^{**} are given by the following theorem.

THEOREM 3. *Let $m = n - 2\eta$, where η is a correction factor. The null distributions of T^* and T^{**} can be asymptotically approximated by*

$$(4.1) \quad \Pr(mT^* \leq x) \\ = \Pr(\chi_r^2 \leq x) + \frac{f}{4m} \{ (4\eta - a - 2b + 2c - 1) \Pr(\chi_r^2 \leq x) \\ - 2(2\eta - 2b + c) \Pr(\chi_{r+2}^2 \leq x) + (a - 2b + 1) \Pr(\chi_{r+4}^2 \leq x) \} + o(m^{-1}),$$

$$(4.2) \quad \Pr(mT^{**} \leq x) \\ = \Pr(\chi_r^2 \leq x) + \frac{f}{4m} \{ (4\eta - a - 2b + 2c - 1) \Pr(\chi_r^2 \leq x) \\ - 2(2\eta - a + c - 1) \Pr(\chi_{r+2}^2 \leq x) - (a - 2b + 1) \Pr(\chi_{r+4}^2 \leq x) \} + o(m^{-1}),$$

where f and a, b, c are defined by (3.2) and (3.3), respectively. Furthermore, putting

$$(4.3) \quad m^* = n - a + c - 1 \quad \text{and} \quad m^{**} = n - 2b + c, \\ E(m^* T^*) = f + o((m^*)^{-1}) \quad \text{and} \quad E(m^{**} T^{**}) = f + o((m^{**})^{-1}).$$

PROOF. From THEOREM 2 of Fujikoshi [4], we see that the characteristic functions of the conditional null distribution of mT_j^* and mT_j^{**} for given $\mathbf{W}_{(j-1)(j-1)}$ can be expanded as

$$\phi_j^*(t | \mathbf{W}_{(j-1)(j-1)}) \\ = (1 - 2it)^{-J_j/2} \left[1 + \frac{f_j}{4m} \left\{ \left(4\eta - \sum_{l=1}^j p_l - 1 + 2 \sum_{l=1}^{k_j} \theta_{jl}^{-1} \right) - 2 \left(2\eta - k_j + \sum_{l=1}^{k_j} \theta_{jl}^{-1} \right) \right. \right. \\ \left. \left. \cdot (1 - 2it)^{-1} + \left(\sum_{l=1}^j p_l + 1 - 2k_j \right) (1 - 2it)^{-2} \right\} + o(m^{-1}) \right]$$

and

$$\phi_j^{**}(t | \mathbf{W}_{(j-1)(j-1)}) \\ = (1 - 2it)^{-J_j/2} \left[1 + \frac{f_j}{4m} \left\{ \left(4\eta - \sum_{l=1}^j p_l - 1 + 2 \sum_{l=1}^{k_j} \theta_{jl}^{-1} \right) - 2 \left(2\eta - k_j + \sum_{l=1}^{k_j} \theta_{jl}^{-1} \right. \right. \right. \\ \left. \left. - \sum_{l=1}^j p_l - 1 \right) (1 - 2it)^{-1} - \left(\sum_{l=1}^j p_l + 1 - 2k_j \right) (1 - 2it)^{-2} \right\} + o(m^{-1}) \right],$$

respectively. By a method similar to that in the proof of Theorem 2 of this paper, the characteristic functions of the null distributions of mT^* and mT^{**} can be expanded as

$$\begin{aligned} \phi^*(t) = (1-2it)^{-f/2} & \left[1 + \frac{f}{4m} \{ (4\eta - a - 2b + 2c - 1) - 2(2\eta - 2b + c)(1-2it)^{-1} \right. \\ & \left. + (a - 2b + 1)(1-2it)^{-2} \} + o(m^{-1}) \right] \end{aligned}$$

and

$$\begin{aligned} \phi^{**}(t) = (1-2it)^{-f/2} & \left[1 + \frac{f}{4m} \{ (4\eta - a - 2b + 2c - 1) - 2(2\eta - a + c - 1)(1-2it)^{-1} \right. \\ & \left. - (a - 2b + 1)(1-2it)^{-2} \} + o(m^{-1}) \right], \end{aligned}$$

respectively. Inverting $\phi^*(t)$ and $\phi^{**}(t)$, we obtain (4.1) and (4.2). Choosing m^* and m^{**} as (4.3), we have

$$(4.4) \quad \begin{aligned} \Pr(m^*T^* \leq x) = & \Pr(\chi_f^2 \leq x) + \frac{f}{4m^*} (a - 2b + 1) \{ \Pr(\chi_f^2 \leq x) \\ & - 2\Pr(\chi_{f+2}^2 \leq x) + \Pr(\chi_{f+4}^2 \leq x) \} + o((m^*)^{-1}) \end{aligned}$$

$$(4.5) \quad \begin{aligned} \Pr(m^{**}T^{**} \leq x) = & \Pr(\chi_f^2 \leq x) + \frac{f}{4m^{**}} (2b - a - 1) \{ \Pr(\chi_f^2 \leq x) \\ & - 2\Pr(\chi_{f+2}^2 \leq x) + \Pr(\chi_{f+4}^2 \leq x) \} + o((m^{**})^{-1}). \end{aligned}$$

