Journal of Multivariate Analysis 100 (2009) 888–901

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

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

ARTICLE INFO

Article history:
Received 28 November 2007
Available online 16 September 2008

AMS 2000 subject classifications:
primary 62H20
secondary 62H15

Keywords:
Asymptotic expansion
Additional information
Canonical correlation analysis
Tests for dimensionality
High-dimensional framework

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.

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 $\mu = (\mu_u', \mu_v')'$ 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 $p_1 \le p_2$. Let $\rho_1 \ge \cdots \ge \rho_{p_1} \ge 0$ be the possible nonzero population canonical correlations between $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$. Note that $\rho_1^2 \ge \cdots \ge \rho_{p_1}^2 \ge 0$ are the characteristic roots of $\Sigma_{uu}^{-1}\Sigma_{uv}\Sigma_{vv}^{-1}\Sigma_{vu}$. The coefficient vectors α_{ui} and α_{vi} of the ith canonical variables are defined as the solutions of

$$\Sigma_{uv}\Sigma_{vv}^{-1}\Sigma_{vu}\alpha_{ui} = \rho_i^2\Sigma_{uu}\alpha_{ui}, \qquad \alpha_{ui}'\Sigma_{uu}\alpha_{uj} = \delta_{ij},$$
$$\Sigma_{vu}\Sigma_{uu}^{-1}\Sigma_{uv}\alpha_{vi} = \rho_i^2\Sigma_{vv}\alpha_{vi}, \qquad \alpha_{vi}'\Sigma_{vv}\alpha_{vj} = \delta_{ij},$$

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

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.,

$$H_{\mathrm{dim}} : \rho_k > \rho_{k+1} = \cdots = \rho_{p_1} = 0.$$

E-mail address: s15002@gug.math.chuo-u.ac.jp.

0047-259X/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2008.09.005

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 $\boldsymbol{x}$, S is partitioned as

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

The following test statistics have been considered (e.g., see Siotani, Hayakawa and Fujikoshi [11]):

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

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,

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

some asymptotic results have been obtained (e.g., Anderson [1] and Siotani et al. [11]). 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

$$\text{A1:} \quad p_1; \text{fixed}, \quad p_2 \to \infty, \quad n \to \infty, \quad p_2/n \to c \in (0, 1).$$

Here, we note that $m = n - p_2 \to \infty$ is implied by A1. In our asymptotic framework A1, we assume $n \geq p = p_1 + p_2$. So, we can use the properties of Sugiura and Fujikoshi [12], which are given in the Appendix.

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 $\alpha_{ui}, \alpha_{vi}, \mu_u, \mu_v, \Sigma$ comfortably:

$$\begin{pmatrix} \alpha_{ui} \\ \alpha_{vi} \end{pmatrix} = \begin{pmatrix} \alpha_{1i} \\ \alpha_{2i} \\ \alpha_{3i} \\ \alpha_{4i} \end{pmatrix}, \quad \begin{pmatrix} \mu_u \\ \mu_v \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \Sigma_{34} \\ \Sigma_{41} & \Sigma_{42} & \Sigma_{43} & \Sigma_{44} \end{pmatrix}.$$

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 $\boldsymbol{x}_3$, or of the redundancy of $\boldsymbol{x}_2$ and $\boldsymbol{x}_4$, is formulated as follows:

$$H_{\text{add}} : \alpha_{2i} = \mathbf{0}, \quad \alpha_{4i} = \mathbf{0}, \qquad (i = 1, \ldots, k).$$

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, S is partitioned 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_{add}, we consider the statistic (Fujikoshi [4]) defined by

$$T = \begin{vmatrix} S_{22\cdot13} & S_{24\cdot13} \\ S_{42\cdot13} & S_{44\cdot13} \end{vmatrix} / \{|S_{22\cdot1}||S_{44\cdot3}|\}, \tag{2}$$

which is a likelihood ratio statistic. Here, $S_{22\cdot1} = S_{22} - S_{21}S_{11}^{-1}S_{12}$, $S_{22\cdot13} = S_{22} - S_{2(13)}S_{(13)(13)}^{-1}S_{(13)2}$, $S_{2(13)} = (S_{21}, S_{23})$, etc.

