# NON-ASYMPTOTIC ANALYSIS OF APPROXIMATIONS FOR LAWLEY–HOTELLING AND BARTLETT–NANDA–PILLAI STATISTICS IN HIGH-DIMENSIONAL SETTINGS

Alexander A. Lipatiev and Vladimir V. Ulyanov

(Received January 2020)

Abstract. We consider General Linear Model (GLM) that includes multivariate analysis of variance (MANOVA) and multiple linear regression as special cases. In practice, there are several widely used criteria for GLM: Wilks' lambda, Bartlett–Nanda–Pillai test, Lawley–Hotelling test and Roy maximum root test. Limiting distributions of statistics for the first three mentioned tests are known under different asymptotic settings. In the present paper we get the computable error bounds for normal approximation of Lawley–Hotelling and Bartlett–Nanda–Pillai statistics when dimensionality grows proportionally to sample size. This result enables us to get more precise calculations of the p-values in applications of multivariate analysis.

## 1. Introduction

In last years we encounter more and more problems in applications when dimension $p$ of observations is comparable with sample size $n$ or even exceeds it. Some examples of high-dimensional data include curve data, spectra, images, DNA microarrays, social networks and financial data.

There are many statistical procedures for data of relatively low dimensionality that have become classic. But it is often impossible to use an existing statistical procedure for high-dimensional case just by turning $p$ to infinity, because a limiting distribution of a test statistic becomes different (see Section 6.3.4 "High-Dimensional Approximations" in [1]).

In this paper we find the uniform computable error bounds for approximations of Lawley–Hotelling and Bartlett–Nanda–Pillai statistics by normal distribution in MANOVA model for high-dimensional setting, when the ratio $p/n$ of dimension of observations and sample size tends to constant from the interval $(0,1)$.

Section 2 contains formulation of Theorems 1 and 2 that are main results of the paper. Theorems 3 and 4 are essential for the proofs of the main results.

The article was prepared within the framework of the HSE University Basic Research Program.

2010 *Mathematics Subject Classification.* Primary 62E17; Secondary 62H10.

*Key words and phrases.* accuracy of approximation, MANOVA, computable error bounds, Lawley–Hotelling statistic, Bartlett–Nanda–Pillai statistic, high dimensional data.

They are also of independent interest. Section 4 contains proofs of the theorems that are based on lemmas from Section 3.

## 2. Main results

Multivariate analysis of variance (MANOVA) studies the following linear model:

$$X = Q\mathbb{B} + \mathcal{E},$$

where $X$ is a random $N \times p$ observation matrix, $Q$ is a non-random $N \times k$ design matrix, $\mathbb{B}$ is a non-random $k \times p$ matrix of regression coefficients and $\mathcal{E}$ is a $N \times p$ error matrix distributed as $N_{N\times p}\left(O, I_N \otimes \Sigma\right)$. Consider testing the linear hypothesis:

$$H_0 : C\mathbb{B} = O,$$

where $C$ is a known $q \times k$ matrix of rank $q$. Under a certain group of transformations, the testing problem is invariant, and invariant tests depend on the non-zero eigenvalues of $S_h S_e^{-1}$, where

$$S_h = \hat{\mathbb{B}}^T C^T \left( C \left(Q^T Q\right)^{-1} C^T \right)^{-1} C\hat{\mathbb{B}} \quad \text{and} \quad S_e = \left(X - Q\hat{\mathbb{B}}\right)^T \left(X - Q\hat{\mathbb{B}}\right),$$

with $\hat{\mathbb{B}} = \left(Q^T Q\right)^{-1} Q^T X$ (see [3], Theorem 10.2.1). Among invariant widely used tests, we consider Lawley–Hotelling statistic and Bartlett–Nanda–Pillai statistic:

$$\begin{aligned} T_0^2 &= n \operatorname{tr} S_h S_e^{-1}, \\ V_{BNP} &= (n+q) \operatorname{tr} S_h \left(S_h + S_e\right)^{-1}. \end{aligned}$$

Further we assume that hypothesis $H_0$ is fullfilled.

Anderson in [2], Muirhead in [3], Fujikoshi, Ulyanov and Shimizu in [4] and Lipatiev and Ulyanov in [5] considered large sample case under *A1 condition:*

$$\mathbf{A1} : p \text{ and } q \text{ fixed}, \ n \to \infty$$

In [4] and in [5] we derived non-asymptotic error bounds for approximations of distribution functions of Lawley–Hotelling and Bartlett–Nanda–Pillai statistics:

$$\begin{aligned} \mathbf{P}\{T_0^2 < x\} &= G_a(x) + \frac{a}{4n}\{(q-p-1)\, G_a(x) \\ &\quad -2q G_{a+2}(x) + (q+p+1)\, G_{a+4}(x)\} + O\left(n^{-2}\right), \\ \mathbf{P}\{V_{BNP} < x\} &= G_a(x) + \frac{3a}{4n}\{G_a(x) \\ &\quad -2G_{a+2}(x) + G_{a+4}(x)\} + O\left(n^{-2}\right), \end{aligned}$$

where $a = pq$ and $G_a$ is the distribution function of chi-square distribution with $a$ degrees of freedom. In [4] we give an upper bound for the remainder term of approximation for Lawley–Hotelling statistic and [5] contains an upper bound for the remainder term in the case of Bartlett–Nanda–Pillai statistic.

