Multivariate Analysis of Variance with fewer observations than the dimension

by

Muni S. Srivastava
Department of Statistics
University of Toronto

and

Yasunori Fujikoshi
Graduate School of Science and Engineering
Chuo University

Technical Report No. 0206 May 5, 2002

TECHNICAL REPORT SERIES

University of Toronto

Department of Statistics

Multivariate Analysis of Variance with fewer observations than the dimension

By Muni S. Srivastava, and Yasunori Fujikoshi,
Department of Statistics, University of Toronto, and
Graduate School of Science and Engineering, Chuo University

Abstract

In this article, we consider the problem of testing a linear hypothesis in a multivariate linear regression model which includes the case of testing the equality of mean vectors of several multivariate normal population with common covariance matrix Σ, the so called multivariate analysis of variance or MANOVA problem. However, we have fewer observations than the dimension of the random vectors. Two tests are proposed and their asymptotic distributions under the hypothesis as well as under the alternatives are given under some mild conditions. A theoretical comparsion of these powers is made.

Key words and phrases: Distribution of test statistics, DNA microarray data, fewer observations than dimension, Moore-Penrose Inverse, multivariate analysis of variance, singular Wishart.

Short Title: MANOVA with Fewer Observations
AMS 1991 Subject classifications. Primary 62H10, 62H15, 62H30; Secondary 62H12.

1 Introduction

Consider the multivariate linear regression model in which the $N \times p$ observation matrix Y is related by

$$Y = X\Xi + E \ , \tag{1.1 }$$

where X is the $N \times k$ design matrix of rank $k < N$, assumed known, and Ξ is the $k \times p$ matrix of unknown parameters. We shall assume that the N row vectors of E are independent and identically distributed (hereafter denoted as iid) as multivariate normal with mean vector zero and covariance matrix Σ, denoted as $\boldsymbol{e}_i \sim N_p(\boldsymbol{0}, \Sigma)$, where $E' = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_N)$. Similarly, we write $Y' = (\boldsymbol{y}_1, \dots, \boldsymbol{y}_N)$ where $\boldsymbol{y}_1, \dots, \boldsymbol{y}_N$ are p-vectors independently distributed as multivariate normal with common covariance matrix Σ. We shall assume that

$$N \leq p \ , \tag{1.2 }$$

that is there are fewer observations than the dimension p. Such a situation arises when there are thousands of gene expressions on microarray data but with observations on only few subjects. The maximum likelihood or the least squares estimates of Ξ is given by

$$\hat{\Xi} = (X'X)^{-1}X'Y \ : \ k \times p \ . \tag{1.3 }$$

The $p \times p$ covariance matrix Σ can be unbiasedly estimated by

$$\hat{\Sigma} = n^{-1} W \ ,$$

where $n = N - k$,

$$W = (Y - X\hat{\Xi})'(Y - X\hat{\Xi}) \ , \tag{1.4 }$$

and often called as the matrix of the sum of squares and products due to error or simply 'within' matrix. The $p \times p$ matrix W is, however, a singular matrix of rank n which is less than p; see Srivastava [11] for its distributional results. We consider the problem of testing the linear hypothesis.

$$H \ : \ C\Xi = 0 \text{ vs } A \ : \ C\Xi \neq 0 \ , \tag{1.5 }$$

where C is a $q \times k$ matrix of rank $q \leq k$ of known constants. The matrix of the sum of squares and products due to the hypothesis, or , simply 'between' matrix is given by

$$B = N \left(C\hat{\Xi} \right)' [CGC']^{-1} C\hat{\Xi} \ , \tag{1.6 }$$

where

$$G = \left[N^{-1} X'X \right]^{-1} \ . \tag{1.7 }$$

When normality is not assumed, it is often required that G converges to a $k \times k$ positive definite (p.d.) matrix for asymptotic normality to hold, and although a weaker condition than (1.7) has been given in Srivastava [12] for asymptotic normality to hold, we will assume that G is positive definite. Under the assumption of normality,

$$W \sim W_p(\Sigma, n) \tag{1.8 }$$

and

$$B \sim W_p(\Sigma, q, N\eta\eta') \tag{1.9 }$$

are independently distributed as Wishart and non-central Wishart matrices respectively, where

$$\eta = (\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_q) = (C\Xi)'(CGC')^{-\frac{1}{2}} \ . \tag{1.10 }$$

