Asymptotic expansions of test statistics for dimensionality and additional information in canonical correlation analysis when the dimension is large

Tetsuro Sakurai*

Faculty of Science and Engineering, Chuo University, Kasuga, Bunkyo-ku, 112-8551, Japan

Abstract

This paper examines asymptotic expansions of test statistics for dimensionality and additional information in canonical correlation analysis based on a sample of size $N = n+1$ on two sets of variables, i.e., $\boldsymbol{x}_u; p_1 \times 1$ and $\boldsymbol{x}_v; p_2 \times 1$. These problems are related to dimension reduction. The asymptotic approximations of the statistics have been studied extensively when dimensions p_1 and p_2 are fixed and the sample size N tends to infinity. However, the approximations worsen as p_1 and p_2 increase. This paper derives asymptotic expansions of the test statistics when both the sample size and dimension are large, assuming that $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$ have a joint $(p_1 + p_2)$-variate normal distribution. Numerical simulations revealed that this approximation is more accurate than the classical approximation as the dimension increases.

Key Words and Phrases: Asymptotic expansion, Tests for dimensionality, Additional information, High-dimensional framework.

* *Email address* : s15002@gug.math.chuo-u.ac.jp

1 Introduction

Let $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$ be two random vectors of p_1 and p_2 components with a joint (p_1+p_2)-variate normal distribution with a mean vector $\boldsymbol{\mu} = (\boldsymbol{\mu}_u', \boldsymbol{\mu}_v')'$ and a covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{pmatrix},$$

where Σ_{uv} is a $p_1 \times p_2$ matrix. Without loss of generality we may assume $p_1 \leq p_2$. Let $\rho_1 \geq \cdots \geq \rho_{p_1} \geq 0$ be the possible nonzero population canonical correlations between $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$. Note that $\rho_1^2 \geq \cdots \geq \rho_{p_1}^2 \geq 0$ are the characteristic roots of $\Sigma_{uu}^{-1}\Sigma_{uv}\Sigma_{vv}^{-1}\Sigma_{vu}$. The coefficient vectors $\boldsymbol{\alpha}_{ui}$ and $\boldsymbol{\alpha}_{vi}$ of the canonical variables are defined as the solutions of

$$\begin{aligned} \Sigma_{uv}\Sigma_{vv}^{-1}\Sigma_{vu}\boldsymbol{\alpha}_{ui} &= \rho_i^2\Sigma_{uu}\boldsymbol{\alpha}_{ui}, \qquad & \boldsymbol{\alpha}_{ui}'\Sigma_{uu}\boldsymbol{\alpha}_{uj} &= \delta_{ij}, \\ \Sigma_{vu}\Sigma_{uu}^{-1}\Sigma_{uv}\boldsymbol{\alpha}_{vi} &= \rho_i^2\Sigma_{vv}\boldsymbol{\alpha}_{vi}, \qquad & \boldsymbol{\alpha}_{vi}'\Sigma_{vv}\boldsymbol{\alpha}_{vj} &= \delta_{ij}, \end{aligned}$$

where $\delta_{ij} = 1$ for $i = j$, 0 for $i \neq j$. Let k be the number of nonzero canonical correlations ρ_i. Then $k = \text{rank}(\Sigma_{uv}) \leq p_1$, and the relationships between $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$ can be summarized in terms of the first k canonical variates $(\boldsymbol{\alpha}_{ui}'\boldsymbol{x}_u, \boldsymbol{\alpha}_{vi}'\boldsymbol{x}_v)$, $i = 1, \ldots, k$.

In canonical correlation analysis, the number of nonzero canonical correlations, defines the dimensionality. Consider the problem of testing the hypothesis that the smaller $p_1 - k$ canonical correlations are zero, i.e.,

$$H_0 : \rho_k > \rho_{k+1} = \cdots = \rho_{p_1} = 0. \tag{1.1}$$

This problem is related to reducing the dimension of the canonical variables. Let S be the sample covariance matrix formed from a sample of size $N = n+1$ of $\boldsymbol{x} = (\boldsymbol{x}_u', \boldsymbol{x}_v')'$. Corresponding to a partition of $\boldsymbol{x}$, we partition S as

$$S = \begin{pmatrix} S_{uu} & S_{uv} \\ S_{vu} & S_{vv} \end{pmatrix}.$$

The following test statistics have been considered (e.g., see Sitotani, Hayakawa and Fujikoshi (1985)):

$$LR = -\log \prod_{j=k+1}^{p_1} (1 - r_j^2), \quad LH = \sum_{j=k+1}^{p_1} \frac{r_j^2}{1 - r_j^2}, \quad BNP = \sum_{j=k+1}^{p_1} r_j^2, \tag{1.2}$$

where r_j^2 is the sample canonical correlation. Note that $r_1^2 > \cdots > r_{p_1}^2 > 0$ are the characteristic roots of $S_{uu}^{-1}S_{uv}S_{vv}^{-1}S_{vu}$. Under a large sample framework,

$$\mathrm{A0} : p_1 \text{ and } p_2 \text{ are fixed}, \quad n \to \infty, \tag{1.3}$$

some asymptotic results have been obtained (e.g., see Anderson (2003), Siotani, et al. (1985)). Note that these results will not work well as dimension p_1 or p_2 increases. In order to overcome this weakness, we study the asymptotic distributions of these statistics under a high-dimensional framework such that

$$\begin{array}{ll} \mathrm{A1} & p_1; \text{fixed}, \quad p_2 \to \infty, \quad n \to \infty, \quad m = n - p_2 \to \infty, \\ & p_2/n \to c \in (0,1). \end{array} \tag{1.4}$$

In this paper we also consider asymptotic distributions of test statistics for a hypothesis concerning the sufficiency of the redundancy of a subset of variables from each of $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$. This problem is related to reducing the dimension of the original variables. In order to formulate the hypothesis, we partition $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$ as $\boldsymbol{x}_u = (\boldsymbol{x}_1', \boldsymbol{x}_2')'$, $\boldsymbol{x}_1 : q_1 \times 1$, $\boldsymbol{x}_2 : q_2 \times 1$, $\boldsymbol{x}_v = (\boldsymbol{x}_3', \boldsymbol{x}_4')'$, $\boldsymbol{x}_3 : q_3 \times 1$, $\boldsymbol{x}_4 : q_4 \times 1$ and $\boldsymbol{\alpha}_{ui}, \boldsymbol{\alpha}_{vi}, \boldsymbol{\mu}_u, \boldsymbol{\mu}_v, \Sigma$ comfortably:

$$\begin{pmatrix} \boldsymbol{\alpha}_{ui} \\ \boldsymbol{\alpha}_{vi} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_{1i} \\ \boldsymbol{\alpha}_{2i} \\ \boldsymbol{\alpha}_{3i} \\ \boldsymbol{\alpha}_{4i} \end{pmatrix}, \quad \begin{pmatrix} \boldsymbol{\mu}_u \\ \boldsymbol{\mu}_v \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_3 \\ \boldsymbol{\mu}_4 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \Sigma_{34} \\ \Sigma_{41} & \Sigma_{42} & \Sigma_{43} & \Sigma_{44} \end{pmatrix}.$$

Note that $p_1 = q_1 + q_2$ and $p_2 = q_3 + q_4$. Then, the hypothesis of the sufficiency of $\boldsymbol{x}_1$ and $\boldsymbol{x}_3$, as the redundancy of $\boldsymbol{x}_2$ and $\boldsymbol{x}_4$, is formulated as follows:

$$H_1 : \boldsymbol{\alpha}_{2i} = \boldsymbol{0}, \quad \boldsymbol{\alpha}_{4i} = \boldsymbol{0} \qquad (i = 1, \ldots, k).$$

Let S be the sample covariance matrix formed from a sample of size $N = n+1$ of $(\boldsymbol{x}_u', \boldsymbol{x}_v')'$. Corresponding to a partition of Σ, we partition S as

$$S = \begin{pmatrix} S_{uu} & S_{uv} \\ S_{vu} & S_{vv} \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{pmatrix}.$$

To test H_1, we consider the statistic (Fujikoshi (1982)) defined by

$$T = \begin{vmatrix} S_{22\cdot 13} & S_{24\cdot 13} \\ S_{42\cdot 13} & S_{44\cdot 13} \end{vmatrix} / \{|S_{22\cdot 1}||S_{44\cdot 3}|\}. \tag{1.5}$$

This is a likelihood ratio statistic. Here, $S_{22\cdot 1} = S_{22} - S_{21}S_{11}^{-1}S_{12}$, $S_{22\cdot 13} = S_{22} - S_{2(13)}S_{(13)(13)}^{-1}S_{(13)2}$, $S_{2(13)} = (S_{21}, S_{23})$, etc.

Under a large sample framework A0, an asymptotic expansion is obtained (see Fujikoshi (1982)). However, the result will not work well as dimensions p_1 and p_2 increase. In order to overcome this weakness, we study asymptotic expansions of the statistics under a high-dimensional framework such that

$$\begin{aligned} \text{A2} \qquad & p = p_1 + p_2 \to \infty, \quad n \to \infty, \\ & m = n - p \to \infty, \quad p/n \to c \in (0, 1). \end{aligned} \tag{1.6}$$

Numerical simulations revealed that our approximation becomes more accurate than the classical approximation as the dimension increases. Similar approximations have been proposed in the MANOVA model and discriminant analysis. Fujikoshi, Himeno, and Wakaki (2006) derived asymptotic distributions of test statistics for dimensionality under A1. Tonda and Fujikoshi (2004) derived an asymptotic expansion of the distribution of Wilks' lambda statistic Λ under a high-dimensional framework. Wakaki (2006) derived similar results for Λ under a different high-dimensional framework. For examples of other distributional results in a high-dimensional framework in which both the dimension and sample size are large, see Bai (1999), Johnstone (2001), Ledoit and Wolf (2002), and Raudys and Young (2004).

