Hiroshima Math. J.
40 (2010), 215–228

Statistical inference for functional relationship between the specified and the remainder populations

Yasutomo Maeda

(Received September 10, 2009)
(Revised November 21, 2009)

Abstract. This paper is concerned with discovering linear functional relationships among k p-variate populations with mean vectors $\boldsymbol{\mu}_i$, $i = 1, \dots, k$ and a common covariance matrix Σ. We consider a linear functional relationship to be one in which each of the specified r mean vectors, for example, $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_r$ are expressed as linear functions of the remainder mean vectors $\boldsymbol{\mu}_{r+1}, \dots, \boldsymbol{\mu}_k$. This definition differs from the classical linear functional relationship, originally studied by Anderson [1], Fujikoshi [8] and others, in that there are r linear relationships among k mean vectors without any specification of k populations. To derive our linear functional relationship, we first obtain a likelihood test statistic when the covariance matrix Σ is known. Second, the asymptotic distribution of the test statistic is studied in a high-dimensional framework. Its accuracy is examined by simulation.

1. Introduction

There are many works on linear functional relationship among variables. For its review, see Fuller [12], Cheng and Van Ness [7] and so on. Anderson [3] dealt functional relationship in bivariate case, and derived maximum likelihood estimator (MLE) of functional coefficients. Fuller [11] studied some properties of p-variate case. Gleser [13] expanded large sample theory in p-variate case, Anderson [4] dealt the estimation problem when error variables is independent and component correlation exists. Arellano-Valle, Bolfarine and Gasco [5] studied MLE when component covariance matrix is arbitrary.

This paper is concerned with the linear functional relationship between k p-variate populations with mean vectors $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_k$ and a common covariance matrix Σ. There are many patterns of linear functional relationships among k p-variate populations. The most commonly used functional relationship was studied by Anderson [1], Fujikoshi [8] and others state that there are r linear relationships among k mean vectors without any specification of k popula-

2000 *Mathematics Subject Classification*. Primary 12A34, 98B76; Secondary 23C57.

Key words and phrases. Asymptotic distribution; high-dimensional framework; likelihood ratio test statistics (LR test statistics); linear functional relationship; maximum likelihood estimators (MLE).

tions. In this paper, however, we consider a linear functional relationship model with some specification of k populations. More precisely, we consider a linear functional relationship such that the first r mean vectors $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_r$ are expressed as linear combinations of the remainder mean vectors $\boldsymbol{\mu}_{r+1}, \dots, \boldsymbol{\mu}_k$. For our linear functional relationship model, the maximum likelihood estimators (MLE) of the coefficient vectors are shown to be latent vectors of a certain matrix and the likelihood test is related to the smallest latent roots of this matrix.

In this paper we also consider the asymptotic distributions of the smallest latent root and the test statistic in a high-dimensional situation. Recently, there are some results on the asymptotic distributions of latent roots, latent vectors, and test statistics in a high-dimensional framework. Fujikoshi, Himeno and Wakaki [10] derived asymptotic distributions of test statistics for dimensionality in canonical discriminant analysis. Wakaki [22] derived asymptotic for Λ in MANOVA model. For examples of other results in a high-dimensional framework in which both the dimension and sample size are large, see Bai [6], Johnstone [16], Ledoit and Wolf [19] and Raudys and Young [20] etc.

Our paper is organized as follows: Section 2 defines our model, Section 3 derives the MLE of the coefficient vectors and a likelihood ratio (LR) statistic when Σ is known, Section 4 derives the asymptotic distribution of the test statistic under large sample and high-dimensional frameworks, and Section 5 provides simulation result.

2. The linear functional relationship with some specification of k populations

Consider k p-dimensional normal populations $\Pi_i : N_p(\boldsymbol{\mu}_i, \Sigma)$, $i = 1, \dots, k$. Suppose that there are independent samples $\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{in_i}$ from Π_i. We consider a multivariate linear functional relationship model as follows:

$$H_0 : \begin{pmatrix} \boldsymbol{\mu}_1' \\ \vdots \\ \boldsymbol{\mu}_r' \end{pmatrix} = \begin{pmatrix} \boldsymbol{\delta}_1' \\ \vdots \\ \boldsymbol{\delta}_r' \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu}_{r+1}' \\ \vdots \\ \boldsymbol{\mu}_k' \end{pmatrix} \tag{1}$$

where $\boldsymbol{\delta}_i = (\delta_{i,r+1}, \dots, \delta_{i,k})$ is unknown. The following matrix notations are used:

$$\Delta = (\boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_r)' : r \times (k - r),$$

$$M_1 = (\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_r)' : r \times p, \qquad M_2 = (\boldsymbol{\mu}_{r+1}, \dots, \boldsymbol{\mu}_k)' : (k - r) \times p,$$

$$M = (\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_k)' = (M_1', M_2')' : k \times p.$$