Thus, we may write

$$B = ZZ' \ , \tag{1.11 }$$

where $Z = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_q)$ and $\boldsymbol{z}_i$ are independently distribtued as $N_p(N^{\frac{1}{2}}\boldsymbol{\eta}_i, \Sigma)$. Since $n < p$, the likelihood ratio test is not available. Also, the sample space χ consists of $p \times N$ matrices of rank $N \leq p$, since Σ is positive definite. Thus, any point in χ can be transformed to another point of χ by an element of the group Glp of $p \times p$ nonsingular matrices. Hence, the group Glp acts transitively on the sample space and the only α-level test $\Psi(B, W)$ that is affine invariant is $\Psi(B, W) \equiv \alpha$, see Lehmann [6, pp.318] and Eaton [4]. Thus, we look for tests that are invariant under a smaller group. In particular, we will be considering tests that are invariant under the transformation $\boldsymbol{y}_i \rightarrow c\Gamma\boldsymbol{y}_i$, where $Y' = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N)$,

$c \neq 0, c \in R_{(0)}$ and $\Gamma \in O_p$: $R_{(0)}$ is the real line without zero and O_p is the group of $p \times p$ orthogonal matrices. Clearly $c\Gamma$ is a subgroup of Glp. Define

$$\begin{aligned} \hat{a}_1 &= (\text{tr } W)/np \,, \\ \hat{a}_2 &= \frac{1}{(n-1)(n+2)p}\left[\text{tr } W^2 - \frac{1}{n}(\text{tr } W)^2\right] \,, \end{aligned} \tag{1.12 }$$

and

$$\hat{b} = (\hat{a}_1^2/\hat{a}_2) \; .$$

Let

$$a_i = (\text{tr } \Sigma^i)/p, \;\; i = 1, \ldots, 4, \;\; \text{and } b = (a_1^2/a_2). \tag{1.13 }$$

We shall assume that

$$0 < \lim_{p\to\infty} a_i = a_{i0} < \infty \;, i = 1, \ldots, 4 \; . \tag{1.14 }$$

It has been shown in Srivastava [9] that under the condition (1.14), $\hat{a}_i$ are consistent estimators of a_i as n and $p \to \infty$. Thus, $\hat{b}$ is a consistent estimator of b. To propose tests for the testing problem defined in (1.5), when $N < p$, we note that the likelihood ratio tests or other invariant tests (under a group of nonsingular transformations) such as Lawley-Hotelling test or Bartlett-Nanda-Pillai test described in most text books are not available. However, we may consider a generalization of Dempster [2] test which is given by

$$\tilde{T}_1 = \frac{(pq)^{-1} \text{ tr } B}{\hat{a}_1} = \frac{n \text{ tr } B}{q \text{ tr } W} \; . \tag{1.15 }$$

However, its exact distribution even under the hypothesis is difficult to obtain. An approximate distribution of $\tilde{T}_1$ under the hypothesis is $F_{[q\hat{r}],[n\hat{r}]}$, where $F_{m,n}$ denotes the F-distribution with m and n degrees of freedom, and $[a]$ denotes the largest integer $\leq a$. The above approximate distribution of the $\tilde{T}_1$ statistic under the hypothesis is obtained by assuming that tr $B \sim m\chi^2_{qr}$ and tr $W \sim m\chi^2_{nr}$, both independently distributed. By equating the first two moments of tr B under the hypothesis with that of $m\chi^2_{qr}$, it is found that $r = pb$. Srivastava [10] proposed to estimate r by

$$\hat{r} = p\,\hat{b} \; . \tag{1.16 }$$

It may be noted that since F is invariant under the scalar transformations, no estimate of m is required to obtain the approximate distribution. To study the power of the $\tilde{T}_1$ test, we consider a normalized version of $\tilde{T}_1$ given by

$$\begin{aligned} T_1 &= \left[\frac{p\hat{b}}{2q(1+n^{-1}q)}\right]^{\frac{1}{2}} \left[\frac{\text{tr } B - pq\hat{a}_1}{p\hat{a}_1}\right] \\ &= \left[\frac{p}{2q\hat{a}_2(1+n^{-1}q)}\right]^{\frac{1}{2}} \left[p^{-1} \text{ tr } B - q\hat{a}_1\right] \\ &= \left[2q\hat{a}_2(1+n^{-1}q)\right]^{-\frac{1}{2}} \left[\frac{\text{tr } B}{\sqrt{p}} - \frac{q}{\sqrt{n}}\frac{\text{tr } W}{\sqrt{np}}\right] \; . \end{aligned} \tag{1.17 }$$