2 Distributions of tests for dimensionality

In this section we consider the distribution of the three test statistics (1.2) under framework A1. When we consider the distributions of the statistics in (1.2),

$$\Sigma = \begin{pmatrix} I_{p_1} & \tilde{\mathcal{P}}' \\ \tilde{\mathcal{P}} & I_{p_2} \end{pmatrix}, \quad \tilde{\mathcal{P}} = (\mathcal{P}, O), \quad \mathcal{P} = \mathrm{diag}(\rho_1, \ldots, \rho_{p_1}),$$

since the statistics are expressed as functions of the characteristic roots of $S_{uu}^{-1}S_{uv}S_{vv}^{-1}S_{vu}$ without loss of generality we may assume. Let $A = nS$. Correponding to a partition of $\boldsymbol{x}$, we partition A as

$$A = \begin{pmatrix} A_{uu} & A_{uv} \\ A_{vu} & A_{vv} \end{pmatrix}.$$

For our derivation, we use the following properties (see Sugiura and Fujikoshi (1969)):

(a) $A_{uu\cdot v} \sim W_{p_1}(m, \Delta)$, where $\Delta = I_{p_1} - \mathcal{P}^2$ and $m = n - p_2$.

(b) Let W be the first $p_1 \times p_1$ submatrix of A_{vv}. Then, given W, $A_{uv}A_{vv}^{-1}A_{vu} \sim W_{p_1}(p_2, \Delta; \mathcal{P}W\mathcal{P})$, and $A_{uv}A_{vv}^{-1}A_{vu}$ and $A_{uu\cdot v}$ are independent.

(c) $W \sim W_{p_1}(n, I_{p_1})$, W and $A_{uu\cdot v}$ are independent.

Let

$$\ell_i^2 = \frac{r_i^2}{1 - r_i^2}, \quad i = 1, \ldots, p_1.$$

These are the characteristic roots of $A_{uu\cdot v}^{-1}A_{uv}A_{vv}^{-1}A_{vu}$. We will derive an asymptotic distribution of a function of $\ell_1^2 > \ldots > \ell_{p_1}^2$ that leads to the asymptotic distribution of a function of $r_1^2 > \ldots > r_{p_1}^2$, since $r_i^2 = \ell_i^2/(1+\ell_i^2)$. Note that without loss of generality, we may assume

(1) $A_{uu\cdot v} \sim W_{p_1}(m, I_{p_1})$.

(2) Let W be the first $p_1 \times p_1$ submatrix of A_{vv}. Then, given W, $A_{uv}A_{vv}^{-1}A_{vu} \sim W_{p_1}(p_2, I_{p_2}; \Gamma W \Gamma)$, where $\Gamma = \Delta^{-\frac{1}{2}}\mathcal{P}$, and $A_{uv}A_{vv}^{-1}A_{vu}$ and $A_{uu\cdot v}$ are independent.

(3) $W \sim W_{p_1}(n, I_{p_1})$, W and $A_{uu\cdot v}$ are independent.

Now we consider the perturbation expansion of

$$Q = A_{uu\cdot v}^{-\frac{1}{2}}A_{uv}A_{vv}^{-1}A_{vu}A_{uu\cdot v}^{-\frac{1}{2}}.$$

Let U and V be the matrices defined by

$$U = \frac{1}{\sqrt{p_2}}\left\{A_{uv}A_{vv}^{-1}A_{vu} - (p_2 I_{p_1} + n\Gamma^2)\right\} \text{ and } V = \frac{1}{\sqrt{m}}(A_{uu\cdot v} - mI_{p_1}),$$

respectively. The characteristic function of U can be expressed as

$$\begin{aligned} C_U(T) &= E[\exp(i\mathrm{tr}TU)] \\ &= E_W\Big[E[\exp(i\mathrm{tr}TU)|W]\Big], \end{aligned}$$

where T is a real symmetric matrix whose (i,j) element is given by $(1+\delta_{ij})t_{ij}/2$. Here, δ_{ij} is the Kronecker delta, i.e., $\delta_{ii}=1$, $\delta_{ij}=0$ $(i\neq j)$. The conditional characteristic function can be evaluated as

$$\begin{aligned}
C_U(T|W) &= E[\exp(i\mathrm{tr}TU)|W] \\
&= \exp\left(-\frac{1}{\sqrt{p_2}}i\mathrm{tr}T(p_2I_{p_1}+n\Gamma^2)\right)\left|I_{p_1}-\frac{2i}{\sqrt{p_2}}T\right|^{-\frac{p_2}{2}} \\
&\quad \times\mathrm{etr}\left[\frac{i}{\sqrt{p_2}}\Gamma W\Gamma T\left(I_{p_1}-\frac{2i}{\sqrt{p_2}}T\right)^{-1}\right] \\
&= \mathrm{etr}\left(-T^2+i\sqrt{\frac{n}{p_2}}\Gamma G\Gamma T-2\frac{n}{p_2}\Gamma^2T^2\right)\times\{1+o^*(1)\},
\end{aligned}$$

where $G=\frac{1}{\sqrt{n}}(W-nI)$ and the notation o_i^* denotes a term that tends to 0 under a high-dimensional framework (1.4). Therefore,

$$\begin{aligned}
C_U(T) &= \int C_U(T|W)f(W)dW \\
&= \mathrm{etr}\left(-T^2-2\frac{n}{p_2}\Gamma^2T^2-\frac{n}{p_2}(\Gamma T\Gamma)^2\right)\times\{1+o^*(1)\}.
\end{aligned}$$

Similarly, the characteristic function of V can be expanded as

$$C_V(T) = \mathrm{etr}(-T^2)\times\{1+o^*(1)\}.$$

Using these results we can expand $C_{V,U}(T_1,T_2)$ of the joint characteristic function of V and U as follows:

$$\begin{aligned}
C_{V,U}(T_1,T_2) &= E[\exp(i\mathrm{tr}T_1V+i\mathrm{tr}T_2U)] \\
&= C_V(T_1)\times E_W\Big[C_U(T_2|W)\Big] \\
&= \mathrm{etr}(-T_1^2)\mathrm{etr}\left(-T_2^2-2\frac{n}{p_2}\Gamma^2T_2^2-\frac{n}{p_2}(\Gamma T_2\Gamma)^2\right) \\
&\quad \times\{1+o^*(1)\}.
\end{aligned}$$

Therefore we obtain the following theorem.

Theorem 2.1 *Under assumption* A1, *each of the elements of U and V is asymptotically and independently distributed as a normal distribution, more* *precisely*

$$v_{ii} \xrightarrow{d} N(0,2), \quad v_{ij} \xrightarrow{d} N(0,1), \quad i \neq j,$$
$$u_{ii} \xrightarrow{d} N\left(0, 2\left(1 + 2\frac{n}{p_2}\gamma_i^2 + \frac{n}{p_2}\gamma_i^4\right)\right), \quad u_{ij} \xrightarrow{d} N\left(0, 1 + \frac{n}{p_2}\gamma_i^4\right) \quad i \neq j.$$

where $\gamma_i^2 = \rho_i^2/(1-\rho_i^2)$ *and* $\xrightarrow{d}$ *denotes convergence in distribution.*

We can write $A_{uv}A_{vv}^{-1}A_{vu}$ and $A_{uu\cdot v}$ in terms of U and V as

$$A_{uv}A_{vv}^{-1}A_{vu} = p_2\left(I_{p_1} + \frac{n}{p_2}\Gamma^2\right) + \sqrt{p_2}U, \quad A_{uu\cdot v} = m\left(I_{p_1} + \frac{1}{\sqrt{m}}V\right), \tag{2.1}$$

and hence

$$\begin{aligned} Q &= A_{uu\cdot v}^{-1/2}A_{uv}A_{vv}^{-1}A_{vu}A_{uu\cdot v}^{-1/2} \\ &= \frac{1}{m}\left(I_{p_1} + \frac{1}{\sqrt{m}}V\right)^{-1/2}\left\{p_2\left(I_{p_1} + \frac{n}{p_2}\Gamma^2\right) + \sqrt{p_2}U\right\}\left(I_{p_1} + \frac{1}{\sqrt{m}}V\right)^{-1/2}. \end{aligned}$$

Therefore, Q can be expanded as

$$\begin{aligned} Q &= \frac{p_2}{m}\left(I_{p_1} - \frac{1}{\sqrt{m}}V + O_1^*)\right)\left\{\left(I_{p_1} + \frac{n}{p_2}\Gamma^2\right) + \sqrt{p_2}U\right\} \\ &\quad \times \left(I_{p_1} - \frac{1}{\sqrt{m}}V + O_1^*)\right) \\ &= \frac{p_2}{m}\left(I_{p_1} + \frac{n}{p_2}\Gamma^2\right) \\ &\quad + \frac{1}{\sqrt{m}}\left\{\sqrt{\frac{p_2}{m}} - \frac{1}{2}V\frac{p_2}{m}\left(I_{p_1} + \frac{n}{p_2}\Gamma^2\right) - \frac{1}{2}\frac{p_2}{m}\left(I_{p_1} + \frac{n}{p_2}\Gamma^2\right)V\right\} \\ &\quad + O_1^*. \end{aligned} \tag{2.2}$$

Here, the notation O_i^* denotes a term of the i-th order with respect to $(n^{-1}, p_2^{-1}, m^{-1})$.