From these, we can see that

$$(4.6) \quad E(m^*T^*) = f + o((m^*)^{-1}) \text{ and } E(m^{**}T^{**}) = f + o((m^{**})^{-1}).$$

□

In the above theorem, m^* and m^{**} is chosen such that (4.6) hold. Note that

$$\begin{aligned} \Pr(m^*T^* \leq x) &= \Pr(\chi_f^2 \leq x) + o((m^*)^{-1}) \text{ and} \\ \Pr(m^{**}T^{**} \leq x) &= \Pr(\chi_f^2 \leq x) + o((m^{**})^{-1}) \end{aligned}$$

do not hold. Consequently, chi-squared approximation of the null distributions of m^*T^* and $m^{**}T^{**}$ for small sample cases should not be used. From (4.4) and (4.5), however, we can see that the null distributions of m^*T^* and $m^{**}T^{**}$ do not depend on the nuisance parameters λ_{ji} , $j=2, \dots, q$, $i=1, \dots, k_j$ up to the order n^{-1} . This is meaningful from the point of similar testing for composite hypothesis.

5. A numerical example

First, note that m_1 , m^* and m^{**} defined in THEOREM 2 and THEOREM 3 depend on c defined by (3.3), which is unknown under $H_{k_2, \dots, k_q}$ because λ_{ji} , $j=2, \dots, q$, $i=1, \dots, k_j$ are nuisance parameters. So we estimate c as

$$\hat{c} = f^{-1} \sum_{j=2}^q \sum_{i=1}^{k_j} f_j l_{ji}^{-2}$$

and use $\hat{m}_1$, $\hat{m}^*$ and $\hat{m}^{**}$ instead of m_1 , m^* and m^{**} , respectively, where $\hat{m}_1 = n - (1/2)(a+1) - b + \hat{c}$, $\hat{m}^* = n - a + \hat{c} - 1$ and $\hat{m}^{**} = n - 2b + \hat{c}$.

When $q=2$, we test the null hypothesis $H_i: \text{rank}(\Sigma_{21}) = i$ sequentially for

$i=0, 1, \dots, p_2-1$ to determine the number of useful canonical variates. If we reject H_i for $i=0, 1, \dots, k_2-1$ and accept H_{k_2} , then we can estimate the number of useful canonical variate as k_2 . To generalize this for $q \geq 2$, we put $\tilde{H}(i) = \{H_{i_2, i_3, \dots, i_q}; \sum_{j=2}^q i_j = i\}$ for $i=0, 1, \dots, \sum_{j=2}^q p_j-1$. If all hypotheses in $\tilde{H}(i)$ are rejected for $i=0, 1, \dots, \sum_{j=2}^q k_j-1$ but some hypothesis in $\tilde{H}(\sum_{j=2}^q k_j)$ is accepted, then we adopt $(k_2, \dots, k_q)$ corresponding to the accepted hypothesis as a estimator of $(\text{rank}(\Sigma_{21}), \text{rank}(\Sigma_{3(2)}), \dots, \text{rank}(\Sigma_{q(q-1)}))$. We will refer to this as the "stepwise" method in this paper. Note that the $(k_2, \dots, k_q)$ adopted by this method is not necessarily unique. The total dimension $\sum_{j=2}^q k_j$ is the same for all of the adopted $(k_2, \dots, k_q)$ however, and it can be considered that it is worth interpreting each of them.

We put

$$A(k_2, \dots, k_q) = \sum_{j=2}^q A_j(k_j),$$

where

$$A_j(k_j) = \begin{cases} -N \log \prod_{i=k_j+1}^{p_j} (1-l_{ji}^2) - 2f_j, & \text{if } k_j < p_j, \\ 0 & \text{if } k_j = p_j, \end{cases}$$

and f_j is defined by (3.2). The AIC method is thus equivalent to adopting the model $H_{k_2, \dots, k_q}$ minimizing $A(k_2, \dots, k_q)$. This is a generalization of Theorem 13.6.1 of Siotani, Hayakawa and Fujikoshi [7]. The proof is similar and omitted here.

We will now give a numerical example. For three sets of variates:

Preliminary entrance examination score vector: $\mathbf{X}^{(1)}$

= (Japanese, Sociology, Mathematics-I, Science-I, Foreign Language-I)',

Final entrance examination score vector: $\mathbf{X}^{(2)}$

= (Foreign Language-II, Mathematics-II, Science-II)'

Score vector after entrance: $\mathbf{X}^{(3)}$

= (Score in general educational course, Score in subject course)',

we perform a canonical correlation analysis simultaneously both between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, and between $(\mathbf{X}^{(1)}, \mathbf{X}^{(2)})'$ and $\mathbf{X}^{(3)}$. Suppose we have the following sample correlation matrix $\mathbf{R}$ of $\mathbf{X} = (\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \mathbf{X}^{(3)})'$ computed from the results of 50 students who were into a university department.