Wakaki, Fujikoshi and Ulyanov in [6] considered the case of large dimensionional data under *A2 condition:*

$$\mathbf{A2} : q \text{ fixed}, \; p \to \infty, \; n \to \infty, \; \frac{p}{n} \to c \in (0,1)$$

and derived the following approximation:

$$\begin{aligned} \mathbf{P}\Big(\frac{1}{\sigma}T_G < z\Big) &= \Phi(z) - \phi(z) \left[ \frac{1}{\sqrt{p}} \left\{ \frac{1}{\sigma} b_1 + \frac{1}{\sigma^3} b_3 \, H_2(z) \right\} \right. \\ & \left. + \frac{1}{p} \left\{ \frac{1}{\sigma^2} b_2 \, H_1(z) + \frac{1}{\sigma^4} b_4 \, H_3(z) + \frac{1}{\sigma^6} b_6 \, H_5(z) \right\} \right] \\ & + O\left( \frac{1}{p\sqrt{p}} \right), \end{aligned}$$

where $T_G$ is one of the following statistics:

$$\begin{aligned} T_{LH} &= \sqrt{p}\{\frac{m}{np} T_0^2 - q\}, \\ T_{BNP} &= \sqrt{p}\left(1 + \frac{p}{m}\right) \left\{ \frac{1}{p} V_{BNP} - q \right\}, \end{aligned}$$

$\Phi(z)$ and $\phi(z)$ are distribution function and probability density function of the standard normal law respectively; $m = n - p + q$; $r = p/m; \sigma = \sqrt{2q(1+r)}$; $b_i = b_i(r,q)$ are some functions of $r$ and $q$, different for $T_{LH}$ and $T_{BNP}$; $H_i(z)$ are the Hermite polynomials. Yet that result was asymptotic: the upper bounds for the remainder terms were not derived.

The main results of the present paper are the following theorems which give the upper bounds for accuracy of approximation for the Lawley–Hotelling and Bartlett–Nanda–Pillai statistics by normal law in case of large dimension, i.e. under condition $A2$ :

Theorem 1. *There exists $M_1 = M_1(r,q)$ such that for all $m > M_1$ one has*

$$\sup_z \left| \mathbf{P}\Big( \frac{T_{LH}}{\sqrt{2q\,(1+r)}} < z \Big) - \Phi(z) \right| \leqslant \frac{K_1(r,q) \ln m}{\sqrt{m}},$$

*where $K_1(r,q)$ is a constants depending on $r$ and $q$ only.*

Theorem 2. *There exists $M_2 = M_2(r,q)$ such that for all $m > M_2$ one has*

$$\sup_z \left| \mathbf{P}\Big( \frac{T_{BNP}}{\sqrt{2q\,(1+r)}} < z \Big) - \Phi(z) \right| \leqslant \frac{K_2(r,q) \ln m}{\sqrt{m}},$$

*where $K_2(r,q)$ is a constants depending on $r$ and $q$ only.*

Note that in Theorems 1 and 2 we have a logarithmic factor comparing with [6], but they are superior in terms of giving computable upper bound for accuracy of approximation. The proofs are also new.

The next two theorems play the key role in the proofs of the main results. They are also of independent interest.

Introduce additional notation. Let random matrices $B$ and $W$ be defined on the same probability space. Assume that $B$ and $W$ are independent and have the Wishart distributions $W_q\left(p, I_q\right)$ and $W_q\left(m, I_q\right)$ respectively, where $m = n-p-q$. Denote the normed versions of $B$ and $W$ by $U$ and $V$ respectively:

$$U = (B - pI_q)/\sqrt{p}, \quad V = (W - mI_q)/\sqrt{m}, \tag{1}$$

Theorem 3. *If* $\operatorname{tr} V^2 < m$ *then*

$$\begin{aligned} &\left|\sqrt{m}\left(\operatorname{tr} BW^{-1} - rq\right) - \left(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\right)\right| \\ \leqslant\ & \frac{\sqrt{r}\left|\operatorname{tr} UV\right| + \left(rq + \left|\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\right|/\sqrt{m}\right)\operatorname{tr} V^2}{\sqrt{m} - \operatorname{tr} V^2/\sqrt{m}}. \end{aligned} \tag{2}$$