It may be noted that $p\hat{b}$ is ratio consistent estimator of r. Dempster [2,3] has proposed two other ratio consistent estimators of $r = pb$. However, these estimators are iterative solutions of two equations. Irrespective of which consistent estimator of $r = pb$ is used, the asymptotic theory remains the same due to Slutsky's theorem, see Rao [7]. The expression in (1.17) is a generalization of Bai and Saranadasa [1] test for the two-sample problem. Next, we describe another test statistic proposed by Srivastava [10] for the testing problem described in (1.5). This statistic uses the Moore-Penrose inverse of W. The Moore-Penrose inverse of a matrix A is defined by A^+ which satisfies the following four conditions:

$$\begin{array}{ll} (i) & AA^+A = A \ , \\ (ii) & A^+AA^+ = A^+ \ , \\ (iii) & (AA^+)' = AA^+ \ , \\ (iv) & (A^+A)' = A^+A \ . \end{array}$$

The Moore-Penrose inverse is unique. The statistic proposed by Srivastava [10] is given by

$$T_2 = -p\hat{b} \log \Pi_{i=1}^{q} \ (1 + c_i)^{-1} \ , \tag{1.18 }$$

where c_i are the non-zero eigenvalues of BW^+. In a sense, it is an adapted version of the likelihood ratio test. Other tests which may be considered as adapted versions of Laweley-Hotelling's, and Bartlett-Nanda-Pillai tests are given by

$$T_3 = p\hat{b} \sum_{i=1}^{q} c_i \ , \tag{1.19 }$$

and

$$T_4 = p\hat{b} \sum_{i=1}^{q} \frac{c_i}{1 + c_i} \ , \tag{1.20 }$$

respectively. It will be shown in Section 2 that as $p \to \infty$, the tests T_2, T_3 and T_4 are asymptotically equivalent. Thus, in the final analysis, we only consider the two test statistics T_1 and T_2. The distributions of these two statistics under the null hypothesis are given in Section 3, and under local alternatives in Section 4. The power comparison is carried out in Section 5. We may note that all the five tests $\tilde{T}_1$, T_1, T_2, T_3, T_4 are invariant under the group of linear transformations $\boldsymbol{y}_i \to c\Gamma \boldsymbol{y}_i$, for $c \neq 0$, $c \in R_{(0)}$ and $\Gamma \in O_p$, where $R_{(0)}$ is the real line except zero and O_p is the group of $p \times p$ orthogonal matrices. Thus, without any loss of generality, we may assume that the population covariance matrix Σ is a diagonal matrix. Thus

$$\Sigma = \wedge = \mathrm{diag}(\lambda_1, \ldots, \lambda_p) \ . \tag{1.21 }$$

2 Asymptotic Equivalence of the test Statistics T_2, T_3 and T_4

In this section, we show that as $p \to \infty$, the three test statistics T_2, T_3 and T_4 are asymptotically equivalent. Since , the $p \times p$ matrix W is of rank $n < p$, there exists a semi-orthogonal $n \times p$ matrix H such that

$$W = H'LH \ , HH' = I_n \ , \tag{2.1 }$$

where $L = \text{diag}(\ell_1, \ldots, \ell_n)$ is a diagonal matrix whose diagonal elements are the non-zero eigenvalues of the matrix W. The Moore-Penrose of the $p \times p$ matrix W is given by

$$W^+ = H' L^{-1} H \ . \tag{2.2 }$$

Since,

$$B = ZZ' \ ,$$

the non-zero eigenvalues of BW^+ are the same as the non-zero eigenvalues of the matrix

$$\begin{aligned} Z'W^+Z &= Z'H'L^{-1}HZ \\ &= Z'H'A(ALA)^{-1}AHZ \\ &= U'(ALA)^{-1}U \ , \end{aligned} \tag{2.3 }$$

where

$$A = (H \wedge H')^{-\frac{1}{2}} \ , \tag{2.4 }$$

and

$$U = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_q) = AHZ \ . \tag{2.5 }$$

Given $H, \boldsymbol{u}_i$ are independently distributed as $N_n(N^{\frac{1}{2}} AH\eta_i, I)$, and in the notation of Srivastava and Khatri [13, pp. 54],

$$U \sim N_{n,q}(N^{\frac{1}{2}} AH\eta, I_n, I_q) \ . \tag{2.6 }$$

From Lemma A.1 given in the Appendix, we get in probability

$$\lim_{p\to\infty} \frac{ALA}{p} = \frac{a_{10}^2}{a_{20}} I_n = \lim_{p\to\infty} bI_n \ , \tag{2.7 }$$

where

$$0 < a_{i0} = \lim_{p\to\infty} (\text{tr } \Sigma^i/p) < \infty, \ i = 1, \ldots, 4 \ . \tag{2.8 }$$