2.1 Null distributions

In this section we consider the null distribution of the three test statistics under framework A1 and

$$\text{A1.1}: \quad \rho_1^2 > \cdots > \rho_k^2 > \rho_{k+1}^2 = \cdots = \rho_{p_1}^2 = 0. \tag{2.3}$$

Consider the transformed test statistics of LR, LH and BNP in (1.2) defined by

$$
\begin{aligned}
T_{LR} &= \sqrt{p_2}\left(1+\frac{m}{p_2}\right)\left\{\log\prod_{j=k+1}^{p_1}(1+\ell_j^2)-(p_1-k)\log\left(1+\frac{p_2}{m}\right)\right\},\\
T_{LH} &= \sqrt{p_2}\left\{\frac{m}{p_2}\sum_{j=k+1}^{p_1}\ell_j^2-(p_1-k)\right\}, \qquad (2.4)\\
T_{BNP} &= \sqrt{p_2}\left(1+\frac{m}{p_2}\right)\left\{\left(1+\frac{m}{p_2}\right)\sum_{j=k+1}^{p_1}\frac{\ell_j^2}{1+\ell_j^2}-(p_1-k)\right\}.
\end{aligned}
$$

Note that $\ell_1^2,\ldots,\ell_{p_1}^2$ are the characteristic roots of Q, and the three test statistics are symmetric functions of the last p_1-k characteristic roots $\ell_{k+1}^2,\ldots,\ell_{p_1}^2$. Using the fact that Q has a perturbation expansion as in (2.2), it can be seen (see Lawley (1956, 1959) and Fujikoshi (1977)) that the last p_1-k characteristic roots $\ell_{k+1}^2,\ldots,\ell_{p_1}^2$ are the characteristic roots of

$$
D=\frac{p_2}{m}I_{p_1-k}+\frac{1}{\sqrt{m}}\left(\sqrt{\frac{p_2}{m}}U_{22}-\frac{p_2}{m}V_{22}\right)+O_1^*, \qquad (2.5)
$$

where

$$
U=\begin{bmatrix} U_{11} & U_{12}\\ U_{21} & U_{22}\end{bmatrix},\ V=\begin{bmatrix} V_{11} & V_{12}\\ V_{21} & V_{22}\end{bmatrix},
$$

U_{22} and V_{22} are $(p_1-k)\times(p_1-k)$ matrices.

Using these results we can expand T_{LR}, T_{LH}, and T_{BNP} as follows:

$$
\begin{aligned}
T_{LR} &= \sqrt{p_2}\left(1+\frac{m}{p_2}\right)\left\{\sum_{j=k+1}^{p_1}\left\{\log\left(1+\frac{p_2}{m}+\ell_j^2-\frac{p_2}{m}\right)\right\}\right.\\
&\qquad\left.-(p_1-k)\log\left(1+\frac{p_2}{m}\right)\right\}\\
&= \sqrt{p_2}\left(1+\frac{m}{p_2}\right)\left\{\sum_{j=k+1}^{p_1}\left\{\log\left(1+\frac{p_2}{m}\right)+\frac{1}{1+\frac{p_2}{m}}\left(\ell_j^2-\frac{p_2}{m}\right)+O_{1/2}^*\right\}\right.\\
&\qquad\left.-(p_1-k)\log\left(1+\frac{p_2}{m}\right)\right\}\\
&= \mathrm{tr}\left(U_{22}-\sqrt{\frac{p_2}{m}}V_{22}\right)+O_{1/2}^*,
\end{aligned}
$$

$$
\begin{aligned}
T_{LH} &= \sqrt{p_2}\Bigg\{\frac{m}{p_2}\Bigg\{\frac{p_2}{m}(p_1-k)+\frac{1}{\sqrt{m}}tr\left(\sqrt{\frac{p_2}{m}}U_{22}-\frac{p_2}{m}V_{22}\right)+O^*_{1/2}\Bigg\} \\
&\qquad -(p_1-k)\Bigg\} \\
&= \mathrm{tr}\left(U_{22}-\sqrt{\frac{p_2}{m}}V_{22}\right)+O^*_{1/2},
\end{aligned}
$$

$$
\begin{aligned}
T_{BNP} &= \sqrt{p_2}\left(1+\frac{m}{p_2}\right)\Bigg\{\left(1+\frac{m}{p_2}\right)\left(\frac{p_2}{m}\left(1+\frac{m}{p_2}\right)^{-1}(p_1-k)\right. \\
&\quad +\frac{1}{\sqrt{m}}\left(1+\frac{m}{p_2}\right)^{-2}\mathrm{tr}\left(\sqrt{\frac{p_2}{m}}U_{22}-\frac{p_2}{m}V_{22}\right)\Bigg)-(p_1-k)+O^*_{1/2}\Bigg\} \\
&= \mathrm{tr}\left(U_{22}-\sqrt{\frac{p_2}{m}}V_{22}\right)+O^*_{1/2}.
\end{aligned}
$$

Each of the diagonal elements of U_{22} and V_{22} is asymptotically distributed as $N(0,2)$. Therefore, we obtain the following theorem.

Theorem 2.2 *Under assumptions* (1.4) *and* (2.3),

$$
\frac{T_G}{\sigma_G} \xrightarrow{d} N(0,1),
$$

where G=LR, LH, BNP, and

$$
\sigma_G = \sqrt{2(p_1-k)\left(1+\frac{p_2}{m}\right)}.
$$

2.2 Non-null distribution

In this section we derive the asymptotic non-null distributions of the three test statistics for dimensionality under the alternative hypothesis:

$$
\tilde{H}_0 : \rho_b > \rho_{b+1} = \ldots = \rho_{p_1} = 0, \qquad k < b \le p_1.
$$

For simplicity, we assume that the first b canonical correlations are differ, i.e.,

$$
\text{A1.2}: \quad \rho_1^2 > \cdots > \rho_b^2 > \rho_{b+1}^2 = \ldots = \rho_{p_1}^2 = 0. \tag{2.6}
$$

This is equivalent to $\gamma_1^2 > \cdots > \gamma_b^2 > \gamma_{b+1}^2 = \cdots = \gamma_{p_1}^2 = 0$. Note that $\ell_1^2, \ldots, \ell_{p_1}^2$ are the characteristic roots of $A_{uu\cdot v}^{-1} A_{uv} A_{vv}^{-1} A_{vu}$, which has a perturbation expansion in (2.2).

$$\begin{aligned} &\sqrt{m}\left\{\ell_j^2 - \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)\right\} \\ &= \sqrt{\frac{p_2}{m}} u_{jj} - \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right) v_{jj} + O_{1/2}^*, \quad j = k+1, \ldots, b. \end{aligned} \tag{2.7}$$

Further, from (2.6) the last $p_1 - b$ characteristic roots $\ell_{b+1}^2, \ldots, \ell_{p_1}^2$ are the characteristic roots of

$$\tilde{Q} = \frac{p_2}{m} I_{p_1 - b} + \frac{1}{\sqrt{m}}\left(\sqrt{\frac{p_2}{m}}\tilde{U}_{22} - \frac{p_2}{m}\tilde{V}_{22}\right) + O_{1/2}^*, \tag{2.8}$$

where

$$U = \begin{bmatrix} \tilde{U}_{11} & \tilde{U}_{12} \\ \tilde{U}_{21} & \tilde{U}_{22} \end{bmatrix}, \ V = \begin{bmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\ \tilde{V}_{21} & \tilde{V}_{22} \end{bmatrix},$$

and $\tilde{U}_{22}$ and $\tilde{V}_{22}$ are $(p_1 - b) \times (p_1 - b)$ matrices. Let

$$\begin{aligned} T_{LR}^* &= \sqrt{p_2}\left(1 + \frac{m}{p_2}\right)\left\{\log \prod_{j=k+1}^{p_1}(1 + \ell_j^2) \right. \\ &\qquad \left. - \log \prod_{j=k+1}^{p_1}\left(1 + \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)\right)\right\}, \\ T_{LH}^* &= \sqrt{p_2}\left\{\frac{m}{p_2}\sum_{j=k+1}^{p_1}\ell_j^2 - \frac{m}{p_2}\sum_{j=k+1}^{p_1}\frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)\right\}, \\ T_{BNP}^* &= \sqrt{p_2}\left(1 + \frac{m}{p_2}\right)\left\{\left(1 + \frac{m}{p_2}\right)\sum_{j=k+1}^{p_1}\frac{\ell_j^2}{1 + \ell_j^2} \right. \\ &\qquad \left. - \left(1 + \frac{m}{p_2}\right)\sum_{j=k+1}^{p_1}\frac{\frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)}{1 + \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)}\right\}. \end{aligned} \tag{2.9}$$

Using (2.7) and (2.8) we can express T^*_{LR}, T^*_{LH} and T^*_{BNP} as follows.