Theorem 4. *If* $\operatorname{tr}\left(\sqrt{r}U + V\right)^2 < (r+1)\, m$ *then*

$$\begin{aligned} &\left|\sqrt{m}\,(r+1)\left((r+1)\operatorname{tr} B\,(B+W)^{-1} - rq\right) - \left(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\right)\right| \\ &\leqslant \frac{(r+1)\sqrt{r}\left(\left|\operatorname{tr} UV\right| + \sqrt{r}\operatorname{tr} U^2\right)}{(r+1)\sqrt{m} - \operatorname{tr}\left(\sqrt{r}U + V\right)^2/\sqrt{m}} \\ &\quad + \frac{\left(rq\,(r+1) + \left(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\right)/\sqrt{m}\right)\operatorname{tr}\left(\sqrt{r}U + V\right)^2}{(r+1)\sqrt{m} - \operatorname{tr}\left(\sqrt{r}U + V\right)^2/\sqrt{m}}. \end{aligned} \tag{3}$$

Note that the probabilities of complements of both events $\left\{\operatorname{tr} V^2 < m\right\}$ and $\left\{\operatorname{tr}\left(\sqrt{r}U + V\right)^2 < (r+1)\, m\right\}$, mentioned in Theorems 3 and 4, have order $O\left(1/\sqrt{m}\right)$. It follows from the results of Section 3 below.

## 3. Lemmas

This section contains auxiliary lemmas used in the proofs of theorems.

Let

$$\begin{aligned} Z_1 &= \operatorname{tr} UV, \\ Z_2 &= \operatorname{tr} V^2, \\ Z_3 &= \operatorname{tr} U - \sqrt{r}\operatorname{tr} V, \\ Z_4 &= \operatorname{tr} U^2, \end{aligned} \tag{4}$$

where random matrices $U$ and $V$ are defined in (1).

Put

$$B = B(q, r, m) = 4\left(q^2 + \sqrt{r}\right)(\sqrt{\ln m} + \sqrt{\ln p})^2. \tag{5}$$

We define the random events $A_{i,m}$ as $A_{i,m} = \{\omega : |Z_i(\omega)| \leqslant B\}$ for $i = 1, 2, 3, 4$ and positive integers $m$.

Let

$$Z_{LH} = \frac{\sqrt{r}\,|\operatorname{tr} UV| + (rq + |\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V|/\sqrt{m})\operatorname{tr} V^2}{\sqrt{m} - \operatorname{tr} V^2/\sqrt{m}},$$

$$Z_{BNP} = \frac{(r+1)\sqrt{r}\left(|\operatorname{tr} UV| + \sqrt{r}\operatorname{tr} U^2\right) + \left(rq\,(r+1) + \sqrt{r}\frac{\left|\operatorname{tr} U - \sqrt{r}\operatorname{tr} V\right|}{\sqrt{m}}\right) S_Z}{(r+1)\sqrt{m} - S_Z/\sqrt{m}},$$

where $S_Z = \left(r\operatorname{tr} U^2 + 2\sqrt{r}\,|\operatorname{tr} UV| + \operatorname{tr} V^2\right)$.

It is clear that there exist the positive integer numbers $M_1 = M_1(r,q)$ and $M_2 = M_2(r,q)$ such that if $m \geqslant M_1$ and $\omega \in \cap_{i=1}^{3} A_{i,m}$ then

$$\begin{aligned} Z_{LH}(\omega) &\leqslant \frac{\sqrt{r}\,B + (rq + \sqrt{r}\,B/\sqrt{m})\,B}{\sqrt{m} - B/\sqrt{m}} \\ &\leqslant 48\,(2\sqrt{r} + rq)\,(q^2 + \sqrt{r})\frac{(\ln m + \ln\sqrt{r})}{\sqrt{m}}. \end{aligned} \tag{6}$$

Moreover, if $m \geqslant M_2$ and $\omega \in \cap_{i=1}^{4} A_{i,m}$, then

$$\begin{aligned} Z_{BNP}(\omega) &\leqslant \frac{(r+1)\sqrt{r}(1+\sqrt{r})B + \left(rq(r+1) + \sqrt{r}\frac{B}{\sqrt{m}}\right) B(1+\sqrt{r})^2}{(r+1)\sqrt{m} - B(1+\sqrt{r})^2/\sqrt{m}} \\ &\leqslant 16\,\frac{(1+\sqrt{r})^2\,(2\sqrt{r} + r\,(rq+q+1))(q^2+\sqrt{r})}{r+1}\,\frac{(\ln m + \ln\sqrt{r})}{\sqrt{m}}. \end{aligned} \tag{7}$$

In the next two lemmas we estimate probabilities $\mathbf{P}(A_{i,m}^c)$ for $i = 1,2,3,4$.

LEMMA 1. *Let* $Z_1$ *and* $B$ *be defined in* (4) *and* (5) *respectively. Then*

$$\mathbf{P}(|Z_1| > B) \leqslant 6.45\,q^2\,\frac{1 + 1/\sqrt{r}}{\sqrt{m}}. \tag{8}$$

*Proof of Lemma 1.* Since $B$ has the Wishart distribution $W_q\,(p, I_q)$, we can represent $B$ as $X^T X$, where

$$X = \begin{pmatrix} X_{11} & \dots & X_{1q} \\ \vdots & \ddots & \vdots \\ X_{p1} & \dots & X_{pq} \end{pmatrix}$$

and $\{X_{ij}\}$ are independent standard normal variables.