$$\mathbf{R} = \left[\begin{array}{ccccc|ccc|cc} 1.0 & & & & & & & & & \\ 0.13 & 1.0 & & & & & & & & \\ 0.02 & -0.10 & 1.0 & & & & & & & \\ 0.06 & 0.02 & 0.06 & 1.0 & & & & & & \\ 0.08 & -0.02 & 0.03 & -0.03 & 1.0 & & & & & \\ \hline 0.10 & 0.10 & 0.01 & -0.07 & 0.37 & 1.0 & & & & \\ -0.14 & -0.08 & 0.11 & 0.09 & -0.06 & -0.04 & 1.0 & & & \\ -0.06 & 0.12 & 0.08 & 0.25 & -0.02 & 0.00 & 0.08 & 1.0 & & \\ \hline 0.03 & 0.14 & 0.10 & 0.05 & 0.08 & 0.18 & 0.16 & 0.18 & 1.0 & \\ -0.29 & -0.02 & 0.32 & -0.30 & 0.16 & 0.31 & 0.24 & 0.04 & 0.82 & 1.0 \end{array} \right]$$

Table 1

$k_2 + k_3$	(k_2, k_3)	$\hat{m}_1 T$	$\hat{m}^* T^*$	$\hat{m}^{**} T^{**}$	f	$\chi^2_f(0.05)$	$\chi^2_f(0.01)$	$A(k_2, k_3)$
0	(0, 0)	67.66	98.18	55.12	31	44.99	52.19	14.92
1	(1, 0)	62.05	94.75	45.44	24	36.42	42.98	20.09
1	(0, 1)	18.39	17.66	18.98	22	33.92	40.29	-23.27
2	(1, 1)	11.20	10.78	11.57	15	25.00	30.58	-18.09
2	(2, 0)	57.83	90.74	40.97	19	30.14	36.19	24.95
2	(0, 2)	13.52	13.05	13.91	15	25.00	30.58	-14.80
3	(1, 2)	6.33	6.15	6.49	8	15.51	20.09	-9.63
3	(2, 1)	6.56	6.33	6.76	10	18.31	23.21	-13.23
3	(3, 0)	53.70	83.35	38.21	16	26.30	32.00	29.72
4	(2, 2)	1.45	1.40	1.49	3	7.81	11.35	-4.77
4	(3, 1)	4.87	4.63	5.08	7	14.07	18.48	-8.46

Then the i th sample canonical correlation l_{2i} between $X^{(1)}$ and $X^{(2)}$ and the i th sample canonical correlation l_{3i} between $(X^{(1)'}, X^{(2)'})'$ and $X^{(3)}$ are calculated as $l_{21}=0.40$, $l_{22}=0.31$, $l_{23}=0.16$, $l_{31}=0.82$ and $l_{32}=0.32$. Table 1 gives the numerical results of testing for H_{k_2, k_3} and criterion $A(k_2, k_3)$. From Table 1, we can see that the stepwise method with a significance level of 0.05, 0.01 based on any statistic of $\hat{m}_1 T$, $\hat{m}^* T^*$ and $\hat{m}^{**} T^{**}$ adopts $(k_2, k_3) = (0, 1)$ as the number pair of useful canonical variates. The AIC method also adopts $(k_2, k_3) = (0, 1)$. This implies that the preliminary entrance examinations and the final entrance examinations are independent and that the linear relationship between the entrance examinations (preliminary and final) and scores after entrance can be interpreted by only the first pair of canonical variates. The first pair is

$$[(-0.53, -0.08, 0.46, -0.54, 0.08, 0.33, 0.17, -0.07)(X^{(1)'}, X^{(2)'})', (-0.95, 1.44)X^{(3)}].$$

Table 1 is also useful in the sense that from the three rows of $(k_2, k_3) = (0, 2)$, $(1, 2)$, $(2, 2)$ in Table 1, we can see that the number of useful canonical variates in the canonical correlation analysis between only $X^{(1)}$ and $X^{(2)}$ is zero. Similarly from the two rows of $(k_2, k_3) = (3, 0)$, $(3, 1)$, we can see that the number of useful canonical variates in the canonical correlation analysis between only $(X^{(1)'}, X^{(2)'})'$ and $X^{(3)}$ is one.

REFERENCES

- [1] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis (2nd ed)*, Wiley, New York.
- [2] Fujikoshi, Y. (1974). The Likelihood Ratio Tests for the Dimensionality of Regression Coefficients, *J. Multiv. Anal.*, **4**, 327-340.
- [3] Fujikoshi, Y. (1977). Asymptotic Expansion of the Distributions of the Latent Roots in MANOVA and the Canonical Correlations, *J. Multiv. Anal.* **7**, 386-396.
- [4] Fujikoshi, Y. (1977). Asymptotic Expansions for the Distributions of Some Multivariate Tests, *In Multivariate Analysis—IV* (P. R. Krishnaiah, Ed.), North-Holland, Amsterdam. 55-71.

- [5] Glynn, W. J., and Muirhead, R. J. (1978). Inference in Canonical Correlation Analysis, *J. Multiv. Anal.*, **8**, 468–478.
- [6] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley, New York.
- [7] Siotani, M., Hayakawa, T., and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis*. Amer. Sci., Columbus, OH.