$$T^*_{LR} = \sum_{j=k+1}^{p_1} \left\{ \frac{1+\dfrac{p_2}{m}}{1+\dfrac{p_2}{m}\left(1+\dfrac{n}{p_2}\gamma_j^2\right)} \sqrt{\frac{m}{p_2}} \left(\sqrt{\frac{p_2}{m}} u_{jj} - \frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right) v_{jj} \right) \right\}$$
$$+ \operatorname{tr}\left(\tilde{U}_{22} - \sqrt{\frac{p_2}{m}}\tilde{V}_{22}\right) + O^*_{1/2},$$

$$T^*_{LH} = \sum_{j=k+1}^{b} \left\{ \sqrt{\frac{m}{p_2}} \left(\sqrt{\frac{p_2}{m}} u_{jj} - \frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right) v_{jj} \right) \right\}$$
$$+ \operatorname{tr}\left(\tilde{U}_{22} - \sqrt{\frac{p_2}{m}}\tilde{V}_{22}\right) + O^*_{1/2},$$

$$T^*_{BNP} = \sum_{j=k+1}^{p_1} \left\{ \frac{\left(1+\dfrac{p_2}{m}\right)^2}{\left(1+\dfrac{p_2}{m}\left(1+\dfrac{n}{p_2}\gamma_j^2\right)\right)^2} \sqrt{\frac{m}{p_2}} \left(\sqrt{\frac{p_2}{m}} u_{jj} - \frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right) v_{jj} \right) \right\}$$
$$+ \operatorname{tr}\left(\tilde{U}_{22} - \sqrt{\frac{p_2}{m}}\tilde{V}_{22}\right) + O^*_{1/2}.$$

Therefore we can combine the above three expressions as

$$\begin{aligned} T^*_G &= \sum_{j=k+1}^{b} d_j^c \left(u_{jj} - \sqrt{\frac{p_2}{m}}\left(1+\frac{n}{p_2}\gamma_j^2\right) v_{jj} \right) \\ &\quad + \operatorname{tr}\left(\tilde{U}_{22} - \sqrt{\frac{p_2}{m}}\tilde{V}_{22}\right) + O^*_{1/2}. \end{aligned}$$

where

$$d_j = \frac{1+\dfrac{p_2}{m}}{1+\dfrac{p_2}{m}\left(1+\dfrac{n}{p_2}\gamma_j^2\right)}.$$

Here, the notation G and c is used such that

$$c = \begin{cases} 1, & \text{when } G = LR, \\ 0, & \text{when } G = LH, \\ 2, & \text{when } G = BNP. \end{cases}$$

Using Theorems 2.1 and 2.2, we obtain the following theorem.

Theorem 2.3 *Let T_G^* be the transformed test statistics defined by* (2.9), *where $G = LR, LH, BNP$. Then, under assumptions* (1.4) *and* (2.6),

$$\frac{T_G^*}{\sigma_G^*} \xrightarrow{d} N(0,1),$$

where

$$\begin{aligned} {\sigma_G^*}^2 &= 2\sum_{j=k+1}^{b} d_j^{2c}\left\{\left(1+2\frac{n}{p_2}\gamma_j^2+\frac{n}{p_2}\gamma_j^4\right)\right\} \\ &\quad +2(p_1-b)\left(1+\frac{p_2}{m}\right). \end{aligned}$$

2.3 Asymptotic power

Based on the asymptotic distributions of the three statistics in Theorem 2.3, we obtain their asymptotic powers. Let $\delta_G = T_G - T_G^*$. Then

$$\delta_{LR} = \sqrt{m}\left(\frac{1+\frac{p_2}{m}}{\sqrt{\frac{p_2}{m}}}\right)\sum_{j=k+1}^{b}\log\left(1+\frac{\frac{p_2}{m}n\gamma_j^2}{\left(1+\frac{p_2}{m}\right)p_2}\right),$$

$$\delta_{LH} = \sqrt{p_2}\sum_{j=k+1}^{b}\frac{n\gamma_j^2}{p_2},$$

$$\delta_{BNP} = \sqrt{p_2}\left(1+\frac{p_2}{m}\right)\left\{\frac{1+\frac{p_2}{m}}{\frac{p_2}{m}}\sum_{j=k+1}^{b}\frac{\frac{p_2}{m}\left(1+\frac{p_2}{m}\right)}{1+\frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right)}-(b-k)\right\}.$$

We have

$$P_D = Pr(T_G > \sigma_G z_\alpha) = Pr(T_G^* > \sigma_G z_\alpha - \delta_G),$$

where z_α is the upper 100α % points of the standard normal distribution.

Using Theorem 2.3, the asymptotic power with a level of significance α is expressed as

$$\lim_{p_2\to\infty} P_D = \lim_{p_2\to\infty}\Phi\left(\frac{\delta_G-\sigma_G z_\alpha}{\sigma_G^*}\right),$$

where Φ is the distribution function of the standard normal distribution. Under assumption (1.4),

$$\frac{p_2}{m} = \frac{p_2}{n-p_2} = \frac{1}{\frac{n}{p_2}-1} \to \frac{1}{\frac{1}{c}-1} = \frac{c}{1-c} > 0.$$

Therefore, we can obtain $\delta_G \to \infty$ so that the asymptotic power is 1.

3 Distributions of tests for additional information

We are interested in the distribution of T in (1.5). According to Theorem 2 in Fujikoshi (1982), T under H_1 is expressed as a product of two independent variables, i.e.,

$$T = T_1 \times T_2, \quad T_1 \sim \Lambda(p_2, q_2, n - p_1), \quad T_2 \sim \Lambda(q_4, q_1, n - r), \tag{3.1}$$

where $p_1 = q_1 + q_2$, $p_2 = q_3 + q_4$, $r = q_1 + q_3$. Here, we denote the distribution of $\Lambda = |A|/|A + B|$ by $\Lambda(p, q, n)$, where A and B have independent Wishart distributions $W_p(n, \Sigma)$ and $W_p(q, \Sigma)$, respectively. Fujikoshi (1982) derived an asymptotic expansion for the distribution of T under A0. The approximation can be written as

$$P(-m \log T \le x) = G_f(x) + \frac{\beta}{m^2}\{G_{f+4}(x) - G_f(x)\} + O(m^{-2}), \tag{3.2}$$

where $p = q_1 + q_2 + q_3 + q_4$, $r = q_1 + q_3$, $f = (q_1 + q_2)(q_3 + q_4) - q_1 q_3$,

$$\begin{aligned} m &= n - \frac{1}{2}(p+1) - \frac{1}{2}\frac{q_1 q_3 (p - r)}{f}, \\ \beta &= \frac{1}{48}\Bigg[\Big\{q_1^2 + q_2^2 + q_3^2 + q_4^2 - 5\Big\} f + 2q_1^2 q_2 p_2 + 2p_1 q_3^2 q_4 \\ &\qquad + 2q_2 q_4 \Big\{q_1 q_2 + q_3 q_4 - 3q_1 q_3\Big\} - 3(q_1 q_3)^2 (p - r)^2 / f\Bigg]. \end{aligned}$$

Let Λ be a statistic that is distributed as $\Lambda(p, q, n)$. Tonda and Fujikoshi (2004) derived an asymptotic expansion formula of the distribution of Λ when q is fixed, $n \to \infty$, $p \to \infty$ with $p/n \to c \in (0, 1)$. For our derivation, we use their result. Let

$$T_F = \frac{-\log \Lambda - m_F}{d_F},$$

where

$$m_F = \sum_{j=1}^{q} \log \frac{n + j}{n - p + j}, \qquad d_F^2 = \frac{2}{p} \sum_{j=1}^{q} \frac{p^2}{(n + j)(n - p + j)}.$$

Then, the characteristic function of T_F can be expanded as

$$C_{T_F}(t) = e^{-\frac{1}{2}t^2}\left\{1+\sum_{\alpha=1}^{2}\kappa_F^{(2\alpha-1)}(it)^{2\alpha-1}+\sum_{\alpha=1}^{3}\kappa_F^{(2\alpha)}(it)^{2\alpha}\right\}+o(p^{-1}),$$

where $\kappa_F^{(\alpha)}$s are defined by

$$\begin{aligned}
\kappa_F^{(1)} &= \frac{1}{\sqrt{p}}\tau_1, \quad \kappa_F^{(3)} = \frac{1}{\sqrt{p}}\tau_3, \quad \kappa_F^{(2)} = \frac{1}{p}\left(\tau_2+\frac{\tau_1^2-\tau_{(11)}}{2}\right),\\
\kappa_F^{(4)} &= \frac{1}{p}\left(\tau_4+\tau_1\tau_3-\tau_{(13)}\right), \quad \kappa_F^{(6)} = \frac{1}{p}\left(\tau_6+\frac{\tau_3^2-\tau_{(33)}}{2}\right),\\
\tau_i &= \sum_{k=1}^{q}\omega_k^i\tau_{ik}, \quad \tau_{(ij)} = \sum_{k=1}^{q}\omega_k^{i+j}\tau_{ik}\tau_{jk}.
\end{aligned}$$

Here the coefficients τ_{ij} and ω_j are given by

$$\begin{aligned}
\tau_{1j} &= \frac{a_j}{\sqrt{2(1+a_j)}}, \quad \tau_{3j} = \frac{2+a_j}{3\sqrt{2(1+a_j)}}, \quad \tau_{2j} = \frac{a_j(4+3a_j)}{4(1+a_j)},\\
\tau_{4j} &= \frac{3+5a_j+2a_j^2}{6(1+a_j)}, \quad \tau_{6j} = \frac{(2+a_j)^2}{36(1+a_j)}, \quad \omega_j = \frac{a_j}{d\sqrt{1+a_j}},
\end{aligned}$$

where $a_j = p/(n-p+j)$, $\quad d = \sum_{j=1}^{q} p^2\{(n+j)(n-p+j)\}^{-1}$.