Therefore, matrix $B$ can be written as:

$$B = \left\{\sum_{k=1}^{p} X_{ki}X_{kj}\right\}_{ij}.$$

Similarly, we can write matrix $W$ as:

$$W = \left\{\sum_{l=1}^{m} Y_{li}Y_{lj}\right\}_{ij},$$

where $\{Y_{ij}\}$ are independent standard normal variables.

Further, we have

$$Z_1 = \operatorname{tr} UV = \sum_{a=1}^{q} (UV)_{aa} = \sum_{a=1}^{q} U_{aa}V_{aa} + \sum_{a\neq b} U_{ab}V_{ab}, \tag{9}$$

where r.v. $U_{aa} = \left(\sum_{k=1}^{p} X_{ka}^2 - p\right)/\sqrt{p}$ has zero mean and variance $\mathbf{Var}U_{aa} = 2$.

If $a \neq b$, r.v. $U_{ab} = \sum_{k=1}^{p} X_{ka}X_{kb}/\sqrt{p}$ has zero mean and variance $\mathbf{Var}U_{ab} = 1$.

Analogous results take place for the matrix $V$.

Lemma 6 and (9) imply

$$\begin{aligned}\mathbf{P}(|Z_1| \geqslant B) \quad &\leqslant \quad q\big(\mathbf{P}(|U_{11}| \geqslant \sqrt{B/(2q)}) + \mathbf{P}(|V_{11}| \geqslant \sqrt{B/(2q)})\big)\\ &\quad +(q^2-q)\mathbf{P}(|U_{12}| \geqslant \sqrt{B/(2(q^2-q))})\\ &\quad +(q^2-q)\mathbf{P}(|V_{12}| \geqslant \sqrt{B/(2(q^2-q))}).\end{aligned} \tag{10}$$

For $U_{aa}$ Lemma 5 implies the following inequality:

$$\sup_x \left|\mathbf{P}(U_{aa}/\sqrt{2} < x) - \Phi(x)\right| \leqslant 6.22/\sqrt{p} = 6.22/\sqrt{rm}. \tag{11}$$

Analogous inequality takes place for $V_{aa}$:

$$\sup_x \left|\mathbf{P}(V_{aa}/\sqrt{2} < x) - \Phi(x)\right| \leqslant 6.22/\sqrt{m}. \tag{12}$$

Next consider $U_{ab}$, $a \neq b$. From Lemma 4 we get

$$\sup_x |\mathbf{P}(U_{ab} < x) - \Phi(x)| \leqslant C_{BE} \cdot \mathbf{E}\,|X_{1a}X_{1b}|^3/\sqrt{p} \leqslant 2.55/\sqrt{rm}, \tag{13}$$

as $\mathbf{E}\,|X_{1a}X_{1b}|^3 = 8/\pi$ and constant $C_{BE}$ in Berry–Esseen inequality in Lemma 4 is less then 1.

Similar inequality takes place for $V_{ab}$:

$$\sup_x |\mathbf{P}(V_{ab} < x) - \Phi(x)| \leqslant 2.55/\sqrt{m}. \tag{14}$$

Further, for all $x \geqslant 0$ one has (see for example 7.1.13. in [7])

$$1 - \Phi(x) \leqslant \frac{2}{\pi}\frac{e^{-x^2/2}}{x + \sqrt{x^2 + 8/\pi}}$$

In particular, for $x \geqslant 1$ we get

$$1 - \Phi(x) \leqslant 0.23\, e^{-x^2/2} \tag{15}$$

Combining (10)–(15) we conclude the proof of Lemma 1. □

Lemma 2. *For $Z_2$, $Z_3$ and $Z_4$ from* (4) *and $B$ from* (5) *the following inequality holds:*

$$\mathbf{P}(Z_2 > B) + \mathbf{P}(|Z_3| > B) + \mathbf{P}(Z_4 > B) \leqslant 19.35\, q^2\, \frac{1 + 1/\sqrt{r}}{\sqrt{m}}. \tag{16}$$

*Proof of Lemma 2.* Analogous to (9) and (10) we deduce

$$\mathbf{P}(Z_2 \geqslant B) \leqslant q\mathbf{P}(|V_{11}| \geqslant \sqrt{B/(2q)}) + (q^2 - q)\mathbf{P}(|V_{12}| \geqslant \sqrt{B/(2(q^2-q))}) \tag{17}$$

and

$$\mathbf{P}(|Z_3| \geqslant B) \leqslant q\big(\mathbf{P}(|U_{11}| \geqslant B/(2q))\big) + q\mathbf{P}\big(|V_{11}| \geqslant B/(2q\sqrt{r})\big) \tag{18}$$

and also

$$\mathbf{P}(Z_4 \geqslant B) \leqslant q\mathbf{P}(|U_{11}| \geqslant \sqrt{B/(2q)}) + (q^2 - q)\mathbf{P}(|U_{12}| \geqslant \sqrt{B/(2(q^2-q))}). \tag{19}$$