A consistent estimator of b, as n and $p \to \infty$, is given by $\hat{b}$, defined in (1.12) see Srivastava [9]. Thus, for the statistic T_2, we get in probability

$$\begin{aligned} \lim_{p\to\infty} \hat{b}p \ \log \prod_{i=1}^{q} (1 + c_i) &= \lim_{p\to\infty} \hat{b}p \ \log |I_q + Z'W^+Z| \\ &= \lim_{p\to\infty} \hat{b}p \left[\text{tr} Z'W^+Z - \frac{1}{2} \text{tr}(Z'W^+Z)^2 + \ldots \right] \\ &= \lim_{p\to\infty} (\hat{b}/b) \ \text{tr } U'U \ . \end{aligned} \tag{2.9 }$$

Similarly, in probability

$$\begin{aligned} \lim_{p\to\infty} p\hat{b} \sum_{i=1}^{q} c_i &= \lim_{p\to\infty} p\hat{b} \text{tr}(Z'W^+Z) \ , \\ &= \lim_{p\to\infty} (\hat{b}/b) \text{tr} U'U \ . \end{aligned}$$

Thus, as $p \to \infty$, the tests T_2 and T_3 are equivalent. For the test T_4, we note that in probability

$$\begin{aligned}
\lim_{p\to\infty} p\hat{b}\left[q - \sum_{i=1}^{q}(1+c_i)^{-1}\right] &= \lim_{p\to\infty} p\hat{b}\left[q - \operatorname{tr}(I_q + Z'W^{+}Z)^{-1}\right] \\
&= \lim_{p\to\infty} p\hat{b}\left[\operatorname{tr}Z'W^{+}Z - \operatorname{tr}(Z'W^{+}Z)^2 + \ldots,\right] \\
&= \lim_{p\to\infty} (\hat{b}/b)\operatorname{tr}(U'U) .
\end{aligned}$$

Thus, all the three tests T_2, T_3 and T_4 are asymptotically equivalent as $p \to \infty$. Thus, we need to consider only the test T_2 (among the three) which will be compared with the test T_1.

3 Distribution of the test statistics T_1 and T_2 under the hypothesis

Under the hypothesis, we have

$$B = ZZ' \sim W_p(\wedge, q) ,$$

where

$$Z = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_q)$$

and $\boldsymbol{z}_i$ are iid $N_p(\mathbf{0}, \wedge)$. The within matrix

$$W \sim W_p(\wedge, n) ,$$

and B and W are independently distributed. Since $\hat{b} \to b$, and $\hat{a}_2 \to a_2$ in probability, it follows from Slutsky's theorem, see Rao [7], that we need only consider the distribution of

$$\begin{aligned}
T_0 &= \left[\frac{\operatorname{tr} B}{\sqrt{p}} - \frac{q}{\sqrt{n}}\frac{\operatorname{tr} W}{\sqrt{np}}\right] \\
&= \frac{1}{\sqrt{p}}\left[\operatorname{tr} B - \frac{q}{n}\operatorname{tr} W\right] \qquad (3.1) \\
&= \frac{1}{\sqrt{p}}\left[\operatorname{tr} \wedge U_1 - \frac{q}{n}\operatorname{tr} \wedge U_2\right] ,
\end{aligned}$$

where $U_1 \sim W_p(I, q)$, and $U_2 \sim W_p(I, n)$ are independently distributed. Let $U_1 = (u_{1ij})$, and $U_2 = (u_{2ij})$. Then u_{1ii} are independently distributed as a chi-squared random variables with q degrees of freedom, denoted by χ^2_q. Similarly, u_{2ii} are iid χ^2_n. Hence, from (3.1)

$$\begin{aligned}
T_0 &= \frac{1}{\sqrt{p}}\sum_{i=1}^{p}\lambda_i(u_{1ii} - n^{-1}qu_{2ii}) \\
&\equiv \frac{1}{\sqrt{p}}\sum_{i=1}^{p}\lambda_i v_{ii} , \qquad (3.2)
\end{aligned}$$

where v_{ii} are iid with mean 0 and variance $2q + 2n^{-1}q^2$. Hence,

$$\begin{aligned} \mathrm{Var}(T_0) &= \frac{1}{p}\sum_{i=1}^{p} \lambda_i^2 \left[2q + 2n^{-1}q^2\right] \\ &= 2q(1+n^{-1}q)a_2 \\ &= \sigma_1^2 < \infty \ . \end{aligned} \qquad (3.3\ )$$

Let

$$T_1^* = \left\{2q(1+n^{-1}q)a_2\right\}^{-1/2} T_0 \ . \qquad (3.4\ )$$

Then, if

$$\frac{\max_{1\le i\le p} \lambda_i/\sqrt{p}}{\sqrt{a_2}} \to 0 \text{ as } p \to \infty \ ,$$

it follows from Srivastava [12] that T_1^* is asymptotically normally distributed as $p \to \infty$. Since, it is assumed that $a_2 < \infty$, we assume that

$$\lambda_i = O\left(p^{\gamma}\right), \ 0 \le \gamma < \frac{1}{2} \ . \qquad (3.5\ )$$

Hence, from Slutsky's theorem, we get the following theorem.