Fujikoshi [4] derived an asymptotic expansion for the distribution of T under A0. The approximation can be written as

$$P(-m \log T \leq x) = G_f(x) + \frac{\beta}{m^2}\{G_{f+4}(x) - G_f(x)\} + O(m^{-2}), \tag{3}$$

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

$$m = n - \frac{1}{2}(p + 1) - \frac{1}{2}\frac{q_1 q_3 (p - r)}{f},$$

$$\beta = \frac{1}{48}\left[\left\{q_1^2 + q_2^2 + q_3^2 + q_4^2 - 5\right\} f + 2q_1^2 q_2 p_2 + 2p_1 q_3^2 q_4 + 2q_2 q_4 \left\{q_1 q_2 + q_3 q_4 - 3q_1 q_3\right\} - 3(q_1 q_3)^2 (p - r)^2 / f\right].$$

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 statistic under a high-dimensional framework such that

$$\text{A2:} \quad p = p_1 + p_2 \to \infty, \quad n \to \infty, \quad p/n \to c \in (0, 1).$$

Here, we note that $m = n - p \to \infty$ is implied by A2. 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 [5] derived asymptotic distributions of test statistics for dimensionality in canonical discriminant analysis under A1. Tonda and Fujikoshi [13] derived an asymptotic expansion of the distribution of Wilks' lambda statistic Λ in the MANOVA model under A1. Wakaki [14] derived similar results for Λ in the MANOVA model 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 [2], Johnstone [6], Ledoit and Wolf [9], and Raudys and Young [10], etc.

2. Distributions of tests for dimensionality

In this section we consider the distribution of the three test statistics (1) under framework A1. When we consider the distributions of the statistics in (1), without loss of generality we may assume

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

since the statistics are expressed as functions of the characteristic roots of $S_{uu}^{-1} S_{uv} S_{vv}^{-1} S_{vu}$. Let $A = nS$. Corresponding to a partition of $\boldsymbol{x}$, A is partitioned as

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

Let

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

These are the characteristic roots of $A_{uu\cdot v}^{-1} A_{uv} A_{vv}^{-1} A_{vu}$.

Let U and V be the matrices defined by

$$U = \frac{1}{\sqrt{p_2}} \left\{ A_{uv} A_{vv}^{-1} A_{vu} - (p_2 I_{p_1} + n\Gamma^2) \right\} \quad \text{and} \quad V = \frac{1}{\sqrt{m}} (A_{uu\cdot v} - m I_{p_1}), \tag{4}$$

respectively. Then we obtain the following theorem.

Theorem 1. *Under assumption A1, each of the elements of U and V is asymptotically and independently distributed as a normal distribution, more precisely*

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

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

The proof of Theorem 1 is given in the Appendix. From Theorem 1 we can obtain the asymptotic distributions of three test statistics in (1).

2.1. Null distributions

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

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

Consider the transformed test statistics of LR, LH and BNP in (1) defined by

$$\begin{aligned} T_{LR} &= \sqrt{p_2}\left(1 + \frac{m}{p_2}\right)\left\{\log \prod_{j=k+1}^{p_1} (1 + \ell_j^2) - (p_1 - k)\log\left(1 + \frac{p_2}{m}\right)\right\}, \\ T_{LH} &= \sqrt{p_2}\left\{\frac{m}{p_2}\sum_{j=k+1}^{p_1} \ell_j^2 - (p_1 - k)\right\}, \\ T_{BNP} &= \sqrt{p_2}\left(1 + \frac{m}{p_2}\right)\left\{\left(1 + \frac{m}{p_2}\right)\sum_{j=k+1}^{p_1} \frac{\ell_j^2}{1 + \ell_j^2} - (p_1 - k)\right\}. \end{aligned} \tag{5}$$

Then T_{LR}, T_{LH} and T_{BNP} can be expanded (see the Appendix) as

$$T_G = \mathrm{tr}\left(U_{22} - \sqrt{\frac{p_2}{m}} V_{22}\right) + O^*_{1/2}, \tag{6}$$

where $G = LR, LH, BNP$,

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

U_{12} and V_{12} are $k \times (p_1 - k)$ matrices. Here, the notation O^*_i denotes a term of the ith order with respect to $(n^{-1}, p_2^{-1}, m^{-1})$. From Theorem 1, 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. *Under assumption A1 and A1.1,*

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

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:

$$K_{\mathrm{dim}} : \rho_b > \rho_{b+1} = \cdots = \rho_{p_1} = 0, \quad k < b \le p_1.$$