Combining (11), (12), (15), (17), (18) and (19) we get Lemma 2. □

LEMMA 3. *Let random variables $T$, $Y$ and $Z$ be defined on the same probability space $(\Omega, \mathbf{A}, \mathbf{P})$, and the distribution of $Y$ is absolutely continuous with a bounded density $f_Y(z)$. Suppose for some event $A \in \mathbf{A}$ and for all $\omega \in A$ the following inequality holds:*

$$|T(\omega) - Y(\omega)| \leqslant Z(\omega) \leqslant a$$

*with some positive constant $a$. Then one has:*

$$\sup_x |\mathbf{P}(T < x) - \mathbf{P}(Y < x)| \leqslant \mathbf{P}(A^c) + a \sup_x f_Y(x). \tag{20}$$

*Proof of Lemma 3.* Note that

$$\sup_x |\mathbf{P}(T < x) - \mathbf{P}(Y < x)| \leqslant$$
$$\leqslant \mathbf{P}(A^c) + \sup_x |\mathbf{P}((T < x) \cap A) - \mathbf{P}((Y < x) \cap A)|.$$

Hence, Lemma follows from the relations:

$$\{T < x\} = \{T - Y + Y < x\} = \{Y < x - (T - Y)\},$$

and

$$\{Y < x - Z\} \cap A \subset \{Y < x - (T - Y)\} \cap A \subset \{Y < x + Z\} \cap A.$$

□

The following two lemmas contain two known results on convergence rate in the central limit theorem for independent identically distributed random variables. The first result is about the case with moment restrictions on the distribution of the summands. The second result is about the chi-square random variable, that is represented as a sum of independent identically distributed random variables with a known distribution.

LEMMA 4. *Let the random variables $\xi_1, \xi_2, \ldots$ be independent and identically distributed with $\mathbf{Var}\xi_1 = \sigma^2 > 0$ and $\mathbf{E}|\xi_1|^3 < \infty$. Put $S_n = \xi_1 + \ldots + \xi_n$.*

*Then there exists a constant $C_{BE}$ such that for distribution function $F_{T_n}(x)$ of r.v. $T_n = (S_n - \mathbf{E}S_n)/\sqrt{\mathbf{Var}S_n}$ one has:*

$$\sup_x |F_{T_n}(x) - \Phi(x)| \leqslant C_{BE} \frac{\mathbf{E}|\xi_1 - \mathbf{E}\xi_1|^3}{\sigma^3\sqrt{n}}.$$

The following inequalities are known for $C_{BE}$ (see for example [8]):

$$\frac{\sqrt{10}+3}{6\sqrt{2\pi}} \leqslant C_{BE} \leqslant 0.4748.$$

A random variable with distribution function $G_p(x)$, having the chi-square distribution with $p$ degrees of freedom, can be represented as a sum of $p$ independent identically distributed random variables with chi-square distribution with 1 degree of freedom. This fact makes it possible to find more precise normal approximation error bounds than those that followed from the Berry–Esseen inequality. Namely, the following lemma is true (see [9]):

Lemma 5. *For all $\lambda \in \left(0; \sqrt{3}-1\right)$ and integer $p > 1$ the following inequality holds:*

$$\sup_x \left| G_p\left(p + x\sqrt{2p}\right) - \Phi(x) \right| \leqslant \frac{\min_\lambda \widetilde{D}(\lambda, p)}{\sqrt{p}},$$

*where*

$$\begin{aligned} \widetilde{D}(\lambda, p) \quad = \quad & \frac{2}{\pi}\left( \frac{\sqrt{\pi}}{6} + \frac{2(1-\lambda)}{\sqrt{p}\,(2-2\lambda-\lambda^2)^2} \right. \\ & \left. + \frac{(1+\lambda^2)}{\lambda^2\sqrt{p}}\left(1+\lambda^2\right)^{-p/4} + \frac{1}{\lambda^2\sqrt{p}} \exp\left(-\frac{\lambda^2 p}{4}\right) \right). \end{aligned}$$

Taking $\lambda = 0.5$, and using monotonicity of the function $\widetilde{D}(\lambda, p)$ with respect to $p$ we get

$$\min_\lambda \widetilde{D}(\lambda, p) \quad \leqslant \quad \widetilde{D}(0.5, 1) \leqslant 6.22. \tag{21}$$

Lemma 6. *For any random variables $X$ and $Y$ and any real $a > 0$ the following inequalities hold:*

$$\mathbf{P}(|X+Y| \geqslant 2a) \leqslant \mathbf{P}(|X| \geqslant a) + \mathbf{P}(|Y| \geqslant a)$$

*and*

$$\mathbf{P}(|X \cdot Y| \geqslant a^2) \leqslant \mathbf{P}(|X| \geqslant a) + \mathbf{P}(|Y| \geqslant a).$$

*Proof of Lemma 6* is obvious. □

Lemma 7. *Let the random variables $X_1, \ldots, X_k$ be independent and the inequalities $|\mathbf{P}(X_j \leqslant x) - \Phi(x)| \leqslant D_j$ hold for all $x$ and $j = 1, \ldots, k$ with some constants $D_1, \ldots, D_k$. Then*

$$|\mathbf{P}(\sum_{j=1}^k c_j\, X_j \leqslant x) - \Phi(x)| \leqslant \sum_{j=1}^k D_j,$$

*where* $c_1, \dots, c_k$ *are any constants such that* $c_1^2 + \cdots + c_k^2 = 1$.