Then, our model is expressed in matrix form as

$$H_0 : M_1 = \Delta M_2. \tag{2}$$

We note that the model (1) is different from the one in which there exist r linear relationships. The model (1) refers to the relationship with some specification of k populations. When $k = 2$ and $r = 1$, the hypothesis H_0 becomes to $\boldsymbol{\mu}_1 = \boldsymbol{\delta}_1 \boldsymbol{\mu}_2$, which was considered by Kraft, Olkin and van Eeden [15].

3. Maximum likelihood estimators (MLE) and LR test

In this section, we derive the MLEs of the mean matrices M_1, M_2 and the coefficient matrix Δ under H_0, when Σ is known. Further, we derive a LR test for H_0.

The likelihood function $L(M, \Delta)$ is given by

$$\begin{aligned} L(M, \Delta) &= \prod_{i=1}^{k} \prod_{j=1}^{n_i} (2\pi)^{-p/2} |\Sigma|^{-1/2} \operatorname{etr}\left\{ -\frac{1}{2} \Sigma^{-1} (\boldsymbol{x}_{ij} - \boldsymbol{\mu}_i)(\boldsymbol{x}_{ij} - \boldsymbol{\mu}_i)' \right\} \\ &= (2\pi)^{-np/2} |\Sigma|^{-n/2} \exp\left\{ -\frac{1}{2} \operatorname{tr} \Sigma^{-1} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\boldsymbol{x}_{ij} - \boldsymbol{\mu}_i)(\boldsymbol{x}_{ij} - \boldsymbol{\mu}_i)' \right\}, \end{aligned}$$

where $n = n_1 + \cdots + n_k$. Let

$$\bar{\boldsymbol{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \boldsymbol{x}_{ij} \qquad \text{and} \qquad W = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)'.$$

Since

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (\boldsymbol{x}_{ij} - \boldsymbol{\mu}_i)(\boldsymbol{x}_{ij} - \boldsymbol{\mu}_i)' = W + \sum_{i=1}^{k} n_i (\bar{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)(\bar{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)',$$

we have

$$\begin{aligned} -2 \log L(M, \Delta) &= n \log|\Sigma| + \operatorname{tr} \Sigma^{-1} W \\ &\quad + \operatorname{tr} \Sigma^{-1} \sum_{i=1}^{k} n_i (\bar{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)(\bar{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)' + np \log 2\pi. \end{aligned}$$

The MLEs of M and Δ are obtained by minimizing the above $-2 \log L(M, \Delta)$ or

$$g(M, \Delta) = \operatorname{tr} \Sigma^{-1} \sum_{i=1}^{k} n_i (\bar{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)(\bar{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)'.$$

Under H_0, g is a function of $\varDelta$ and M_2. First, we consider to minimize g with respect to M_2. Let

$$z_i = \sqrt{n_i}\Sigma^{-1/2}\bar{\boldsymbol{x}}_i, \quad \gamma_i = \sqrt{n_i}\Sigma^{-1/2}\boldsymbol{\mu}_i, \qquad i = 1, 2, \dots, k.$$

The functional model (1) can be expressed as

$$H_0 : \Gamma_1 = \Xi\Gamma_2, \tag{3}$$

where

$$\begin{aligned}
&\Gamma_1 = D_1 M_1 \Sigma^{-1/2} : r \times p, \qquad D_1 = \mathrm{diag}(\sqrt{n_1}, \dots, \sqrt{n_r}),\\
&\Gamma_2 = D_2 M_2 \Sigma^{-1/2} : (k - r) \times p, \qquad D_2 = \mathrm{diag}(\sqrt{n_{r+1}}, \dots, \sqrt{n_k}),\\
&\Xi = D_1 \varDelta D_2^{-1} = (\xi_{ij}) : r \times (k - r),\\
&\xi_{ij} = \sqrt{n_i/n_j}\delta_{ij}, \qquad i = 1, \dots, r,\ j = r + 1, \dots, k,\\
&\Gamma = (\Gamma_1, \Gamma_2) = (\gamma_1, \dots, \gamma_k)'.
\end{aligned}$$