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

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

This is equivalent to $\gamma_1^2 > \cdots > \gamma_b^2 > \gamma_{b+1}^2 = \cdots = \gamma_{p_1}^2 = 0$. Let

$$\begin{aligned}
T^*_{LR} &= \sqrt{p_2}\left(1 + \frac{m}{p_2}\right)\left\{\log \prod_{j=k+1}^{p_1}(1 + \ell_j^2) - \log \prod_{j=k+1}^{p_1}\left(1 + \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)\right)\right\}, \\
T^*_{LH} &= \sqrt{p_2}\left\{\frac{m}{p_2}\sum_{j=k+1}^{p_1}\ell_j^2 - \frac{m}{p_2}\sum_{j=k+1}^{p_1}\frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)\right\}, \\
T^*_{BNP} &= \sqrt{p_2}\left(1 + \frac{m}{p_2}\right)\left\{\left(1 + \frac{m}{p_2}\right)\sum_{j=k+1}^{p_1}\frac{\ell_j^2}{1 + \ell_j^2} - \left(1 + \frac{m}{p_2}\right)\sum_{j=k+1}^{p_1}\frac{\frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)}{1 + \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)}\right\}.
\end{aligned} \tag{7}$$

Then T_{LR}, T_{LH} and T_{BNP} can be expanded (see the Appendix) as

$$T^*_G = \sum_{j=k+1}^{b} d_j^c\left(u_{jj} - \sqrt{\frac{p_2}{m}}\left(1 + \frac{n}{p_2}\gamma_j^2\right)v_{jj}\right) + \mathrm{tr}\left(\tilde{U}_{22} - \sqrt{\frac{p_2}{m}}\tilde{V}_{22}\right) + O^*_{1/2}, \tag{8}$$

where

$$d_j = \frac{1 + \frac{p_2}{m}}{1 + \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)}, \quad U = \begin{pmatrix} \tilde{U}_{11} & \tilde{U}_{12} \\ \tilde{U}_{21} & \tilde{U}_{22} \end{pmatrix}, \quad V = \begin{pmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\ \tilde{V}_{21} & \tilde{V}_{22} \end{pmatrix},$$

$\tilde{U}_{12}$ and $\tilde{V}_{12}$ are $b \times (p_1 - b)$ matrices. 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 1 and 2, we obtain the following theorem.

Theorem 3. *Let T_G^* be the transformed test statistics defined by (7), where $G = LR, LH, BNP$. Then, under assumption A1 and A1.2,*

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

where

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

2.3. Asymptotic power

On the basis of the asymptotic distributions of the three statistics in Theorem 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 $100\alpha\%$ points of the standard normal distribution.

Using Theorem 3, the asymptotic power with a level of significance α is expressed under A1 as

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

where Φ is the distribution function of the standard normal distribution. Under assumption A1,

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

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 (2). According to Theorem 2 in [4], T under H_{add} is expressed as a product of two independent variables, i.e.,

$$T = T_1 \times T_2, \quad T_1 \sim \Lambda(p_2, q_2, n - p_1), \quad T_2 \sim \Lambda(q_4, q_1, n - r), \tag{9}$$

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.

Let Λ be a statistic that is distributed as $\Lambda(p, q, n)$. Tonda and Fujikoshi [13] 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

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

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

$$\begin{aligned}
&\kappa_F^{(1)} = \frac{1}{\sqrt{p}}\tau_1, \qquad \kappa_F^{(3)} = \frac{1}{\sqrt{p}}\tau_3, \qquad \kappa_F^{(2)} = \frac{1}{p}\left(\tau_2 + \frac{\tau_1^2 - \tau_{(11)}}{2}\right),\\
&\kappa_F^{(4)} = \frac{1}{p}\left(\tau_4 + \tau_1\tau_3 - \tau_{(13)}\right), \qquad \kappa_F^{(6)} = \frac{1}{p}\left(\tau_6 + \frac{\tau_3^2 - \tau_{(33)}}{2}\right),\\
&\tau_i = \sum_{k=1}^{q}\omega_k^i\tau_{ik}, \quad \tau_{(ij)} = \sum_{k=1}^{q}\omega_k^{i+j}\tau_{ik}\tau_{jk}.
\end{aligned} \tag{10}$$

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

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

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

On the other hand, Wakaki [14] derived an asymptotic expansion formula for 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 also 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

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

Then, the characteristic function of T_W can be expanded as

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

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

$$\kappa_W^{(3)} = \tau^{(3)}/\left(\tau^{(2)}\right)^{3/2}, \qquad \kappa_W^{(4)} = \tau^{(4)}/\left(\tau^{(2)}\right)^2, \qquad \kappa_W^{(6)} = \left(\kappa_W^{(3)}\right)^2. \tag{11}$$

Using these results, we can obtain the asymptotic distribution of T in (2) 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 $p_1 \le p_2$, the condition can be realized as one of the following 12 cases.