*Proof of Lemma 7* see for example in Theorem 3.1 in [10]. □

## 4. Proofs of theorems

We will prove the theorems in the reverse order. The inequalities (2) and (3) are key components in the proofs of Theorems 1 and 2.

*Proof of Theorem 4.* Matrix equality

$$(I + A)^{-1} - (I - A) = A^2 (I + A)^{-1},$$

and definition (1) give

$$\begin{aligned}
&(B + W)^{-1} = \left(\sqrt{p}U + pI_q + \sqrt{m}V + mI_q\right)^{-1} \\
= \;& \frac{1}{p+m}\left(I_q - \frac{1}{p+m}\left(\sqrt{p}U + \sqrt{m}V\right) + \frac{\left(\sqrt{p}U + \sqrt{m}V\right)^2}{p+m}(B + W)^{-1}\right)
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sqrt{m}\,(r+1)\Big((r+1)\,B\,(B+W)^{-1} - rI_q\Big) - \left(\sqrt{r}U - rV\right) \\
= \;& \frac{\sqrt{r}}{\sqrt{m}}U\left(\sqrt{r}U + V\right) + \frac{1}{\sqrt{m}}B\left(\sqrt{r}U + V\right)^2 (B+W)^{-1}.
\end{aligned}$$

Hence, for the traces of the matrices one has:

$$\begin{aligned}
&\left|\sqrt{m}\,(r+1)\Big((r+1)\operatorname{tr} B\,(B+W)^{-1} - rq\Big) - \left(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\right)\right| \\
\leqslant \;& \frac{1}{\sqrt{m}}\sqrt{r}\left(|\operatorname{tr} UV| + \sqrt{r}\operatorname{tr} U^2\right) + \frac{1}{\sqrt{m}}\left|\operatorname{tr}\left[B\left(\sqrt{r}U + V\right)^2 (B+W)^{-1}\right]\right| \\
\leqslant \;& \frac{1}{\sqrt{m}}\sqrt{r}\left(|\operatorname{tr} UV| + \sqrt{r}\operatorname{tr} U^2\right) + \frac{1}{\sqrt{m}}\operatorname{tr}\left(\sqrt{r}U + V\right)^2 \operatorname{tr} B\,(B+W)^{-1}. \quad (22)
\end{aligned}$$

Here we have used the fact that both random matrices $\left(\sqrt{r}U + V\right)^2$ and $B\,(B+W)^{-1}$ are symmetric and positive semi-definite, as for symmetric and positive semi-definite matrices $X$ and $Y$ the following inequality holds (see [11]): $\operatorname{tr} XY \leqslant \operatorname{tr} X \operatorname{tr} Y$.

Since $\operatorname{tr} B\,(B+W)^{-1}$ is involved in both sides of (22), after proper transformations we get (3) provided that condition $\operatorname{tr}\left(\sqrt{r}U + V\right)^2 < (r+1)\,m$ is met. Thus, Theorem 4 is proved. □

*Proof of Theorem 3.* Similarly to the proof of Theorem 4, we can write

$$\begin{aligned}
&\sqrt{m}\left(BW^{-1} - rI_q\right) - \left(\sqrt{r}U - rV\right) \\
= \;& -\sqrt{\frac{r}{m}}UV + \sqrt{m}\left(rI_q + \frac{1}{\sqrt{m}}\sqrt{r}U\right)\left[\left(I_q + \frac{1}{\sqrt{m}}V\right)^{-1} - \left(I_q - \frac{1}{\sqrt{m}}V\right)\right] \\
= \;& -\sqrt{\frac{r}{m}}UV + \frac{1}{\sqrt{m}}\left(rI_q + \frac{1}{\sqrt{m}}\sqrt{r}U\right)V^2\left(I_q + \frac{1}{\sqrt{m}}V\right)^{-1}.
\end{aligned}$$

Therefore, for the traces of matrices we get:

$$\begin{aligned}
&\left|\sqrt{m}\left(\operatorname{tr} BW^{-1} - rq\right) - \left(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\right)\right| \\
\leqslant &\frac{1}{\sqrt{m}}\sqrt{r}\left|\operatorname{tr} UV\right| + \frac{1}{\sqrt{m}}\left|\operatorname{tr}\left[\left(rI_q + \frac{1}{\sqrt{m}}\sqrt{r}U\right)V^2\left(I_q + \frac{1}{\sqrt{m}}V\right)^{-1}\right]\right| \\
\leqslant &\frac{1}{\sqrt{m}}\sqrt{r}\left|\operatorname{tr} UV\right| + \frac{1}{\sqrt{m}}\operatorname{tr} V^2 \operatorname{tr}\left[\left(rI_q + \frac{1}{\sqrt{m}}\sqrt{r}U\right)\left(I_q + \frac{1}{\sqrt{m}}V\right)^{-1}\right] \\
= &\frac{1}{\sqrt{m}}\sqrt{r}\left|\operatorname{tr} UV\right| + \frac{1}{\sqrt{m}}\operatorname{tr} V^2 \operatorname{tr} BW^{-1}. \qquad (23)
\end{aligned}$$