Theorem 3.1 Under the null hypothesis and condition (3.5)

$$\lim_{n\to\infty} \lim_{p\to\infty} \left[P_0(T_1 < z) - \Phi(z)\right] = 0 \ ,$$

where P_0 denotes that the probability has been computed under the hypothesis that $\eta = 0$.

Next, we consider the asymptotic distribution of the statistic T_2 given by

$$\begin{aligned} T_2 &= p\hat{b}\log|I_p + BW^+| \\ &= p\hat{b}\log|I_q + Z'W^+Z| \end{aligned} \qquad (3.6\ )$$

under the hypothesis H, where Z and W are independently distributed and $Z \sim N_{p,q}(0, \Sigma, I_q)$. From (2.9)

$$\lim_{p\to\infty} T_2 = \lim_{p\to\infty} (\hat{b}/b) \text{ tr } U'U \ , \qquad (3.7\ )$$

where under the hypothesis, $U' = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \ : \ q \times n$, and $\boldsymbol{u}_i$ are iid $N_q(0, I_q)$. Thus, we get the following theorem.

Theorem 3.2 Under the null hypothesis

$$\lim_{n\to\infty} \lim_{p\to\infty} \left[P_0\left(\frac{T_2 - nq}{\sqrt{2nq}} < z\right) - \Phi(z)\right] = 0 \ ,$$

where P_0 denotes that the probability is being calculated under the hypothesis.

4 Distribution of the statistic T_1 and T_2 under the alternative

Before we derive the non-null distribution of the statistics T_1 and T_2, we shall first consider the statistic $\tilde{T}_1$ defined in (1.15). As mentioned earlier, the statistic $\tilde{T}_1$ is also invariant under the transformation $\boldsymbol{y}_i \to c\Gamma\boldsymbol{y}_i$, $c \neq 0$, $\Gamma\Gamma' = I_p$. Thus, the covariance matrix Σ can be assumed to be diagonal as given in (1.21). Futhermore, since all the statistics are invariant under scalar transformations, the assumption that $\Sigma = \sigma^2 I_p$ is equivalent to assuming that $\Sigma = I_p$ for the distributional purposes. Thus, when $\Sigma = I_p$, the $\tilde{T}_1$ statistic has a non-central F distribution with pq and np degrees of freedom with non-centrality parameter

$$\gamma^2 = N \text{ tr } \eta\eta' \ . \tag{4.1 }$$

Next, we easily obtain the following theorem from Simaika [8].

Theorem 4.1 Assume that $\wedge = \lambda I_p$. Then, for testing the hypothesis that tr $\eta\eta' = 0$ against the alternative that tr $\eta\eta' \neq 0$, the $\tilde{T}_1$ test is uniformly most powerful among all tests whose powers depend on γ^2.

From the above theorem, it implies that any other test whose power depends on γ^2, will have power no more than the $\tilde{T}_1$-test. It will be shown in the next two theorems that the power of the two tests T_1 and T_2 depends only on γ^2 when $\Sigma = I_p$.

Thus, before using either of the two tests T_1 and T_2, the spherecity hypothesis should be tested by a test proposed by Srivastava [9], when $n = O(p^\delta)$, $0 < \delta \leq 1$.

4.1 Non-null distribution of the test statistic T_1

Let

$$\Omega = N \wedge^{-\frac{1}{2}} \eta\eta' \wedge^{-\frac{1}{2}} \ . \tag{4.2 }$$

We shall assume that

$$0 \leq \frac{\text{tr} \wedge^i \Omega}{p} < \infty, \ i = 1, 2 \ . \tag{4.3 }$$

Under the alternative hypothesis, B and W are independently distributed where

$$W \sim W_p(\wedge, n) \ ,$$

and

$$B \sim W_p(\wedge, q, N\eta\eta') \ ,$$

a non-central Wishart distribution with non-centrality matrix

$$N\eta\eta' = \wedge^{\frac{1}{2}} \Omega \wedge^{\frac{1}{2}} \ . \tag{4.4 }$$

Define

$$u = \frac{1}{\sqrt{p}}[\text{tr}B - q\text{tr} \wedge - \text{tr} \wedge \Omega]$$

and

$$v = \frac{1}{\sqrt{np}}[\mathrm{tr}W - n\mathrm{tr}\wedge] \ .$$

Then the following lemma is required to derive the asymptotic non-null distribution of T_1.