Wakaki (2006) derived an asymptotic expansion formula of the distribution of Λ when all three values of p, q and n tend to infinity with $p/n \to c_1 \in (0,1)$ and $q/n \to c_2 \in (0,1)$. For our derivation, we use his result. Let

$$T_W = \frac{-\log\Lambda - m_W}{d_W},$$

where $m_W = \tau^{(1)}$, $d_W^2 = \tau^{(2)}$,

$$\begin{aligned}
\tau^{(s)} &= (-1)^s\left\{\psi_q^{(s-1)}\left(\frac{n-p+q}{2}\right)-\psi_q^{(s-1)}\left(\frac{n+q}{2}\right)\right\},\\
\psi_q^{(s)} &= \sum_{j=1}^{q}\psi^{(s)}\left(a-\frac{j-1}{2}\right), \quad (s=0,1,\ldots;a>0),
\end{aligned}$$

and $\psi^{(s)}(a)$ is the polygamma function defined as

$$\psi^{(s)}(a) = \left(\frac{d}{da}\right)^s\log\Gamma(a) = \begin{cases} -C+\sum_{k=0}^{\infty}\left(\frac{1}{1+k}-\frac{1}{k+a}\right), & (s=0),\\ \sum_{k=0}^{\infty}\frac{(-1)^{s+1}s!}{(k+a)^{s+1}}, & (s=1,2,\ldots). \end{cases}$$

Then, the characteristic function of T_W can be expanded as

$$C_{T_W}(t) = e^{-\frac{1}{2}t^2}\left\{1 + \sum_{\alpha=1}^{2} \kappa_W^{(2\alpha-1)}(it)^{2\alpha-1} + \sum_{\alpha=1}^{3} \kappa_W^{(2\alpha)}(it)^{2\alpha}\right\} + o(p^{-1}),$$

where $\kappa_W^{(\alpha)}$s are defined by $\kappa_W^{(1)} = \kappa_W^{(2)} = 0$,

$$\kappa_W^{(3)} = \tau^{(3)}/\left(\tau^{(2)}\right)^{3/2}, \quad \kappa_W^{(4)} = \tau^{(4)}/\left(\tau^{(2)}\right)^2, \quad \kappa_W^{(6)} = \left(\kappa_W^{(3)}\right)^2.$$

Using these results, we can obtain the asymptotic distribution of T in (1.5) under various high-dimensional frameworks satisfying A2.

3.1 Null distribution under A2

In our framework A2, the conditions "$p \to \infty$" and "one of $q_1, q_2, q_3, q_4 \to \infty$" are equivalent. Under $p_1 \leq p_2$, the condition can be realized as one of the following 12 cases.

(i) q_1:fixed, q_2:fixed, q_3:fixed, $q_4 \to \infty$.
(ii) q_1:fixed, q_2:fixed, $q_3 \to \infty$, q_4:fixed.
(iii) q_1:fixed, q_2:fixed, $q_3 \to \infty$, $q_4 \to \infty$.
(iv) q_1:fixed, $q_2 \to \infty$, q_3:fixed, $q_4 \to \infty$.
(v) q_1:fixed, $q_2 \to \infty$, $q_3 \to \infty$, q_4:fixed.
(vi) q_1:fixed, $q_2 \to \infty$, $q_3 \to \infty$, $q_4 \to \infty$.
(vii) $q_1 \to \infty$, q_2:fixed, q_3:fixed, $q_4 \to \infty$.
(viii) $q_1 \to \infty$, q_2:fixed, $q_3 \to \infty$, q_4:fixed.
(ix) $q_1 \to \infty$, q_2:fixed, $q_3 \to \infty$, $q_4 \to \infty$.
(x) $q_1 \to \infty$, $q_2 \to \infty$, q_3:fixed, $q_4 \to \infty$.
(xi) $q_1 \to \infty$, $q_2 \to \infty$, $q_3 \to \infty$, q_4:fixed.
(xii) $q_1 \to \infty$, $q_2 \to \infty$, $q_3 \to \infty$, $q_4 \to \infty$.

We can obtain an asymptotic expansion of the distribution of T in all of the cases except (ii). Note that even in situation (ii), our approximations are good based on the numerical simulation.

We have seen that T_1 and T_2 have the following asymptotic means and variances:

$$\mathrm{E}(-\log T_j) \approx m_j, \qquad \mathrm{Var}(-\log T_j) \approx d_j,$$

and hence

$$\mathrm{E}(-\log T) \approx m_1 + m_2, \qquad \mathrm{Var}(-\log T) \approx d,$$

where $d = \left(d_1^2 + d_2^2\right)^{-1/2}$. Note that m_j and d_j are given by Tonda and Fujikoshi (2004) and Wakaki (2006), respectively. Let T_H be the standardization of T defined by

$$T_H = \frac{-\log T - (m_1 + m_2)}{d}.$$

Noting that $T_H = w_1\tilde{T}_1 + w_2\tilde{T}_2$ with $w_j = d_j/d$, the characteristic function $C_H(t)$ of T_H is expressed as

$$C_{T_H}(t) = C_{\tilde{T}_1}(w_1 t)C_{\tilde{T}_2}(w_2 t),$$

where $\tilde{T}_j$ are defined by the standardization of $-\log T_j$. Then the characteristic function of T_H is expressed by

$$C_{T_H}(t) = e^{-\frac{1}{2}t^2}\left\{1 + \sum_{\alpha=1}^{2}\kappa^{(2\alpha-1)}(it)^{2\alpha-1} + \sum_{\alpha=1}^{3}\kappa^{(2\alpha)}(it)^{2\alpha}\right\} + o(p^{-1}),$$

where

$$\begin{aligned}&\kappa^{(1)} = \tilde{\kappa}_1^{(1)} + \tilde{\kappa}_2^{(1)}, \quad \kappa^{(3)} = \tilde{\kappa}_1^{(3)} + \tilde{\kappa}_2^{(3)}, \quad \kappa^{(2)} = \tilde{\kappa}_1^{(2)} + \tilde{\kappa}_2^{(2)} + \tilde{\kappa}_1^{(1)}\tilde{\kappa}_2^{(1)},\\ &\kappa^{(4)} = \tilde{\kappa}_1^{(4)} + \tilde{\kappa}_2^{(4)} + \tilde{\kappa}_1^{(1)}\tilde{\kappa}_2^{(3)} + \tilde{\kappa}_1^{(3)}\tilde{\kappa}_2^{(1)}, \quad \kappa^{(6)} = \tilde{\kappa}_1^{(6)} + \tilde{\kappa}_2^{(6)} + \tilde{\kappa}_1^{(3)}\tilde{\kappa}_2^{(3)}.\end{aligned} \tag{3.3}$$

Here $\tilde{\kappa}_j^{(\alpha)}$'s are defined by $\tilde{\kappa}_j^{(\alpha)} = w_j^{\alpha}\kappa_j^{(\alpha)}$, where $\kappa_j^{(\alpha)}$'s are given by Tonda and Fujikoshi (2004), Wakaki (2006).

Let T_{H1} denote the standardization of T under (i), (iii), and (viii), T_{H2} under (iv), (vi), and (xi), T_{H3} under (v), T_{H4} under (vii), and (ix), and T_{H5} under (x) and (xii). For convenience, let (q_{inf}, q_{fix}) be defined by

$$(q_{inf}, q_{fix}) = \begin{cases}(q_4, q_1) & \text{under } q_4 \to \infty,\ q_1 : \text{fixed},\\ (q_1, q_4) & \text{under } q_1 \to \infty,\ q_4 : \text{fixed}.\end{cases}$$

Then, m_j, d_j, $\kappa_j^{(\alpha)}$ are obtained from the following table.

	$m_1, d_1, \kappa_1^{(\alpha)}$	$m_2, d_2, \kappa_2^{(\alpha)}$
T_{H1}	$F(p_2, q_2, n - p_1)$	$F(q_{inf}, q_{fix}, n - r)$
T_{H2}	$W(p_2, q_2, n - p_1)$	$F(q_{inf}, q_{fix}, n - r)$
T_{H3}	$F(p_1, q_4, n - p_2)$	$W(q_2, q_3, n - r)$
T_{H4}	$F(p_2, q_2, n - p_1)$	$W(q_4, q_1, n - r)$
T_{H5}	$W(p_2, q_2, n - p_1)$	$W(q_4, q_1, n - r)$

Here, $F(a, b, c)$ and $W(a, b, c)$ show that $(m_j, d_j, \kappa_j^{(\alpha)})$ are defined from $(m_F, d_F, \kappa_F^{(\alpha)})$ or $(m_W, d_W, \kappa_W^{(\alpha)})$ by substituting (n, p, q) for (a, b, c), respectively. Using $\Lambda(p, q, n) = \Lambda(q, p, n + q - p)$, cases (viii) and (xi) can be obtained. Note that case (v) is obtained by using the fact that

$$T = T_1' \times T_2', \quad T_1' \sim \Lambda(p_1, q_4, n - p_1), \quad T_2' \sim \Lambda(q_2, q_3, n - r),$$

Finally, by inverting the characteristic function of T_G we have the following theorem.