Our model may be also formulated in the term of the matrix Γ as

$$\Gamma = (\Gamma_1', \Gamma_2')' = (\Gamma_2'\Xi', \Gamma_2')' = (\Xi', I_{k-r})'\Gamma_2 = \begin{pmatrix} \Xi \\ I_{k-r} \end{pmatrix}\Gamma_2.$$

Put

$$\begin{aligned}
&\bar{X} = (\bar{\boldsymbol{x}}_1, \dots, \bar{\boldsymbol{x}}_k)' : k \times p,\\
&Z = D\bar{X}\Sigma^{-1/2} : k \times p, \qquad D = \mathrm{diag}(\sqrt{n_1}, \dots, \sqrt{n_k}),\\
&A = (\Xi', I_{k-r})' = \begin{pmatrix} \Xi \\ I_{k-r} \end{pmatrix} : k \times (k - r).
\end{aligned}$$

Then, under H_0 we can write $g(M, \varDelta)$ as

$$\sum_{i=1}^{k} \mathrm{tr}(z_i - \gamma_i)(z_i - \gamma_i)' = \mathrm{tr}(Z - A\Gamma_2)'(Z - A\Gamma_2) \equiv g^*(\Gamma_2, \Xi).$$

For any fixed Ξ, we have

$$\mathrm{tr}(Z - A\Gamma_2)'(Z - A\Gamma_2) \geq \mathrm{tr}(Z - A\hat{\Gamma}_2)'(Z - A\hat{\Gamma}_2) = \mathrm{tr}\, Z'(I_k - P_A)Z,$$

where

$$\hat{\Gamma}_2 = (A'A)^{-1}A'Z : (k - r) \times p, \qquad P_A = A(A'A)^{-1}A' : k \times k.$$

Here, P_A is a projection matrix of the space $\mathscr{R}[A]$ spanned by the column vectors of A. The equality holds when $\Gamma = \Gamma_2$.

Next, we consider to minimize tr $Z'(I_k - P_A)Z$ with respect to Ξ. Since $(I_k - P_A)$ is idempotent, the eigenvalues of $(I_k - P_A)$ are 1 or 0. Let the column vectors of $U_1 : k \times r$ be orthonormal eigenvectors of $(I_k - P_A)$ corresponding to 1. Then, considering a spectrum decomposition of $I_k - P_A$, we have

$$I_k - P_A = U_1 U_1', \qquad \text{and} \qquad U_1 U_1' = I_r.$$

Therefore

$$\text{tr } Z'(I_k - P_A)Z = \text{tr } Z' U_1 U_1' Z = \text{tr } U_1' Z Z' U_1.$$

Let

$$\ell_1 \geq \cdots \geq \ell_k > 0 \tag{4}$$

be the eigenvalues of $ZZ' = D\bar{X}\Sigma^{-1}\bar{X}'D$, and let

$$\boldsymbol{h}_1, \ldots, \boldsymbol{h}_k \tag{5}$$

be the corresponding eigenvectors satisfying $\boldsymbol{h}_i'\boldsymbol{h}_i = 1$ and $\boldsymbol{h}_i'\boldsymbol{h}_j = 0$ $(i \neq j)$. We denote the latent vectors in matrix forms as

$$H = (H_1, H_2), \qquad H_1 = (\boldsymbol{h}_1, \ldots, \boldsymbol{h}_{k-r}), \qquad H_2 = (\boldsymbol{h}_{k-r+1}, \ldots, \boldsymbol{h}_k), \tag{6}$$

and further we decompose H_2 as

$$H_2 = \begin{pmatrix} H_{12} \\ H_{22} \end{pmatrix}, \qquad H_{12} : r \times r. \tag{7}$$

Then, note that

$$\min_{L_1'L_1 = I_r} \text{tr } L_1' Z Z' L_1 = \text{tr } H_2' Z Z' H_2 = \sum_{i=1}^{r} \ell_{k-r+i},$$

(see, e.g., Seber [21]). Thus, we have

$$\min_H \{-2 \log L\} = np \log 2\pi + n \log|\Sigma| + \text{tr } \Sigma^{-1} W + \sum_{i=1}^{r} \ell_{k-r+i}.$$