Lemma 4.1 As $p \to \infty$, and under the conditions (4.3) and (1.14),

$$\begin{aligned} u &\xrightarrow{d} N\left[0, 2qa_2 + 4\mathrm{tr}(\wedge^2\Omega/p)\right] \ , \\ v &\xrightarrow{d} N(0, 2a_2) \ , \end{aligned}$$

where $\xrightarrow{d}$ denotes 'in distribution'.

Proof. The result has been essentially used in Fujikoshi et al. [5]. Here, we give a detailed derivation. The characteristic function of u is given by

$$\begin{aligned} \Phi_u(t) &= E(e^{itu}) \\ &= e^{-\frac{it}{\sqrt{p}}(q\mathrm{tr}\wedge + \mathrm{tr}\wedge\Omega)} \times E\left(e^{\frac{it}{\sqrt{p}}\mathrm{tr}B}\right) \\ &= e^{-\frac{it}{\sqrt{p}}(q\mathrm{tr}\wedge + \mathrm{tr}\wedge\Omega)} \times |I_p - \frac{2it}{\sqrt{p}} \wedge|^{-\frac{1}{2}q} \left(e^{\frac{it}{\sqrt{p}}\mathrm{tr}\wedge(I - \frac{2it}{\sqrt{p}}\wedge)^{-1}\Omega}\right) \ , \end{aligned}$$

see Srivastava and Khatri [13, Theorem 3.3.10, pp. 85]. Now expanding, see Srivastava and Khatri [13, pp. 33, pp. 37],

$$\begin{aligned} \log\left|I_p - \frac{2it}{\sqrt{p}}\wedge\right|^{-\frac{1}{2}q} &= -\frac{1}{2}q\log\left|I_p - \frac{2it}{\sqrt{p}}\wedge\right| \\ &= \frac{1}{2}q\left[\frac{2it}{\sqrt{p}}\mathrm{tr}\wedge + \frac{1}{2}\left(\frac{2it}{\sqrt{p}}\right)^2 \mathrm{tr}\wedge^2\right] + o(1) \\ &= \frac{itq}{\sqrt{p}}\mathrm{tr}\wedge + \frac{q(it)^2}{p}\mathrm{tr}\wedge^2 + o(1) \ , \end{aligned}$$

and

$$\begin{aligned} \frac{it}{\sqrt{p}}\mathrm{tr}\wedge\left(I_p - \frac{2it\wedge}{\sqrt{p}}\right)^{-1}\Omega &= \frac{it}{\sqrt{p}}\mathrm{tr}\wedge\left[I + \frac{2it}{\sqrt{p}}\wedge + \frac{1}{2}\left(\frac{2it}{\sqrt{p}}\right)^2\wedge^2\right]\Omega + o(1) \\ &= \frac{it\ \mathrm{tr}\wedge\Omega}{\sqrt{p}} + \frac{2(it)^2}{p}\mathrm{tr}\wedge^2\Omega + o(1) \ . \end{aligned}$$

Hence,

$$E\left(e^{itu}\right) = e^{\frac{1}{2}(it)^2[2qa_2 + 4(\mathrm{tr}\wedge^2\Omega/p)]} \times (1 + o(1)) \ . \tag{4.5 }$$

Thus, $u \sim N(0, 2qa_2 + 4\mathrm{tr}(\wedge^2\Omega/p))$ as $p \to \infty$. The charactertistic function of v is given by

$$\begin{aligned} \Phi_v(t) &= E\left(e^{itv}\right) \\ &= \left[e^{-\frac{it}{\sqrt{np}}n\mathrm{tr}\wedge}\right]\left[\left|I - \frac{2it}{\sqrt{np}}\wedge\right|^{-\frac{n}{2}}\right] \ . \end{aligned}$$

As before, we have

$$-\frac{n}{2}\log\left|I_p - \frac{2it}{\sqrt{np}}\wedge\right| = \frac{n}{2}\left[\frac{2it}{\sqrt{np}}\mathrm{tr}\wedge + \frac{(2it)^2}{2np}\mathrm{tr}\wedge^2\right] + o(1) \ .$$

Hence,

$$\Phi_v(t) = e^{(it)^2(\mathrm{tr}\wedge^2)/p}\,(1+o(1))\ .$$

Thus, as $p \to \infty, v \to N(0, 2a_2)$. This proves both parts of the lemma.