Theorem 3.1 *Let T_G be the normalization of T defined by* (1.5)*. Then, the null distribution of T_G can be expanded as*

$$P(T_G \leq x) = \Phi(x) - \phi(x)\Big[a_1(x) + a_2(x)\Big] + o(p^{-1}), \qquad (3.4)$$

where $G = H1 \sim H5$, $\Phi(x)$ and $\phi(x)$ are the distribution and density function of the standard normal distribution, respectively and $a_j(x)$'s are defined by

$$a_1(x) = \kappa^{(1)} + \kappa^{(3)}h_2(x), \quad a_2(x) = \kappa^{(2)}h_1(x) + \kappa^{(4)}h_3(x) + \kappa^{(6)}h_5(x). \quad (3.5)$$

Here, $h_j(x)$ is the jth Hermite polynomial; in particular, $h_1(x) = x$, $h_2(x) = x^2 - 1$, $h_3(x) = x^3 - x$, $h_4(x) = x^4 - 6x^2 + 3$, $h_5(x) = x^5 - 10x^3 + 15x$, and $\kappa^{(\alpha)}s$ are given by (3.3)*.*

Using the coefficients $a_j(x)$s of the asymptotic expansion (3.4), we can obtain the Cornish-Fisher expansion. Let x and $\tilde{t}(x)$ denote the percentage point of the limiting distribution of T_G and the corresponding percentage point of T_G, respectively, that is

$$P(T_G \leq \tilde{t}(x)) = \Phi(x), \qquad G = H1 \sim H5.$$

Then from (3.4), $\tilde{t}_G(x)$ can be expanded as

$$\begin{aligned}\tilde{t}(x) &= x + a_1(x) + \left\{a_1(x)a_1'(x) - \frac{1}{2}xa_1(x)^2 + a_2(x)\right\} + o(p^{-1}) \\ &= \tilde{t}_G(x) + o(p^{-1}) \qquad (3.6)\end{aligned}$$

4 Simulation results

In this section we compare our high-dimensional approximations (denoted as H) with the classical approximations (denoted as C) based on the asymptotic distribution under a large sample framework such that p_1 and p_2 are fixed and n tends to infinity. The numerical accuracy is studied for the upper percentage points and the actual test size.

4.1 Null distributions of tests for dimensionality

It is well-known that under the large sample framework, the three statistics

$$-n\log\prod_{j=k+1}^{p_1}(1-r_j^2), \quad n\sum_{j=k+1}^{p_1}\frac{r_j^2}{1-r_j^2}, \quad n\sum_{j=k+1}^{p_1}r_j^2.$$

are asymptotically distributed as the χ^2-distribution with $(p_1 - k)(p_2 - k)$ degrees of freedom (e.g., see Siotani et al. (1985)). Under the high-dimensional framework the three statistics T_G/σ_G are distributed asymptotically (see Theorem 2.2) as $N(0,1)$, where $\sigma_G = \sqrt{2(p_1-k)(1+p_2/m)}$, $G = LR, LH, BNP$. To facilitate, let $t_C = n^{-1}\chi^2_{(p_1-k)(p_2-k),\alpha}$,

$$t_{LR\cdot H} = (p_1-k)\log\left(1+\frac{p_2}{m}\right) + p_2^{-1/2}\left(1+\frac{m}{p_2}\right)^{-1} z_\alpha,$$

$$t_{LH\cdot H} = \frac{p_2}{m}\left\{(p_1-k) + \sigma_{LH}\times p_2^{-1/2} z_\alpha\right\},$$

$$t_{BNP\cdot H} = \left(1+\frac{m}{p_2}\right)^{-1}\left\{(p_1-k) + \sigma_{BNP}\times p_2^{-1/2}\left(1+\frac{m}{p_2}\right)^{-1} z_\alpha\right\},$$

where $\chi^2_{(p_1-k)(p_2-k),\alpha}$ and z_α are $100(1-\alpha)\%$ point of the χ^2-distribution with $(p_1-k)(p_2-k)$ degrees of freedom and the standard normal distribution, respectively.

The values of p_1, n, p_2 and $\mathcal{P}$ were chosen as follows:

$$\begin{aligned}(p_1+p_2, n);\ &(10,50),(20,50),(30,50),(40,50),(10,100),(15,100),\\ &(20,100),(50,100),(70,100),(90,100)(10,100),\\ &(30,200),(50,200),(80,200),(100,200)(120,200),\\ &(150,200),(170,200),(190,200),\\ (p_1,\mathcal{P}) :\ &(3,\mathrm{diag}(0.9,0.6,0.0)),(p_1,\mathcal{P}) : (4,\mathrm{diag}(0.9,0.6,0.0,0.0)),\end{aligned}$$

Table 1 shows the estimated upper 5% points based on a Monte Carlo simulation, the approximated critical points using our method, $t_{LR\cdot H}$, $t_{LH\cdot H}$, $t_{BNP\cdot H}$, and the classical approximations t_C. Table 2 shows the corresponding actual test sizes. We are interested in the behavior when the dimension is large and close to the sample size.

From Tables 1 and 2, the chi-square type approximation t_C, $\alpha_{LR\cdot C}$, $\alpha_{LH\cdot C}$, $\alpha_{BNP\cdot C}$ performs well when p is less than 20. However, the chi-square type approximation is poor when p is greater than 20. When p is large, $\alpha_{LR\cdot C}$ and $\alpha_{LH\cdot C}$ are close to 1 and $\alpha_{BNP\cdot C}$ is close to 0. The normal type approximation $\alpha_{LR\cdot C}$, $\alpha_{LH\cdot C}$ and $\alpha_{BNP\cdot C}$ performs well when the dimension p is close to half of N. When $k=2$, $\alpha_{LH\cdot C}$ performs well when the dimension p close to N.

Table 1: Upper 5% point

$p_1 = 3, \quad k = 2, \quad , \mathcal{P} = \text{diag}(0.9, 0.6, 0.0)$

N	p	p_2	Simu$_{LR}$	$t_{LR \cdot H}$	Simu$_{LH}$	$t_{LH \cdot H}$	Simu$_{BNP}$	$t_{BNP \cdot H}$	t_C
50	10	7	0.24**	0.29*	0.28**	0.32*	0.23**	0.28*	0.23
50	20	17	0.60**	0.65*	0.83	0.88*	1.30	1.69	0.51
50	30	27	1.04	1.11*	1.83	1.91*	3.63	5.12	0.77
50	40	37	1.71	1.79*	4.67	4.65*	7.69	11.44	1.02
100	10	7	0.12**	0.14*	0.12**	0.15*	0.09**	0.10*	0.11
100	15	12	0.20**	0.22*	0.22**	0.24*	0.22**	0.25*	0.18
100	20	17	0.28**	0.29*	0.32**	0.33*	0.43	0.48*	0.25
100	50	47	0.82	0.84*	1.28	1.28*	4.10	4.84	0.62
100	70	67	1.33	1.38*	2.82	2.86*	8.82	11.97	0.86
100	90	87	2.35	2.41*	9.84	9.73*	15.61	23.85	1.09
200	10	7	0.06**	0.07*	0.06**	0.07*	0.04**	0.04*	0.06
200	30	27	0.20**	0.21*	0.22**	0.23*	0.34	0.37*	0.19
200	50	47	0.35**	0.36*	0.42	0.43*	1.12	1.23	0.31
200	80	77	0.61	0.61*	0.83	0.83*	3.65	3.99	0.48
200	100	97	0.80	0.81*	1.23	1.22*	6.34	7.13	0.60
200	120	117	1.04	1.05*	1.83	1.81*	10.11	11.55	0.71
200	150	147	1.51	1.53*	3.56	3.54*	16.97	21.03	0.87
200	170	167	2.01	2.03*	6.52	6.49*	22.15	29.55	0.98
200	190	187	2.99	3.03*	19.59	19.28*	26.27	40.06	1.09

* Denotes $|\text{Simu}_G - t_{G \cdot H}| \leq 10^{-1}$, where $G = LR, LH, BNP$.

** Denotes $|\text{Simu}_G - t_C| \leq 10^{-1}$, where $G = LR, LH, BNP$.

$p_1 = 4, \quad k = 2, \quad , \mathcal{P} = \text{diag}(0.9, 0.6, 0.0, 0.0)$

N	p	p_2	Simu_{LR}	$t_{LR\cdot H}$	Simu_{LH}	$t_{LH\cdot H}$	Simu_{BNP}	$t_{BNP\cdot H}$	t_C
50	10	7	0.24**	0.29*	0.28**	0.32*	0.23**	0.28*	0.23
50	20	17	0.60**	0.65*	0.83	0.88*	1.30	1.69	0.51
50	30	27	1.04	1.11*	1.83	1.91*	3.63	5.12	0.77
50	40	37	1.71	1.79*	4.67	4.65*	7.69	11.44	1.02
100	10	7	0.12**	0.14*	0.12**	0.15*	0.09**	0.10*	0.11
100	15	12	0.20**	0.22*	0.22**	0.24*	0.22**	0.25*	0.18
100	20	17	0.28**	0.29*	0.32**	0.33*	0.43	0.48*	0.25
100	50	47	0.82	0.84*	1.28	1.28*	4.10	4.84	0.62
100	70	67	1.33	1.38*	2.82	2.86*	8.82	11.97	0.86
100	90	87	2.35	2.41*	9.84	9.73*	15.61	23.85	1.09
200	10	7	0.06**	0.07*	0.06**	0.07*	0.04**	0.04*	0.06
200	30	27	0.20**	0.21*	0.22**	0.23*	0.34	0.37*	0.19
200	50	47	0.35**	0.36*	0.42	0.43*	1.12	1.23	0.31
200	80	77	0.61	0.61*	0.83	0.83*	3.65	3.99	0.48
200	100	97	0.80	0.81*	1.23	1.22*	6.34	7.13	0.60
200	120	117	1.04	1.05*	1.83	1.81*	10.11	11.55	0.71
200	150	147	1.51	1.53*	3.56	3.54*	16.97	21.03	0.87
200	170	167	2.01	2.03*	6.52	6.49*	22.15	29.55	0.98
200	190	187	2.99	3.03*	19.59	19.28*	26.27	40.06	1.09

* Denotes $|\text{Simu}_G - t_{G\cdot H}| \leq 10^{-1}$, where $G = LR, LH, BNP$.