	q_1	q_2	q_3	q_4		q_1	q_2	q_3	q_4
(i)	f	f	f	∞	(vii)	∞	f	f	∞
(ii)	f	f	∞	f	(viii)	∞	f	∞	f
(iii)	f	f	∞	∞	(ix)	∞	f	∞	∞
(iv)	f	∞	f	∞	(x)	∞	∞	f	∞
(v)	f	∞	∞	f	(xi)	∞	∞	∞	f
(vi)	f	∞	∞	∞	(xii)	∞	∞	∞	∞

Here f and ∞ in the q_i column denote "q_i is fixed" and "$q_i \to \infty$", respectively. We can obtain an asymptotic expansion of the distribution of T in all of the cases except (ii). For case (ii), we apply the approximations in the other cases. On the basis of a numerical simulation, we shall see that our approximations even in situation (ii) are good.

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 [13,14], respectively, depending on situations (i)–(xii). Let T_H be the standardization of T defined by

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

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). Then we obtain the following theorem.

Theorem 4. *Let T_G be the standardization of T defined by (2). Then, the null distribution of T_G can be expanded as*

$$P(T_G \le x) = \Phi(x) - \phi(x)\left[a_1(x) + a_2(x)\right] + o(p^{-1}), \tag{12}$$

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

$$a_1(x) = \kappa^{(1)} + \kappa^{(3)} h_2(x), \qquad a_2(x) = \kappa^{(2)} h_1(x) + \kappa^{(4)} h_3(x) + \kappa^{(6)} h_5(x). \tag{13}$$

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 the $\kappa^{(\alpha)}$ s are given by

$$\kappa^{(1)} = \tilde{\kappa}_1^{(1)} + \tilde{\kappa}_2^{(1)}, \qquad \kappa^{(3)} = \tilde{\kappa}_1^{(3)} + \tilde{\kappa}_2^{(3)}, \qquad \kappa^{(2)} = \tilde{\kappa}_1^{(2)} + \tilde{\kappa}_2^{(2)} + \tilde{\kappa}_1^{(1)}\tilde{\kappa}_2^{(1)},$$
$$\kappa^{(4)} = \tilde{\kappa}_1^{(4)} + \tilde{\kappa}_2^{(4)} + \tilde{\kappa}_1^{(1)}\tilde{\kappa}_2^{(3)} + \tilde{\kappa}_1^{(3)}\tilde{\kappa}_2^{(1)}, \qquad \kappa^{(6)} = \tilde{\kappa}_1^{(6)} + \tilde{\kappa}_2^{(6)} + \tilde{\kappa}_1^{(3)}\tilde{\kappa}_2^{(3)},$$

where $\tilde{\kappa}_j^{(\alpha)} = w_j^{\alpha}\kappa_j^{(\alpha)}$ and $w_j = d_j/d$. Furthermore the $\kappa_j^{(\alpha)}$ s are given by (10) and (11), respectively. For a detailed definition, see Table A.1 in the Appendix.

The proof of Theorem 4 is given in the Appendix.

Using the coefficients $a_j(x)$ of the asymptotic expansion (12), we can obtain the Cornish–Fisher expansion. Let x and $t_G(x)$ denote the percentage point of the limiting distribution of T_G and the corresponding percentage points of T_G, respectively, that is

$$P(T_G \le t_G(x)) = \Phi(x), \qquad G = H1\text{–}H5.$$

Then from (12), $t_G(x)$ can be expanded as

$$\begin{aligned} t_G(x) &= x + a_1(x) + \left\{a_1(x)a_1'(x) - \frac{1}{2}x a_1(x)^2 + a_2(x)\right\} + o(p^{-1}) \\ &= \tilde{t}_G(x) + o(p^{-1}). \end{aligned} \tag{14}$$

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

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

are asymptotically distributed as the χ^2-distribution with $(p_1-k)(p_2-k)$ degrees of freedom (e.g., see Siotani et al. [11]). Under the high-dimensional framework the three statistics T_G/σ_G are distributed asymptotically (see Theorem 2) as $N(0, 1)$, where $\sigma_G = \sqrt{2(p_1-k)(1+p_2/m)}$ and $G = LR, LH, BNP$. To facilitate understanding, 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 the 100(1–α)% points 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 $\mathscr{P}$ were chosen as follows:

(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),

$(p_1, \mathscr{P})$: (3, diag(0.9, 0.6, 0.0)), $(p_1, \mathscr{P})$: (4, diag(0.9, 0.6, 0.0, 0.0)),

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 10. However, the chi-square type approximation is poor when p is greater than 10. 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 approximations $\alpha_{LR\cdot C}$, $\alpha_{LH\cdot C}$ and $\alpha_{BNP\cdot C}$ perform 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.