Thus, we have

$$\begin{aligned}\left(u - \frac{q}{\sqrt{n}}v\right) &= \frac{1}{\sqrt{p}}\left[\mathrm{tr}B - q\mathrm{tr}\wedge - \mathrm{tr}\wedge\Omega - \frac{q}{n}\mathrm{tr}W + q\mathrm{tr}\wedge\right] \\ &= \frac{1}{\sqrt{p}}\left[\mathrm{tr}B - \frac{q}{n}\mathrm{tr}W - \mathrm{tr}\wedge\Omega\right]\ .\end{aligned}$$

Note that u and v are independently distributed. Hence, from Lemma 4.1 ,

$$u - \frac{q}{\sqrt{p}}v \sim N\left(0, \sigma_1^{*2}\right)\ ,$$

as $p \to \infty$, where

$$\begin{aligned}\sigma_1^{*2} &= 2qa_2 + 4\mathrm{tr}(\wedge^2\Omega/p) + \frac{2q^2}{n}a_2 \\ &= 2qa_2(1+n^{-1}q) + 4\mathrm{tr}(\wedge^2\Omega/p) \\ &= \sigma_1^2 + 4\mathrm{tr}\wedge^2\Omega/p\ .\end{aligned} \tag{4.6 }$$

Hence, as $p \to \infty$,

$$\frac{\mathrm{tr}B - n^{-1}q\mathrm{tr}W - \mathrm{tr}\wedge\Omega}{\sigma_1^*\sqrt{p}} \xrightarrow{d} N(0,1)\ .$$

Thus,

$$P\left\{\frac{\mathrm{tr}B - n^{-1}q\mathrm{tr}W}{\sigma_1\sqrt{p}} > z_\alpha | A\right\} = P\left\{\frac{\mathrm{tr}B - n^{-1}q\mathrm{tr}W - \mathrm{tr}\wedge\Omega}{\sigma_1^*\sqrt{p}} > \frac{\sigma_1}{\sigma_1^*}z_\alpha - \frac{\mathrm{tr}\wedge\Omega}{\sigma_1^*\sqrt{p}}\right\}\ .$$

Theorem 4.2 Assume that conditions (1.14) and (4.3) holds. Then, when $\eta \neq 0$,

$$\lim_{n\to\infty}\lim_{p\to\infty} P_1\left[T_1 > z_\alpha\right] = \lim_{n\to\infty}\lim_{p\to\infty}\Phi\left[-\frac{\sigma_1}{\sigma_1^*}z_\alpha + \frac{\mathrm{tr}\wedge\Omega}{\sigma_1^*\sqrt{p}}\right]\ .$$

For local alternatives, we assume that

$$\mathrm{tr}\wedge\Omega = O(\sqrt{p})\ . \tag{4.7 }$$

Writing

$$\eta = (nN)^{-1/2}\Delta\ , \tag{4.8 }$$

we have that the assumption (4.7) is equivalent to

$$\frac{1}{n\sqrt{p}}\mathrm{tr}\Delta\Delta' = O(1) \tag{4.9 }$$

which is satisfied when $\Delta = O(1)$ and $n = O(\sqrt{p})$. Then, from (3.5) we have

$$\frac{\mathrm{tr}\wedge^2\Omega}{p} = \frac{O(p^\gamma\mathrm{tr}\wedge\Omega)}{p} = \frac{O(p^\gamma\sqrt{p})}{p} \to 0 \text{ as } p \to \infty\ ,$$

and

$$\sigma_1^* \to \sigma_1\ .$$

Hence, we get the following Corollary.

Corollary 4.1 For local alternatives satisfying (4.7) or (4.9), the asymptotic power of the T_1-test is given by

$$\begin{aligned}\beta(T_1) &\simeq \Phi\left(-z_\alpha + \frac{\mathrm{tr}\wedge\Omega}{\sigma_1\sqrt{p}}\right)\\ &\simeq \Phi\left(-z_\alpha + \frac{\mathrm{tr}\Delta\Delta'}{n\sqrt{2pqa_2}}\right) .\end{aligned}$$

Thus, when $\Sigma = I_p$, $a_2 = 1$, and

$$\beta(T_1) \simeq \Phi\left(-z_\alpha + \frac{\mathrm{tr}\Delta\Delta'}{n\sqrt{2pq}}\right) .$$

4.2 Non-null distribution of the test statistic T_2.

For the statistic T_2, we derive the power of the T_2 test under the local alternatives where η is given in (4.8) with $\Delta = O(1)$. From (2.9), it follows that

$$\lim_{p\to\infty} T_2 = \lim_{p\to\infty}\left(\frac{\hat{b}}{b}\right)\mathrm{tr}U'U ,$$

where given H

$$U'U \sim W_q(I_q, n, N\eta'H'A^2H\eta), \quad A = (H\wedge H')^{-\frac{1}{2}} .$$