In order to complete the above result, we need to show that there exists an $\hat{A} = (\hat{\Xi}', I_k)'$ such that $I_k - P_A = U_1 U_1'$ or $(I_k - P_{\hat{A}})H_2 = H_2$. Next, we seek $\hat{\Delta}$, and $\hat{M}_2$. Note that H_2 satisfies $(I_k - P_{\hat{A}})H_2 = H_2$. The equation $(I_k - P_{\hat{A}})H_2 = H_2$ is expressed as $H_2'\hat{A} = O$. Since $\hat{\Xi} = D_1 \hat{\Delta} D_2^{-1}$ and

$$H_2'\hat{A} = (H_{12}' \quad H_{22}') \begin{pmatrix} \hat{\Xi} \\ I_{k-r} \end{pmatrix} = H_{12}'\hat{\Xi} + H_{22}' = O.$$

Further, it holds with probability 1 that

$$A0 : H_{12} \text{ is nonsingular.} \tag{8}$$

This result has been shown by Gleser [13]. By using this property, we have

$$\hat{\Delta} = -D_1^{-1}(H_{22}H_{12}^{-1})'D_2.$$

Similarly,

$$\hat{M}_2 = D_2^{-1}\hat{\Gamma}_2\Sigma^{1/2} = D_2^{-1}(\hat{A}'\hat{A})^{-1}\hat{A}'Z\Sigma^{1/2} = D_2^{-1}(\hat{A}'\hat{A})^{-1}\hat{A}'D\overline{X}.$$

We can thus derive the following theorem:

THEOREM 1. *When Σ is known, the MLE of M_2 and Δ under the model hypothesis $H_0 : M_1 = \Delta M_2$ are given as follows:*

$$\hat{\Delta} = -D_1^{-1}(H_{22}H_{12}^{-1})'D_2, \qquad H_2 = (H_{12}', H_{22}')',$$

$$\hat{M}_2 = D_2^{-1}(\hat{A}'\hat{A})^{-1}\hat{A}'D\overline{X}, \qquad \hat{A} = (\hat{\Xi}', I_{k-r})', \qquad \hat{\Xi} = -(H_{22}H_{12}^{-1})',$$

where H_{12}, H_{22} are the submatrices of H_2 (see (7)) partitioned with the smallest r eigenvalues of $D\overline{X}\Sigma^{-1}\overline{X}'D$, that is, $H_2 = (H_{12}', H_{22}')'$.

On the other hand, we note that the maximum likelihood for the no restriction model is

$$\min\{-2 \log L\} = np \log 2\pi + n \log|\Sigma| + \mathrm{tr}\, \Sigma^{-1} W.$$

The LR statistic is given by the following theorem:

THEOREM 2. *A likelihood ratio (LR) statistic is*

$$LR = \sum_{i=1}^{r} \ell_{k-r+i},$$

where $\ell_1 \geq \cdots \geq \ell_k$ are the eigenvalues of $ZZ' = D\overline{X}\Sigma^{-1}\overline{X}'D$.

The LR statistic is the summation of the smallest eigenvalues ℓ_{k-r+1}, $\ell_{k-r+2}, \ldots, \ell_k$ and does not depend on a set of specified r populations. Thus we show that the LR statistic gives the same result for testing the dimensionality model.

4. Asymptotic distribution of LR test under high-dimensional framework

In this section, we derive the limiting distribution of LR test statistic under a high-dimensional framework

$$\mathrm{A1} : n \to \infty, \quad k : \text{fix}, \quad p \to \infty, \quad n - p \to \infty, \quad c = p/n \to c_0 \in (0, 1).$$

We use $\bar{\boldsymbol{x}}_i \sim N_p\big(\boldsymbol{\mu}_i, \frac{1}{n_i}\Sigma\big)$, and

$$\overline{X} = (\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_k)' \sim N_{k,p}(M, D^{-2} \otimes \Sigma),$$

where $D = \mathrm{diag}(\sqrt{n_1}, \ldots, \sqrt{n_k})$. Therefore,

$$Z = D\overline{X}\Sigma^{-1/2} \sim N_{k,p}(DM\Sigma^{-1/2}, I_k \otimes I_p),$$

and it is easy (see, e.g., Gupta and Nagar [14]) to see that

$$ZZ' = D\overline{X}\Sigma^{-1}\overline{X}'D \sim W_k(p, I_k; DM\Sigma^{-1}M'D).$$