4.2. Test for additional information

It is easy to obtain Cornish–Fisher expansion of the large sample approximation (3). In fact, the expansion (3) can be written as

$$P(-m\log T \le x) = G_f(x) + g_f(x)\frac{1}{m^2}\tilde{p}_1(x) + o(m^{-2}), \tag{15}$$

where $g_f(x)$ is the density function of the chi-square variable with f degrees of freedom and the 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

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

Then the Cornish–Fisher expansion can be written in the same way as in (14), that is,

$$\begin{aligned}\tilde{t}(x) &= x + \frac{1}{m^2}\tilde{p}_1(x) + o(m^{-2})\\ &= \tilde{t}_C + O(m^{-2}).\end{aligned} \tag{16}$$

Table 1
Upper 5% points of *LR*, *LH*, *BNP* for dimensionality

N	p	p_2	Simu_{LR}	$t_{LR\cdot H}$	Simu_{LH}	$t_{LH\cdot H}$	Simu_{BNP}	$t_{BNP\cdot H}$	t_C
$p_1 = 3, k = 2, \mathcal{P} = \mathrm{diag}(0.9, 0.6, 0.0)$									
50	10	7	0.24[b]	0.29[a]	0.28[b]	0.32[a]	0.23[b]	0.28[a]	0.23
50	20	17	0.60[b]	0.65[a]	0.83	0.88[a]	1.30	1.69	0.51
50	30	27	1.04	1.11[a]	1.83	1.91[a]	3.63	5.12	0.77
50	40	37	1.71	1.79[a]	4.67	4.65[a]	7.69	11.44	1.02
100	10	7	0.12[b]	0.14[a]	0.12[b]	0.15[a]	0.09[b]	0.10[a]	0.11
100	15	12	0.20[b]	0.22[a]	0.22[b]	0.24[a]	0.22[b]	0.25[a]	0.18
100	20	17	0.28[b]	0.29[a]	0.32[b]	0.33[a]	0.43	0.48[a]	0.25
100	50	47	0.82	0.84[a]	1.28	1.28[a]	4.10	4.84	0.62
100	70	67	1.33	1.38[a]	2.82	2.86[a]	8.82	11.97	0.86
100	90	87	2.35	2.41[a]	9.84	9.73[a]	15.61	23.85	1.09
$p_1 = 4, k = 2, \mathcal{P} = \mathrm{diag}(0.9, 0.6, 0.0, 0.0)$									
50	10	7	0.24[b]	0.29[a]	0.28[b]	0.32[a]	0.23[b]	0.28[a]	0.23
50	20	17	0.60[b]	0.65[a]	0.83	0.88[a]	1.30	1.69	0.51
50	30	27	1.04	1.11[a]	1.83	1.91[a]	3.63	5.12	0.77
50	40	37	1.71	1.79[a]	4.67	4.65[a]	7.69	11.44	1.02
100	10	7	0.12[b]	0.14[a]	0.12[b]	0.15[a]	0.09[b]	0.10[a]	0.11
100	15	12	0.20[b]	0.22[a]	0.22[b]	0.24[a]	0.22[b]	0.25[a]	0.18
100	20	17	0.28[b]	0.29[a]	0.32[b]	0.33[a]	0.43	0.48[a]	0.25
100	50	47	0.82	0.84[a]	1.28	1.28[a]	4.10	4.84	0.62
100	70	67	1.33	1.38[a]	2.82	2.86[a]	8.82	11.97	0.86
100	90	87	2.35	2.41[a]	9.84	9.73[a]	15.61	23.85	1.09

[a] Denotes $|\mathrm{Simu}_G - t_{G\cdot H}| \le 10^{-1}$, where $G = LR, LH, BNP$.
[b] Denotes $|\mathrm{Simu}_G - t_C| \le 10^{-1}$, where $G = LR, LH, BNP$.

Table 2
The corresponding actual test sizes of *LR*, *LH*, *BNP* for dimensionality.