** Denotes $|\text{Simu}_G - t_C| \leq 10^{-1}$, where $G = LR, LH, BNP$.

Table 2: The corresponding actual test size.
$p_1 = 3, \quad k = 2, \quad , \mathcal{P} = \text{diag}(0.9, 0.6, 0.0)$

N	p	p_2	$\alpha_{LR \cdot H}$	$\alpha_{LR \cdot C}$	$\alpha_{LH \cdot H}$	$\alpha_{LH \cdot C}$	$\alpha_{BNP \cdot H}$	$\alpha_{BNP \cdot C}$
50	10	7	0.021	0.070	0.023	0.102	0.017	0.041*
50	20	17	0.026	0.155	0.035	0.351	0.017	0.016
50	30	27	0.024	0.406	0.040*	0.787	0.012	0.001
50	40	37	0.025	0.837	0.051*	0.991	0.004	0.000
100	10	7	0.023	0.058*	0.025	0.074	0.021	0.046*
100	15	12	0.032	0.074	0.035	0.111	0.028	0.043*
100	20	17	0.038	0.097	0.042*	0.173	0.032	0.041*
100	50	47	0.039	0.465	0.052*	0.845	0.025	0.006
100	70	67	0.026	0.895	0.044*	0.998	0.011	0.000
100	90	87	0.024	1.000	0.054*	1.000	0.002	0.000
200	10	7	0.024	0.056*	0.025	0.062	0.024	0.050*
200	30	27	0.034	0.087	0.038	0.159	0.031	0.037
200	50	47	0.041*	0.168	0.046	0.393	0.035	0.030
200	80	77	0.045*	0.451	0.054*	0.865	0.036	0.018
200	100	97	0.043*	0.727	0.054*	0.984	0.033	0.007
200	120	117	0.044*	0.924	0.057*	1.000	0.030	0.002
200	150	147	0.036	0.999	0.054*	1.000	0.018	0.000
200	170	167	0.029	1.000	0.053*	1.000	0.009	0.000
200	190	187	0.026	1.000	0.061	1.000	0.003	0.000

* Denotes the approximation in $[0.040, 0.060]$

$p_1 = 4, \quad k = 2, \quad , \mathcal{P} = \mathrm{diag}(0.9, 0.6, 0.0, 0.0)$

N	p	p_2	$\alpha_{LR \cdot H}$	$\alpha_{LR \cdot C}$	$\alpha_{LH \cdot H}$	$\alpha_{LH \cdot C}$	$\alpha_{BNP \cdot H}$	$\alpha_{BNP \cdot C}$
50	10	7	0.011	0.072	0.016	0.114	0.006	0.039
50	20	17	0.023	0.209	0.038	0.502	0.011	0.019
50	30	27	0.022	0.623	0.053*	0.948	0.008	0.004
50	40	37	0.034	0.979	0.092	1.000	0.007	0.000
100	10	7	0.009	0.061	0.012	0.078	0.007	0.044*
100	15	12	0.018	0.078	0.022	0.131	0.014	0.039
100	20	17	0.025	0.115	0.032	0.214	0.018	0.040*
100	50	47	0.028	0.684	0.051*	0.975	0.017	0.008
100	70	67	0.033	0.990	0.063	1.000	0.016	0.000
100	90	87	0.039	1.000	0.109	1.000	0.007	0.000
200	10	7	0.011	0.051*	0.012	0.058*	0.010	0.044*
200	30	27	0.029	0.110	0.035	0.221	0.024	0.042*
200	50	47	0.035	0.244	0.043*	0.572	0.028	0.031
200	80	77	0.041*	0.676	0.055*	0.982	0.032	0.020
200	100	97	0.037	0.925	0.054*	1.000	0.025	0.009
200	120	117	0.041*	0.996	0.063	1.000	0.027	0.003
200	150	147	0.036	1.000	0.062	1.000	0.019	0.000
200	170	167	0.036	1.000	0.076	1.000	0.014	0.000
200	190	187	0.045*	1.000	0.126	1.000	0.007	0.000

* Denotes the approximation in $[0.040, 0.060]$

4.2 Additional information

The Cornish-Fisher expansion of the large sample approximation (3.2) is well known. The expansion is obtained by using an alternative form of (3.2) written as

$$P(-m \log T \leq x) = G_f(x) + g_f(x)\frac{1}{m^2}\tilde{p}_1(x) + O(m^{-2}), \tag{4.1}$$

where $g_f(x)$ is the density function of the chi-square variable with f degrees of freedom and coefficient $\tilde{p}_1(x)$ is defined by

$$\tilde{p}_1(x) = \beta \sum_{i=1}^{2} \frac{2x^i}{\prod_{j=1}^{i} f + 2(j-1)}.$$

Similarly let

$$P(-m \log T \leq \tilde{t}(x)) = G_f(x).$$

Then the Cornish-Fisher expansion can be written in the same way as in (3.6), that is,

$$\begin{aligned} \tilde{t}(x) &= x + \frac{1}{m^2}\tilde{p}_1(x) + O(m^{-2}) \\ &= \tilde{t}_C + O(m^{-2}). \end{aligned} \tag{4.2}$$

For comparison, let $t_C = m^{-1} \times \tilde{t}_C$ $t_G = m_1 + m_2 + d \times \tilde{t}_G(x)$ where $G = H1 \sim H5$.

Table 3 gives the upper 5% of the points based on a Monte Carlo simulation (Simu), and the approximated critical points of our method, $t_{H1} \sim t_{H5}$, and the classical approximations t_{A0}. Table 4 gives the corresponding actual test sizes. We are interested in the behavior when the dimension is large and close to the sample size.

From Tables 3 and 4, the chi-square type approximation t_C, α_C performs well when p is less than 8. In contrast, the chi-square type approximations are poor when the smallest of q_1, q_2, q_3, and q_4 is large. When p is large, the normal type approximation $t_{H1} \sim t_{H5}$, $\alpha_{H1} \sim \alpha_{H5}$ performs better than the chi-square type approximation. Furthermore, when the sample size is much larger than the dimension, the performance of the normal type approximation is similar to that of a large sample approximation. In particular, the approximation t_{H5}, α_{H5} is the best of these approximation for all cases.

Table 3: Upper 5% point

N	p	q_1	q_2	q_3	q_4	Simu	t_C	t_{H1}	t_{H2}	t_{H3}	t_{H4}	t_{H5}
50	8	2	2	2	2	0.47	0.47*	0.47*	0.47*	0.47*	0.47*	0.47*
50	48	2	2	2	42	13.12	10.36	12.78	13.08*	12.94	12.88	13.17*
50	48	2	2	42	2	8.97	6.55	8.60	8.93*	8.77	8.62	8.95*
50	48	2	2	22	22	12.05	10.14	11.63	11.93	11.78	11.72	12.02*
50	48	2	22	2	22	31.83	28.70	31.52	31.88*	31.52	31.52	31.89*
50	48	2	22	22	2	30.75	27.75	30.37	30.74*	30.52	30.37	30.74*
50	48	22	2	2	22	30.68	27.75	30.37	30.59*	30.37	30.52	30.74*
50	48	22	2	22	2	10.79	10.21	10.53	10.82*	10.58	10.58	10.86*
50	48	12	12	12	12	28.05	27.69	27.61	27.97*	27.62	27.62	27.98*
100	8	2	2	2	2	0.22	0.22*	0.22*	0.22*	0.22*	0.22*	0.22*
100	48	12	12	12	12	6.81	6.88*	6.81*	6.82*	6.81*	6.81*	6.82*
100	88	2	2	42	42	7.98	7.55	7.96*	7.98*	7.98*	7.97*	8.00*
100	88	2	42	2	42	42.28	40.69	42.22*	42.28*	42.22*	42.22*	42.28*
100	88	2	42	42	2	41.22	39.69	41.14*	41.19*	41.18*	41.14*	41.19*
100	88	42	2	2	42	41.20	39.69	41.14*	41.15*	41.14*	41.18*	41.19*
100	88	42	2	42	2	6.90	6.83*	6.87*	6.89*	6.87*	6.87*	6.89*
100	96	24	24	24	24	51.66	51.45	51.39	51.58*	51.39	51.39	51.59*
100	98	2	2	2	92	16.10	10.92	15.72	16.04*	15.90	15.83	16.13*
100	98	2	2	92	2	10.41	6.26	10.06	10.41*	10.26	10.09	10.44*
100	98	2	32	32	32	15.13	15.15*	15.14*	15.14*	15.14*	15.14*	15.14*
100	98	32	2	32	32	45.60	41.89	45.12	45.34	45.12	45.28	45.50*

* Denotes the approximation in Simu $\pm 10^{-1}$.