Under the linear relationship model $H_0 : M_1 = \Delta M_2$, $M = (\Delta', I_{k-r})'M_2$, and $\mathrm{rank}(M) = k - r$. Therefore,

$$\mathrm{rank}(DM\Sigma^{-1}M'D) = k - r.$$

Moreover, since the LR statistic is a function of the eigenvalues of ZZ', we may use $HZZ'H$ in stead of ZZ'. Now we use an orthogonal matrix $H : k \times k$ such that

$$HZZ'H' \sim W_k(p, I_k; n\Omega_0), \qquad \Omega_0 = \mathrm{diag}(\omega_1, \ldots, \omega_{k-r}, 0, \ldots, 0) : k \times k,$$

where $\omega_1 \geq \cdots \geq \omega_{k-r}$ are the nonzero eigenvalues of $DM\Sigma^{-1}M'D$. In the derivation of our asymptotic distribution, we assume

$$\begin{aligned} \text{A2:} \quad & \omega_1 > \cdots > \omega_{k-r} > \omega_{k-r+1} = \cdots = \omega_k = 0, \\ & \omega_i = O(1), \qquad i = 1, \ldots, k - r. \end{aligned}$$

Therefore, we may start from the following set-up: $\ell_1 \geq \cdots \geq \ell_k$ are the eigenvalues of $HZZ'H'$, LR statistic is

$$LR = \sum_{i=1}^{r} \ell_{k-r+i},$$

and

$$HZZ'H' \sim W_k(p, I_k; n\Omega_0), \qquad \Omega_0 = \mathrm{diag}(\omega_1, \ldots, \omega_{k-r}, 0, \ldots, 0) : k \times k.$$

Under the assumption A1, we will study the approximation for

$$W \sim W_k(p, I_k; n\Omega), \qquad \Omega = \mathrm{diag}(\omega_1, \ldots, \omega_k) : k \times k \qquad \textit{and} \qquad \Omega = \Omega_0.$$

Put

$$U = \frac{1}{\sqrt{p}}\{W - (pI_k + n\Omega_0)\}.$$

We thereby use the following lemma (see, e.g., Gupta and Nagar [14]):

Lemma 1. *Suppose that $V \sim W_p(n, \Sigma; \Omega)$, then the characteristic function of V is*

$$C_V(T) = |I_p - 2iT\Sigma|^{-n/2} \operatorname{etr}\left\{-\frac{1}{2}\Omega + \frac{1}{2}\Omega(I_p - 2iT\Sigma)^{-1}\right\},$$

where T is real symmetric matrix with the (i,j) element given by $\frac{1}{2}(1+\delta_{ij})t_{ij}$, and δ_{ij} is Kronecker delta.

Since

$$\begin{aligned}-\frac{1}{2}\Omega + \frac{1}{2}\Omega(I_p - 2iT\Sigma)^{-1} &= \frac{1}{2}\Omega\{-I_p + (I_p - 2iT\Sigma)^{-1}\} \\ &= \frac{1}{2}\Omega\{-(I_p - 2iT\Sigma) + I_p\}(I_p - 2iT\Sigma)^{-1} \\ &= i\Omega T\Sigma(I_p - 2iT\Sigma)^{-1},\end{aligned}$$

we can write the characteristic function $C_V(T)$ as

$$C_V(T) = |I_p - 2iT\Sigma|^{-n/2} \operatorname{etr}\{i\Omega T\Sigma(I_p - 2iT\Sigma)^{-1}\}.$$

Using Lemma 1, the characteristic function U is

$$\begin{aligned}C_U(T) &= \mathrm{E}[\exp(i \operatorname{tr} TU)] \\ &= \exp\left(-\frac{i}{\sqrt{p}} \operatorname{tr}(pI_k + n\Omega_0)\right)\mathrm{E}\left[\exp\left(\frac{i}{\sqrt{p}} \operatorname{tr} TW\right)\right] \\ &= \exp\left(-\frac{i}{\sqrt{p}} \operatorname{tr} T(pI_k + n\Omega_0)\right) \\ &\quad \times \left|I_k - \frac{2i}{\sqrt{p}}T\right|^{-p/2} \operatorname{etr}\left\{\frac{i}{\sqrt{p}}n\Omega_0 T\left(I_k - \frac{2i}{\sqrt{p}}T\right)^{-1}\right\}.\end{aligned}$$