N	p	p_2	$\alpha_{LR\cdot H}$	$\alpha_{LR\cdot C}$	$\alpha_{LH\cdot H}$	$\alpha_{LH\cdot C}$	$\alpha_{BNP\cdot H}$	$\alpha_{BNP\cdot C}$
$p_1 = 3, k = 2, \mathcal{P} = \mathrm{diag}(0.9, 0.6, 0.0)$								
50	10	7	0.021	0.070	0.023	0.102	0.017	0.041[a]
50	20	17	0.026	0.155	0.035	0.351	0.017	0.016
50	30	27	0.024	0.406	0.040[a]	0.787	0.012	0.001
50	40	37	0.025	0.837	0.051[a]	0.991	0.004	0.000
100	10	7	0.023	0.058[a]	0.025	0.074	0.021	0.046[a]
100	15	12	0.032	0.074	0.035	0.111	0.028	0.043[a]
100	20	17	0.038	0.097	0.042[a]	0.173	0.032	0.041[a]
100	50	47	0.039	0.465	0.052[a]	0.845	0.025	0.006
100	70	67	0.026	0.895	0.044[a]	0.998	0.011	0.000
100	90	87	0.024	1.000	0.054[a]	1.000	0.002	0.000
$p_1 = 4, k = 2, \mathcal{P} = \mathrm{diag}(0.9, 0.6, 0.0, 0.0)$								
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[a]	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[a]
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[a]
100	50	47	0.028	0.684	0.051[a]	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

[a] Denotes the approximation in [0.040, 0.060].

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

Table 3 gives the upper 5% 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 approximations $t_{H1} \sim t_{H5}$, $\alpha_{H1} \sim \alpha_{H5}$ perform 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 approximations t_{H5} and α_{H5} are the best of these approximation for all cases.

Table 3
Upper 5% points of test statistic for additional information

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

[a] Denotes the approximation in Simu $\pm 10^{-1}$.

Table 4
The corresponding actual test sizes of the test statistic for additional information

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[a]	0.052[a]	0.051[a]	0.052[a]	0.052[a]	0.051[a]
50	48	2	2	2	42	0.502	0.072	0.052[a]	0.059[a]	0.063	0.047[a]
50	48	2	2	42	2	0.472	0.075	0.052[a]	0.062	0.073	0.051[a]
50	48	2	2	22	22	0.281	0.080	0.057[a]	0.068	0.074	0.052[a]
50	48	2	22	2	22	0.420	0.065	0.048[a]	0.065	0.065	0.048[a]
50	48	2	22	22	2	0.389	0.071	0.051[a]	0.063	0.071	0.051[a]
50	48	22	2	2	22	0.383	0.066	0.055[a]	0.066	0.057[a]	0.047[a]
50	48	22	2	22	2	0.097	0.067	0.049[a]	0.064	0.064	0.046[a]
50	48	12	12	12	12	0.070	0.074	0.054[a]	0.074	0.074	0.053[a]
100	8	2	2	2	2	0.049[a]	0.049[a]	0.049[a]	0.049[a]	0.049[a]	0.049[a]
100	48	12	12	12	12	0.037	0.050[a]	0.050[a]	0.050[a]	0.050[a]	0.050[a]
100	88	2	2	42	42	0.147	0.053[a]	0.050[a]	0.050[a]	0.051[a]	0.048[a]
100	88	2	42	2	42	0.293	0.055[a]	0.050[a]	0.055[a]	0.055[a]	0.050[a]
100	88	2	42	42	2	0.270	0.056[a]	0.052[a]	0.053[a]	0.056[a]	0.052[a]
100	88	42	2	2	42	0.275	0.054[a]	0.053[a]	0.054[a]	0.051[a]	0.051[a]
100	88	42	2	42	2	0.061	0.056[a]	0.052[a]	0.055[a]	0.055[a]	0.052[a]
100	96	24	24	24	24	0.059[a]	0.063	0.053[a]	0.063	0.063	0.053[a]
100	98	2	2	2	92	0.961	0.075	0.053[a]	0.061	0.067	0.048[a]
100	98	2	2	92	2	0.923	0.072	0.050[a]	0.058[a]	0.070	0.048[a]
100	98	2	32	32	32	0.046[a]	0.048[a]	0.048[a]	0.048[a]	0.048[a]	0.048[a]
100	98	32	2	32	32	0.467	0.071	0.061	0.070	0.064	0.053[a]

[a] Denotes the approximation in [0.040, 0.060].

5. Concluding remarks and discussion

In this paper we obtained asymptotic approximations of test statistics for dimensionality and additional information in canonical correlation analysis under high-dimensional frameworks A1 and A2. By means of simulation experiments (Tables 1–4), it was shown that the high-dimensional approximations are better than the large sample approximations for a wide range of (p, q, n) with large p. The high-dimensional asymptotic approximations are useful for the distributions of test statistics for dimensionality and additional information in such situations.