N	p	q_1	q_2	q_3	q_4	Simu	t_C	t_{H1}	t_{H2}	t_{H3}	t_{H4}	t_{H5}
200	8	2	2	2	2	0.11	0.11*	0.11*	0.11*	0.11*	0.11*	0.11*
200	96	24	24	24	24	12.87	12.99	12.87*	12.87*	12.87*	12.87*	12.87*
200	144	36	36	36	36	36.49	37.42	36.51*	36.52*	36.51*	36.51*	36.52*
200	148	2	2	2	142	5.76	5.61	5.76*	5.76*	5.76*	5.76*	5.76*
200	148	2	2	142	2	3.12	3.03*	3.12*	3.12*	3.12*	3.12*	3.12*
200	148	2	2	72	72	4.85	4.82*	4.85*	4.85*	4.85*	4.85*	4.85*
200	148	2	72	2	72	48.12	47.65	48.10*	48.11*	48.10*	48.10*	48.11*
200	148	2	72	72	2	47.24	46.78	47.21*	47.22*	47.22*	47.21*	47.22*
200	148	72	2	2	72	47.20	46.78	47.21*	47.21*	47.21*	47.22*	47.22*
200	148	72	2	72	2	3.95	3.95*	3.94*	3.94*	3.94*	3.94*	3.94*
200	158	2	52	52	52	52.22	51.50	52.18*	52.19*	52.18*	52.18*	52.19*
200	158	52	2	52	52	33.82	33.91*	33.83*	33.83*	33.83*	33.84*	33.84*
200	188	2	2	92	92	10.76	9.06	10.70*	10.73*	10.73*	10.72*	10.75*
200	188	2	92	2	92	104.51	97.34	104.46*	104.53*	104.46*	104.46*	104.53*
200	188	2	92	92	2	103.36	96.24	103.23	103.29*	103.28*	103.23	103.29*
200	188	92	2	2	92	103.28	96.24	103.23*	103.24*	103.23*	103.28*	103.29*
200	188	92	2	92	2	9.48	8.63	9.47*	9.50*	9.48*	9.48*	9.50*
200	188	2	62	62	62	95.03	88.41	95.00*	95.07*	95.01*	95.00*	95.07*
200	188	62	2	62	62	66.62	64.23	66.55*	66.56*	66.55*	66.59*	66.61*
200	192	48	48	48	48	98.99	99.11	98.84	98.94*	98.84	98.84	98.94*
200	198	2	2	2	192	18.942	11.09	18.576	18.896*	18.767	18.689	18.999*
200	198	2	2	192	2	11.914	6.008	11.491	11.846*	11.692	11.52	11.871*

* Denotes the approximation in Simu $\pm 10^{-1}$.

Table 4: The corresponding actual test size.

N	p	q_1	q_2	q_3	q_4	α_C	α_{H1}	α_{H2}	α_{H3}	α_{H4}	α_{H5}
50	8	2	2	2	2	0.051*	0.052*	0.051*	0.052*	0.052*	0.051*
50	48	2	2	2	42	0.502	0.072	0.052*	0.059*	0.063	0.047*
50	48	2	2	42	2	0.472	0.075	0.052*	0.062	0.073	0.051*
50	48	2	2	22	22	0.281	0.080	0.057*	0.068	0.074	0.052*
50	48	2	22	2	22	0.420	0.065	0.048*	0.065	0.065	0.048*
50	48	2	22	22	2	0.389	0.071	0.051*	0.063	0.071	0.051*
50	48	22	2	2	22	0.383	0.066	0.055*	0.066	0.057*	0.047*
50	48	22	2	22	2	0.097	0.067	0.049*	0.064	0.064	0.046*
50	48	12	12	12	12	0.070	0.074	0.054*	0.074	0.074	0.053*
100	8	2	2	2	2	0.049*	0.049*	0.049*	0.049*	0.049*	0.049*
100	48	12	12	12	12	0.037	0.050*	0.050*	0.050*	0.050*	0.050*
100	88	2	2	42	42	0.147	0.053*	0.050*	0.050*	0.051*	0.048*
100	88	2	42	2	42	0.293	0.055*	0.050*	0.055*	0.055*	0.050*
100	88	2	42	42	2	0.270	0.056*	0.052*	0.053*	0.056*	0.052*
100	88	42	2	2	42	0.275	0.054*	0.053*	0.054*	0.051*	0.051*
100	88	42	2	42	2	0.061	0.056*	0.052*	0.055*	0.055*	0.052*
100	96	24	24	24	24	0.059*	0.063	0.053*	0.063	0.063	0.053*
100	98	2	2	2	92	0.961	0.075	0.053*	0.061	0.067	0.048*
100	98	2	2	92	2	0.923	0.072	0.050*	0.058*	0.070	0.048*
100	98	2	32	32	32	0.046*	0.048*	0.048*	0.048*	0.048*	0.048*
100	98	32	2	32	32	0.467	0.071	0.061	0.070	0.064	0.053*

* Denotes the approximation in $[0.040, 0.060]$

N	p	q_1	q_2	q_3	q_4	α_C	α_{H1}	α_{H2}	α_{H3}	α_{H4}	α_{H5}
200	8	2	2	2	2	0.055*	0.055*	0.055*	0.055*	0.055*	0.055*
200	96	24	24	24	24	0.026	0.049*	0.049*	0.049*	0.049*	0.049*
200	144	36	36	36	36	0.003	0.048*	0.048*	0.048*	0.048*	0.047*
200	148	2	2	2	142	0.117	0.051*	0.051*	0.050*	0.051*	0.050*
200	148	2	2	142	2	0.104	0.052*	0.051*	0.051*	0.052*	0.051*
200	148	2	2	72	72	0.064	0.053*	0.052*	0.052*	0.052*	0.051*
200	148	2	72	2	72	0.124	0.053*	0.052*	0.053*	0.053*	0.052*
200	148	2	72	72	2	0.121	0.055*	0.053*	0.053*	0.055*	0.053*
200	148	72	2	2	72	0.118	0.049*	0.049*	0.049*	0.048*	0.048*
200	148	72	2	72	2	0.050*	0.053*	0.052*	0.053*	0.053*	0.052*
200	158	2	52	52	52	0.171	0.054*	0.053*	0.054*	0.054*	0.053*
200	158	52	2	52	52	0.041*	0.049*	0.048*	0.048*	0.048*	0.048*
200	188	2	2	92	92	0.722	0.057*	0.053*	0.053*	0.055*	0.051*
200	188	2	92	2	92	0.991	0.054*	0.049*	0.054*	0.054*	0.049*
200	188	2	92	92	2	0.992	0.057*	0.053*	0.054*	0.057*	0.053*
200	188	92	2	2	92	0.990	0.053*	0.052*	0.053*	0.050*	0.049*
200	188	92	2	92	2	0.301	0.052*	0.048*	0.051*	0.051*	0.048*
200	188	2	62	62	62	0.988	0.053*	0.048*	0.052*	0.053*	0.048*
200	188	62	2	62	62	0.422	0.054*	0.053*	0.054*	0.052*	0.051*
200	192	48	48	48	48	0.043*	0.057*	0.052*	0.057*	0.057*	0.052*
200	198	2	2	2	192	1.000	0.076	0.054*	0.062	0.068	0.047*
200	198	2	2	192	2	1.000	0.080	0.054*	0.063	0.077	0.052*

* Denotes the approximation in $[0.040, 0.060]$

Acknowledgments

The author thanks Prof. T. Sugiyama and Prof. Y. Fujikoshi of Chuo University for many useful comments and suggestions.

References

[1] Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis* (3rd ed.). John Wiley & Sons, New York.

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

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

[4] Fujikoshi, Y. (1982). A test for additional information in canonical correlation analysis. *Ann. Inst. Statist. Math.*, **34**, 523–530

[5] Fujikoshi, Y. , Himeno, T. and Wakaki, H. (2008). Asymptotic results in canonical discriminant analysis when the dimension is large compared to the sample size. *J. Stat. Plann. Inference*, (to appear).

[6] Hsu, P. L. (1941). On the limiting distribution of roots of the canonical correlations. *Biometrika*, **32**, 38–45.

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

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

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

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

[11] Raudys, S. and Young, D. M. (2004). Results in statistical discriminant analysis: A review of the former Soviet Union literature. *J. Multivariate Anal.*, **89**, 1–35.

[12] Siotani, M. , Hayakawa, T. and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*. American Sciences Press, Columbus, Ohio.

[13] Sugiura, N. and Fujikoshi, Y. (1969). Asymptotic expansions of the non-null distributions of the likelihood ratio criteria for multivariate linear hypothesis and independence. *Ann. Math. Statist.*, **40**, 942–952.

[14] Tonda, T. and Fujikoshi, Y. (2004). Asymptotic expansion of the null distribution of LR statistic for multivariate linear hypothesis when the dimension is large. *Comm. Statist. Theory Methods*, **33**, 1205–1220.

[15] Wakaki, H. (2006). Edgeworth expansion of Wilks' lambda statistic. *J. Multivariate Anal.*, **97**, 1958–1964.