Again we have used the fact that both random matrices $V^2$ and

$$\left[\left(rI_q + \frac{1}{\sqrt{m}}\sqrt{r}U\right)\left(I_q + \frac{1}{\sqrt{m}}V\right)^{-1}\right] = BW^{-1}$$

are symmetric and positive semi-definite.

We see that r.v. $\operatorname{tr} BW^{-1}$ is involved in both sides of the inequality (23). Transforming the inequality under condition $\operatorname{tr} V^2 < m$, we get (2). Thus, Theorem 3 is proved. □

*Proof of Theorem 2.* By Lemma 1 from [6] we can represent Bartlett–Nanda–Pillai statistic

$$T_{BNP} = \sqrt{p}\left(1 + \frac{p}{m}\right)\left\{\left(1 + \frac{m}{p}\right)\operatorname{tr}\left[S_h\left(S_h + S_e\right)^{-1}\right] - q\right\}$$

in terms of matrices $B$ and $W$ of size $q \times q$ rather than matrices $S_h$ and $S_e$ of size $p \times p$, where $S_h$ and $S_e$ are defined in (1). Recall that $B$ and $W$ are independent and have Wishart distributions $W_q(p, I_q)$ and $W_q(m, I_q)$ with $m = n - p + q$ respectively. We shall employ below the following equality (see [6]):

$$\operatorname{tr} S_h\left(S_h + S_e\right)^{-1} = \operatorname{tr} B\left(B + W\right)^{-1}.$$

According to (3), for $Z_1, Z_2, Z_3$ and $Z_4$ (see definition in (4)) under condition $(rZ_4 + 2\sqrt{r}\left|Z_1\right| + Z_2) < (r+1)m$ we obtain the following relations

$$\begin{aligned}
&\left|\sqrt{r}\,T_{BNP} - \left(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\right)\right| \\
= \;&\left|\sqrt{m}\,(r+1)\left((r+1)\operatorname{tr} B\left(B+W\right)^{-1} - rq\right) - \left(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\right)\right| \\
\leqslant \;&\frac{(r+1)\sqrt{r}\left(\left|Z_1\right| + \sqrt{r}Z_4\right) + \left(rq\,(r+1) + \sqrt{r}\left|Z_3\right|/\sqrt{m}\right)S_Z}{(r+1)\sqrt{m} - S_Z/\sqrt{m}},
\end{aligned}$$

where $S_Z = \left(r\operatorname{tr} U^2 + 2\sqrt{r}\left|\operatorname{tr} UV\right| + \operatorname{tr} V^2\right).$

Therefore, for $m \geqslant M_2$ the inequalities (7) and (20) imply

$$\sup_z \left| \mathbf{P}\Big( \frac{T_{BNP}}{\sqrt{2q\,(1+r)}} < z \Big) - \mathbf{P}\Big( \frac{\operatorname{tr} U - \sqrt{r}\operatorname{tr} V}{\sqrt{2q\,(1+r)}} < z \Big) \right|$$
$$\leqslant \sum_{i=1}^{4} \mathbf{P}(|Z_i| > B) + K_4(r,q)\,\frac{\ln m}{\sqrt{m}} \cdot \sup_x f(x), \tag{24}$$

where $f(x)$ is probability density function of random variable $\frac{(\operatorname{tr} U - \sqrt{r}\operatorname{tr} V)}{\sqrt{2q(1+r)}}$ and $K_4\,(r,q)$ is some computable constant depending on $r$ and $q$.

Note that since $B$ and $W$ are independent then $U$ and $V$ are also independent. It is known (see for example Ch. 2 in [1]), that $\operatorname{tr} B$ and $\operatorname{tr} W$ have chi-square distributions with $pq$ and $mq$ degrees of freedom respectively. Moreover, probability density function of chi-square distribution with $k \geqslant 3$ degrees of freedom is bounded above by $1/(2\sqrt{\pi\,(k-2)})$. Hence, probability density function $f(x)$ admits the following uniform bound:

$$f(x) \leqslant \min \left( \frac{\sqrt{p}}{\sqrt{(pq-2)}}, \frac{\sqrt{m}}{\sqrt{r\,(mq-2)}} \right) \cdot \frac{\sqrt{q(1+r)}}{\sqrt{2\pi}}. \tag{25}$$

Combining Lemmas 5 and 7 and relations (1), (8), (16), (21), (24) and (25) we get Theorem 2. □