Since

$$\begin{aligned}\log\left|I_k - \frac{2i}{\sqrt{p}}T\right|^{-p/2} &= -\frac{p}{2}\left|I_k - \frac{2i}{\sqrt{p}}T\right| \\ &= \sqrt{p}i \operatorname{tr} T - \operatorname{tr} T^2 + \frac{4i^3}{3\sqrt{p}} \operatorname{tr} T^3 + O(p^{-1}),\end{aligned}$$

we have

$$\begin{aligned}\left|I_k - \frac{2i}{\sqrt{p}}T\right|^{-p/2} &= \exp\left(\sqrt{p}i \operatorname{tr} T - \operatorname{tr} T^2 + \frac{4i^3}{3\sqrt{p}} \operatorname{tr} T^3 + O(p^{-1})\right) \\ &= \exp(\sqrt{p}i \operatorname{tr} T) \times \exp(-T^2) + O(p^{-1/2}).\end{aligned}$$

Moreover, using

$$\left(I_k - \frac{2i}{\sqrt{p}}T\right)^{-1} = I_k + \frac{2i}{\sqrt{p}}T + \frac{4i^2}{p}T^2 + O(p^{-3/2}),$$

we get

$$\begin{aligned}&\mathrm{etr}\left\{\frac{i}{\sqrt{p}}n\Omega_0 T\left(I_k - \frac{2i}{\sqrt{p}}T\right)^{-1}\right\}\\ &\quad = \mathrm{etr}\left(\frac{i}{\sqrt{p}}n\Omega_0 T\right)\,\mathrm{etr}\left(-\frac{2}{p}n\Omega_0 T^2\right) + O(p^{-1/2}).\end{aligned}$$

Therefore,

$$\begin{aligned}C_U(T) &= \exp\left(-\frac{i}{\sqrt{p}}\ \mathrm{tr}\ T(pI_k + n\Omega_0)\right)\\ &\qquad \times \left|I_k - \frac{2i}{\sqrt{p}}T\right|^{-p/2}\mathrm{etr}\left\{\frac{i}{\sqrt{p}}n\Omega_0 T\left(I_k - \frac{2i}{\sqrt{p}}T\right)^{-1}\right\}.\\ &= \mathrm{etr}(-(I_k + 2c\Omega_0)T^2) + O(p^{-1/2}).\end{aligned}$$

It is easy to see that

$$\begin{aligned}\mathrm{tr}(I_k + 2c\Omega)T^2 = &\sum_{1\le i\le k-r}(1+2c\omega_i)t_{ii}^2 + \sum_{k-r+1\le i\le k} t_{ii}^2\\ &+ \sum_{1\le i<j\le k-r}\frac{(1+2c\omega_i)}{2}t_{ij}^2\\ &+ \sum_{1\le i\le k-r,\, k-r+1\le j\le k}\frac{(1+c\omega_i)}{2}t_{ij}^2 + \sum_{k-r+1\le i<j\le k}\frac{1}{2}t_{ij}^2.\end{aligned}$$

We obtain the limiting distribution of u_{ij} is

$$\begin{aligned}&u_{ii} \sim N(0, 2(1+2c\omega_i)), \qquad 1 < i < k-r,\\ &u_{ii} \sim N(0,2), \qquad k-r+1 < i < k,\\ &u_{ij} \sim N(0, 1+2c\omega_i), \qquad i \ne j,\ 1 \le i < j \le k-r,\\ &u_{ij} \sim N(0, 1+c\omega_i), \qquad i \ne j,\ 1 \le i \le k-r,\ k-r+1 \le j \le k,\\ &u_{ij} \sim N(0,1), \qquad i \ne j,\ k-r+1 \le i < j \le k,\end{aligned}$$

whereby the u_{ij}'s are independent. Suppose that $W = HZZ'H'$, we get

$$R = \frac{1}{p}W = I_k + c\Omega_0 + \frac{1}{\sqrt{p}}U.$$

Under assumption A2 we can see by using Lawley [17] [18], Anderson [2] and Fujikoshi [9], the smallest $k - r$ eigenvalues of R are approximated by those of

$$Q = I_r + \frac{1}{\sqrt{p}}U_{22} + O(p^{-1}),$$

where U_{22} is a submatrix of

$$U = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}, \qquad U_{12} : (k - r) \times r.$$

Now we consider a standardized statistics

$$T_{LR} = \sqrt{p}\left(\frac{1}{p}\sum_{i=1}^{r} \ell_{k-r+i} - r\right),$$

which is written as

$$T_{LR} = \sqrt{p}\left(\frac{1}{p}\sum_{i=1}^{r} \ell_{k-r+i} - r\right) = \mathrm{tr}\ U_{22}.$$

Since $u_{ii} \xrightarrow{d} N(0, 2)$, $k - r + 1 \leq i \leq k$, we have

$$T_{LR} \xrightarrow{d} N(0, 2r).$$

THEOREM 3. *Let*

$$T_{LR} = \sqrt{p}\left(\frac{1}{p}\sum_{i=1}^{r} \ell_{k-r+i} - r\right).$$

Then, under H_0 and the assumptions A1 *and* A2,

$$\frac{T_{LR}}{\sigma} \xrightarrow{d} N(0, 1), \qquad \sigma = \sqrt{2r},$$

where $\xrightarrow{d}$ means the convergence in distribution.