However, it is pointed out that the high-dimensional approximations for tests of dimensionality worsen when q is large on test statistics for dimensionality. An approach to overcoming the fault is to derive asymptotic distributions of test statistics for dimensionality in canonical correlations under the following high-dimensional framework:

$$q \to \infty, \quad p \to \infty, \quad n \to \infty, \quad c_1 = p/n,\ c_2 = q/n \to c_{01}, c_{02} \in [0, 1).$$

This problem and the extension to a class of elliptical distributions, etc., are left as future research. In addition, we do not discuss whether the estimated dimension based on a sequential test procedure with high-dimensional approximation is convergent in probability to the true dimension, though the problem is important.

Acknowledgments

The author thanks Prof. Y. Fujikoshi, Prof. T. Sugiyama, Chuo University, and the two reviewers for many useful comments and suggestions.

Appendix. Proofs of the theorem and derivation of the asymptotic expansions

Proof of Theorem 1

For our derivation, we use the following properties (see Sugiura and Fujikoshi [12]):

(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.

When we consider the distribution of a function of the canonical correlations $r_1 > \cdots > r_q$, without loss of generality, we may assume that:

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

The characteristic function of U in (4) can be expressed as

$$\begin{aligned} C_U(T) &= E[\exp(\mathrm{itr}TU)] \\ &= E_W\left[E[\exp(\mathrm{itr}TU)|W]\right], \end{aligned}$$

where T is a real symmetric matrix whose (i, j) element is given by $(1 + \delta_{ij})t_{ij}/2$ for every real value t_{ij}. 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

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

where $G = \frac{1}{\sqrt{n}}(W - nI)$ and o_i^* denotes a term that tends to 0 under a high-dimensional framework A1. Therefore,

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

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 $C_{V,U}(T_1, T_2)$ for the joint characteristic function of V and U as follows:

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

Therefore we can obtain Theorem 1.

Derivation of (6) *and* (8)

For our derivation of (6) and (8), we consider a perturbation expansion of

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

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

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

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

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

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

Derivation of the asymptotic distribution of T_G in (6). Note that $\ell_1^2, \ldots, \ell_{p_1}^2$ are the characteristic roots of $Q = A_{uu\cdot v}^{-\frac{1}{2}} A_{uv} A_{vv}^{-1} A_{vu} A_{uu\cdot v}^{-\frac{1}{2}}$. Using the fact that Q has a perturbation expansion as in (A.2), it can be seen (see Lawley [7,8] and Fujikoshi [3]) that the last $p_1 - k$ characteristic roots $\ell_{k+1}^2, \ldots, \ell_{p_1}^2$ are the characteristic roots of

$$D = \frac{p_2}{m}I_{p_1-k} + \frac{1}{\sqrt{m}}\left(\sqrt{\frac{p_2}{m}}U_{22} - \frac{p_2}{m}V_{22}\right) + O_1^*, \tag{A.3}$$

under A1.2. Here U_{22} and V_{22} are the last $(p_1-k)\times(p_1-k)$ submatrices of U and V, respectively. From (A.3) we can expand T_{LR}, T_{LH}, and T_{BNP} as follows:

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

Therefore we can combine the above three expressions as (6).

Derivation of the asymptotic distribution of T_G^* in (8). Using a perturbation expansion of Q in (A.2) and a general result (e.g., see Siotani, Hayakawa and Fujikoshi [11]) for a perturbation expansion of its characteristic root, we can obtain

$$\sqrt{m}\left\{\ell_j^2 - \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)\right\} = \sqrt{\frac{p_2}{m}}u_{jj} - \frac{p_2}{m}\left(1 + \frac{n}{p_2}\gamma_j^2\right)v_{jj} + O_{1/2}^*, \quad j = k+1, \ldots, b. \tag{A.4}$$

Further, from A1.2 the last $p_1 - b$ characteristic roots $\ell_{b+1}^2, \ldots, \ell_{p_1}^2$ are the characteristic roots of

$$\tilde{Q} = \frac{p_2}{m} I_{p_1-b} + \frac{1}{\sqrt{m}} \left(\sqrt{\frac{p_2}{m}} \tilde{U}_{22} - \frac{p_2}{m} \tilde{V}_{22} \right) + O_{1/2}^*, \tag{A.5}$$

where

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

and $\tilde{U}_{12}$ and $\tilde{V}_{12}$ are $b \times (p_1 - b)$ matrices. Using (A.4) and (A.5) we can express T_{LR}^*, T_{LH}^* and T_{BNP}^* as follows:

$$\begin{aligned} T_{LR}^* &= \sum_{j=k+1}^{p_1} \left\{ \frac{1+\frac{p_2}{m}}{1+\frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right)} \sqrt{\frac{m}{p_2}} \left(\sqrt{\frac{p_2}{m}} u_{jj} - \frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right) v_{jj} \right) \right\} \\ &\quad + \mathrm{tr}\left(\tilde{U}_{22} - \sqrt{\frac{p_2}{m}} \tilde{V}_{22} \right) + O_{1/2}^*, \\ T_{LH}^* &= \sum_{j=k+1}^{b} \left\{ \sqrt{\frac{m}{p_2}} \left(\sqrt{\frac{p_2}{m}} u_{jj} - \frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right) v_{jj} \right) \right\} \\ &\quad + \mathrm{tr}\left(\tilde{U}_{22} - \sqrt{\frac{p_2}{m}} \tilde{V}_{22} \right) + O_{1/2}^*, \\ T_{BNP}^* &= \sum_{j=k+1}^{p_1} \left\{ \frac{\left(1+\frac{p_2}{m}\right)^2}{\left(1+\frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right)\right)^2} \sqrt{\frac{m}{p_2}} \left(\sqrt{\frac{p_2}{m}} u_{jj} - \frac{p_2}{m}\left(1+\frac{n}{p_2}\gamma_j^2\right) v_{jj} \right) \right\} \\ &\quad + \mathrm{tr}\left(\tilde{U}_{22} - \sqrt{\frac{p_2}{m}} \tilde{V}_{22} \right) + O_{1/2}^*. \end{aligned}$$

Therefore we can combine the above three expressions as (8).

Proof of Theorem 4

To prove Theorem 4, we consider the characteristic function $C_H(t)$. 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

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

where

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

Here the $\tilde{\kappa}_j^{(\alpha)}$ are defined by $\tilde{\kappa}_j^{(\alpha)} = w_j^{\alpha} \kappa_j^{(\alpha)}$, where the $\kappa_j^{(\alpha)}$ are given by (10) and (11), respectively.

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.

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

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

By inverting the characteristic function of T_G we can obtain Theorem 4.

Table A.1
The definition of m_i, d_i and $\kappa_i^{(\alpha)}$ in test statistics for additional information

	$m_1, d_1, \kappa_1^{(\alpha)}$	$m_2, d_2, \kappa_2^{(\alpha)}$
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}	$W(p_2, q_2, n-p_1)$	$W(q_4, q_1, n-r)$

References

[1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, 3rd edition, John Wiley & Sons, New York, 2003.
[2] Z.D. Bai, Methodologies in spectral analysis of large dimensional random matrices, a review, Statist. Sinica 9 (1999) 611–677.
[3] Y. Fujikoshi, Asymptotic expansions for the distributions of some multivariate tests, in: P.R. Krishnaiah (Ed.), Multivariate Analysis—IV, North-Holland Publishing Company, Amsterdam, 1977, pp. 55–71.
[4] Y. Fujikoshi, A test for additional information in canonical correlation analysis, Ann. Inst. Statist. Math. 34 (1982) 523–530.
[5] Y. Fujikoshi, T. Himeno, H. Wakaki, Asymptotic results in canonical discriminant analysis when the dimension is large compared to the sample size J. Statist. Plann. Inference (2008) (in press).
[6] I.M. Johnstone, On the distribution of the largest eigenvalue in principal component analysis, Ann. Statist. 29 (2001) 295–327.
[7] D.N. Lawley, Tests of significance for the latent roots of covariance and correlation matrices, Biometrika 43 (1956) 128–136.
[8] D.N. Lawley, Tests of significance in canonical analysis, Biometrika 46 (1959) 59–66.
[9] O. Ledoit, M. Wolf, Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size, Ann. Statist. 30 (2002) 1081–1102.
[10] S. Raudys, D.M. Young, Results in statistical discriminant analysis: A review of the former Soviet Union literature, J. Multivariate Anal. 89 (2004) 1–35.
[11] M. Siotani, T. Hayakawa, Y. Fujikoshi, Modern Multivariate Statistical Analysis: A Graduate Course and Handbook, American Sciences Press, Columbus, Ohio, 1985.
[12] N. Sugiura, Y. Fujikoshi, Asymptotic expansions of the non-null distributions of the likelihood ratio criteria for multivariate linear hypothesis and independence, Ann. Math. Statist. 40 (1969) 942–952.
[13] T. Tonda, Y. Fujikoshi, Asymptotic expansion of the null distribution of LR statistic for multivariate linear hypothesis when the dimension is large, Comm. Statist. Theory Methods 33 (2004) 1205–1220.
[14] H. Wakaki, Edgeworth expansion of Wilks' lambda statistic, J. Multivariate Anal. 97 (2006) 1958–1964.