*Proof of Theorem 1.* Arguing similarly to the proof of Theorem 2 we represent Lawley–Hotelling statistic

$$T_{LH} = \sqrt{p}\left( \frac{m}{p}\operatorname{tr} S_h S_e^{-1} - q \right)$$

in terms of independent matrices $B \sim W_q\,(p, I_q)$ and $W \sim W_q(m, I_q)$ satisfying

$$\operatorname{tr} S_h S_e^{-1} = \operatorname{tr} BW^{-1}.$$

According to (2) for $Z_1, Z_2$ and $Z_3$ (see definition in (4)) under condition $Z_2 < m$ we obtain

$$\begin{aligned} |\sqrt{r}\,T_{LH} - \big(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\big)\,| &= \big|\sqrt{m}\,\big(\operatorname{tr} BW^{-1} - rq\big) - \big(\sqrt{r}\operatorname{tr} U - r\operatorname{tr} V\big)\big| \\ &\leqslant \frac{\sqrt{r}\,|Z_1| + (rq + \sqrt{r}\,|Z_3|\,/\sqrt{m})\,Z_2}{\sqrt{m} - Z_2/\sqrt{m}}. \end{aligned}$$

Therefore, for $m \geqslant M_1$ the inequalities (6) and (20) imply that

$$\sup_z \left| \mathbf{P}\Big( \frac{T_{LH}}{\sqrt{2q\,(1+r)}} < z \Big) - \mathbf{P}\Big( \frac{\operatorname{tr} U - \sqrt{r}\operatorname{tr} V}{\sqrt{2q\,(1+r)}} < z \Big) \right|$$
$$\leqslant \sum_{i=1}^{3} \mathbf{P}(|Z_i| > B) + K_3(r,q)\,\frac{\ln m}{\sqrt{m}} \cdot \sup_x f(x), \tag{26}$$

where $f(x)$ is the probability density function of random variable $(\operatorname{tr} U - \sqrt{r}\operatorname{tr} V)/\sqrt{2q\,(1+r)}$ and $K_3\,(r,q)$ is some computable constant depending on $r$ and $q$.

Combining Lemmas 5 and 7 and relations (1), (8), (16), (21), (26) and (25), we conclude the proof of Theorem 1. $\square$

## References


[1] Y. Fujikoshi, V. V. Ulyanov and R. Shimizu, Multivariate Statistics: High-dimensional and large-sample approximations, John Wiley & Sons, Hoboken, NJ, 2010.

[2] T. W. Anderson, An introduction to multivariate analysis (3rd ed.), John Wiley & Sons, Hoboken, NJ, 2003.

[3] R. J. Muirhead, Aspects of Multivariate Statistical Theory, John Wiley & Sons, Hoboken, NJ, 1982.

[4] Y. Fujikoshi, V. V. Ulyanov and R. Shimizu, $L_1$-norm error bounds for asymptotic expansions of multivariate scale mixtures and their applications to Hotelling's generalized $T_0^2$, J. of Multivariate Analysis, **96** (2005), No. 1, 1–19.

[5] A. A. Lipatiev and V. V. Ulyanov, On Computable Estimates for Accuracy of Approximation for the Bartlett–Nanda–Pillai Statistic, Siberian Adv. Math., **27** (2017), No. 3, 153–159.

[6] H. Wakaki, Y. Fujikoshi and V. V. Ulyanov, Asymptotic expansions of the distributions of MANOVA test statistics when the dimension is large, Hiroshima Math. J., **44** (2014), No. 3, 247–259.

[7] M. Abramowitz and I. A. Stegun, Handbook of mathematical functions with formulas, graphs, and mathematical tables, National Bureau of Standards Applied Mathematics Series, No. 55, U.S. Government Printing Office, Washington, DC, 1964.

[8] I. G. Shevtsova, On the absolute constants in the Berry–Esseen type inequalities for identically distributed summands, Available at:https://arxiv.org/pdf/1111.6554 (accessed November 28, 2011).

[9] Y. Kawaguchi, V. V. Ulyanov and Y. Fujikoshi, Asymptotic distributions of basic statistics in geometric representation for high-dimensional data and their error bounds, Informatics and Applications, **4** (2010), No. 1, 22—27.

[10] V. V. Ulyanov, H. Wakaki and Y. Fujikoshi, Berry–Esseen bound for high dimensional asymptotic approximation of Wilks' Lambda distribution, Statistics and Probability Letters, **76** (2006), No. 12, 1191—1200.

[11] I. D. Coope, On matrix trace inequalities and related topics for products of Hermitian matrices, J. of Mathematical Analysis and Applications, **188** (1949), No. 3, 999–1001.

*Alexander A. Lipatiev*
*Department of Mathematical Statistics*
*Faculty of Computational Mathematics and Cybernetics*
*Lomonosov Moscow State University*
*Moscow, 119991, Russia*
*E-mail: allipatev@cs.msu.ru*

*Vladimir V. Ulyanov*
*Faculty of Computational Mathematics and Cybernetics*
*Lomonosov Moscow State University*
*Moscow, 119991, Russia*
*National Research University Higher School of Economics, Russia*
*E-mail: vulyanov@cs.msu.ru*