5. Simulation result

We simulate an experiment to examine the accuracy of our normal approximation of T_{LR}. For the cases of (a) $k = 2$ and $r = 1$, (b), (c) $k = 3$

and $r = 1, 2$ and (d)$\sim$(f) $k = 4$ and $r = 1\sim 3$, we obtain the accuracy of the normal approximation for the nominal probability 0.05 (Table 1$\sim$Table 6) for $p, n_j = 10$ and $p, n_j = 200$. The simulation with 100,000 repeatitions was performed.

Table 1. (a) The accuracy of the normal approximation for the nominal probability 0.05 for $k = 2$ and $r = 1$

	p					
n_j	10	20	50	100	150	200
10	0.03904	0.04346	0.04714	0.04984	0.04839	0.04866
20	0.04251	0.04616	0.04849	0.04964	0.04948	0.05083
50	0.04306	0.04696	0.04776	0.04935	0.04953	0.05058
100	0.04242	0.04680	0.04840	0.04978	0.04884	0.04918
150	0.04238	0.04621	0.04768	0.04931	0.04921	0.04998
200	0.04273	0.04553	0.04956	0.04870	0.05074	0.05163

Table 2. (b) The accuracy of the normal approximation for the nominal probability 0.05 for $k = 3$ and $r = 1$

	p					
n_j	10	20	50	100	150	200
10	0.00224	0.01570	0.02199	0.03392	0.03802	0.03956
20	0.00839	0.02515	0.03548	0.03800	0.04038	0.04314
50	0.02363	0.03049	0.03838	0.04203	0.04251	0.04416
100	0.02088	0.03341	0.03922	0.04138	0.04432	0.04546
150	0.02472	0.03267	0.03864	0.04308	0.04451	0.04403
200	0.02602	0.03310	0.03983	0.04303	0.04358	0.04558

Table 3. (c) The accuracy of the normal approximation for the nominal probability 0.05 for $k = 3$ and $r = 2$

	p					
n_j	10	20	50	100	150	200
10	0.03215	0.03787	0.04171	0.04489	0.04574	0.04664
20	0.03313	0.03750	0.04334	0.04441	0.04609	0.04522
50	0.03328	0.03730	0.04323	0.04456	0.04600	0.04616
100	0.03307	0.03892	0.04290	0.04523	0.04703	0.04587
150	0.03251	0.03924	0.04259	0.04488	0.04679	0.04718
200	0.03346	0.03831	0.04302	0.04518	0.04752	0.04637

Table 4. (d) The accuracy of the normal approximation for the nominal probability 0.05 for $k = 4$ and $r = 1$

	p					
n_j	10	20	50	100	150	200
10	0.00004	0.00329	0.01328	0.02138	0.02626	0.02880
20	0.00050	0.01016	0.02236	0.03057	0.03311	0.03478
50	0.00883	0.01938	0.02866	0.03424	0.03670	0.03829
100	0.01006	0.02141	0.02895	0.03598	0.03818	0.03983
150	0.01282	0.02240	0.03118	0.03673	0.03795	0.03974
200	0.01121	0.02175	0.03271	0.03547	0.03864	0.04018

Table 5. (e) The accuracy of the normal approximation for the nominal probability 0.05 for $k = 4$ and $r = 2$

	p					
n_j	10	20	50	100	150	200
10	0.00128	0.00670	0.01880	0.02711	0.02942	0.03116
20	0.00469	0.01553	0.02406	0.03193	0.03406	0.03512
50	0.01230	0.01937	0.03019	0.03561	0.03625	0.03814
100	0.00956	0.02334	0.02872	0.03587	0.03754	0.03999
150	0.01507	0.01965	0.03130	0.03625	0.03807	0.03856
200	0.01451	0.02353	0.03153	0.03691	0.03855	0.03963

Table 6. (f) The accuracy of the normal approximation for the nominal probability 0.05 for $k = 4$ and $r = 3$

	p					
n_j	10	20	50	100	150	200
10	0.02643	0.03067	0.03771	0.04291	0.04322	0.04404
20	0.02701	0.03293	0.03897	0.04223	0.04340	0.04506
50	0.02716	0.03336	0.04049	0.04154	0.04420	0.04530
100	0.02743	0.03405	0.03950	0.04184	0.04371	0.04390
150	0.02728	0.03334	0.03951	0.04271	0.04326	0.04486
200	0.02761	0.03357	0.03898	0.04294	0.04425	0.04507

Clearly the case (a) is good approximate than the other cases. As the number k of groups are increased, the approximations become bad. However, for all the cases, we can find to be near 0.05 as n and p are larger if not satisfied $n \geq p$ under A1 and A2.

Acknowledgement

The author wishes to express his gratitude to the two reviewers for their valuable comments and to Prof. T. Sugiyama and Y. Fujikoshi, Chuo University for many useful comments and suggestions.

References

[1] Anderson, T. W., Estimating linear restrictions on regression coefficients for multivariate normal distributions, Ann. Math. Statist., **22** (1951), 327–351.

[2] Anderson, T. W., Asymptotic theory for principal component analysis, Ann. Math. Statist., **34** (1963), 122–148.

[3] Anderson, T. W., Estimation of linear functional Relationships: Approximate distributions and communications with simultaneous equations in econometrics, J. Roy. Statist. Soc. Ser. B, **38** (1976), 1–20.

[4] Anderson, T. W., Estmating linear statistical relationships, Ann. Statist., **12** (1984), 1–45.

[5] Arellano-Valle, R. B., Bolfarine, H. and Gasco, L., Measurement error models with nonconstant covariance matrices, J. Multivariate Anal., **82** (2002), 395–415.

[6] Bai, Z. D., Methodologies in spectral analysis of large dimensional random matrices, a review, Statist. Sinica, **9** (1999), 611–677.

[7] Chi-Lun Cheng and John W. Van Ness, Statistical Regression with Measurement Error, Kendall's Library of Statistics 6, 1999.

[8] Fujikoshi, Y., The likelihood ratio tests for the dimensionality of regression coefficients, J. Multivariate Anal., **4** (1974), 327–340.

[9] Fujikoshi, Y., Asymptotic expansions for the distributions of some multivariate tests, Multivariate Analysis-IV, (P. R. Krishnaiah, Ed.), North-Holland Publishing Company, Amsterdam., (1977), 55–71.

[10] Fujikoshi, Y., Himeno, T. and Wakaki, H., Asymptotic results in canonical discriminant analysis when the dimension is large compared to the sample size, J. Statist. Plann. Inference, **138** (2008), 3457–3466.

[11] Fuller, Wayne A., Properties of some estimators for the errors-in-variables model, Ann. Statist., **8** (1980), 407–422.

[12] Fuller, Wayne A., Measurement Error Models, Wiley, New York, 1987.

[13] Gleser, Leon Jay, Estimation in a multivariate 'errors in variables' regression model: large sample results, Ann. Statist., **9** (1981), 24–44.

[14] Gupta, A. K. and Nagar, D. K., Matrix Variate Distributions, Chapman & Hall, 2000.

[15] Kraft, C. H., Olkin, I. and van Eeden, C., Estimation and testing for differences in magnitude or displacement in the mean vectors of two multivariate normal populations, Ann. Math. Statist., **43** (1972), 455–467.

[16] Johnstone, I. M., On the distribution of the largest eigenvalue in principal component analysis, Ann. Statist., **29** (2001), 295–327.

[17] Lawley, D. N., Tests of significance for the latent roots of covariance and correlation matrices, Biometrika, **43** (1956), 128–136.

[18] Lawley, D. N., Tests of significance in canonical analysis, Biometrika, **46** (1959), 59–66.

[19] Ledoit, O. Wolf, M., Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size, Ann. Statist., **30** (2002), 1081–1102.

[20] Raudys, S. and Young, D. M., Results in statistical discriminant analysis: A review of the former Soviet Union literature, J. Multivariate Anal., **89** (2004), 1–35.
[21] Seber, George A. F., A matrix handbook for statisticians, John Wiley & Sons, 2008.
[22] Wakaki, H., Edgeworth expansion of Wilks' lambda statistic, J. Multivariate Anal., **97** (2006), 1958–1964.

Yasutomo Maeda
Depertment of Mathematics
Chuo University
1-13-27 *Kasuga, Bukyo-ku, Tokyo, Japan*
E-mail: maeda@gug.math.chuo-u.ac.jp