Further, under the local alternative (4.8) we get from Lemma A.1,

$$\begin{aligned}\lim_{p\to\infty} N\eta'H'A^2H\eta &= \lim_{p\to\infty}\Delta'H'A^2H\Delta/n\\ &\to (a_{10}/a_{20})\lim_{p\to\infty}\frac{\Delta'H'H\Delta}{n} .\end{aligned}$$

Writing

$$\Delta = (\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_q) \ : \ p\times q ,$$

we find that

$$n^{-1}\mathrm{tr}\Delta'H'H\Delta = n^{-1}\sum_{i=1}^{q}\boldsymbol{\delta}_i'H'H\boldsymbol{\delta}_i .$$

Hence, from Lemma A.1 ,

$$\lim_{n\to\infty}\lim_{p\to\infty} n^{-1}\boldsymbol{\delta}_i'H'H\boldsymbol{\delta}_i = \lim_{p\to\infty}\left(\frac{\boldsymbol{\delta}_i'\wedge\boldsymbol{\delta}_i}{pa_1}\right) .$$

Thus,

$$\lim_{n\to\infty}\lim_{p\to\infty}\mathrm{tr}N\eta'H'A^2H\eta = \lim_{n\to\infty}\lim_{p\to\infty}\frac{\mathrm{tr}\wedge\Delta\Delta'}{pa_{20}} .$$

Hence, the power of the T_2-test under local alternatives is given in the following theorem.

Theorem 4.3 Under the local alternatives $\eta = (nN)^{-1/2}\Delta$ with $\Delta = O(1)$, the power of the T_2-test is given by

$$\lim_{n\to\infty}\lim_{p\to\infty} P_1\left[\frac{T_2 - nq}{\sqrt{2nq}} > z_\alpha\right] = \lim_{n\to\infty}\lim_{p\to\infty}\Phi\left(-z_\alpha + \frac{\mathrm{tr}\wedge\Delta\Delta'}{pa_2\sqrt{2nq}}\right) .$$

Thus, when $\Sigma = I_p$, the asymptotic power of the T_2 test is given by

$$\beta(T_2) \simeq \Phi\left(-z_\alpha + \frac{\mathrm{tr}\Delta\Delta'}{p\sqrt{2nq}}\right) .$$

5 Power Comparison

We have shown that the power of the T_1 and T_2 tests depends on $\mathrm{tr}\delta\delta' = nN\mathrm{tr}\eta\eta'$ when $\Sigma = I$. Thus, in this case the $\tilde{T}_1$ test will always have a higher power than T_1 and T_2. However, the T_1-test is an asymptotic version of $\tilde{T}_1$ test, which implies that in this case T_2-test will be inferior than T_1. It also follows from the asymptotic powers, since

$$\sqrt{n}\sqrt{2npq} < \sqrt{p}\sqrt{2nqp}$$

for all $n \leq p$. Clearly, T_2-test should be only considered when $\Sigma \neq \sigma^2 I$. For general case T_2-test should be preferred over T_1 if

$$\left[pa_2\sqrt{2nq}\right]^{-1} \mathrm{tr} \wedge \Delta\Delta' > \mathrm{tr}\Delta\Delta'/n\sqrt{2npqa_2} \ ,$$

that is, if

$$\frac{\mathrm{tr} \wedge \Delta\Delta'}{\mathrm{tr}\Delta\Delta'} > (pa_2/n)^{\frac{1}{2}} \ . \tag{5.1 }$$

For example, if $\Delta = (\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_q)$ and $\boldsymbol{\delta}_i = (\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_p})'$, $i = 1, \ldots, q$, then (5.1) becomes

$$\left(\frac{a_2}{a_1^2}\right)^{\frac{1}{2}} > (p/n)^{\frac{1}{2}} \ ,$$

or

$$n < p(a_1^2/a_2) = pb \ ,$$

where $0 < b \leq 1$. Thus, for large p, and small n, the T_2-test appears to perform better. Fujikoshi et al. [5] have considered power comparison when $p/n \to c, 0 \leq c < 1$.

A Appendix

Lemma A.1 Let $V = YY' \sim W_p(\wedge, n)$, where the columns of Y are iid $N_p(\mathbf{0}, \wedge)$. Let $\ell_1, \ldots, \ell_n$ be the n non-zero eigenvalues of $V = H'LH$, $HH' = I_n$, $L = \mathrm{diag}(\ell_1, \ldots, \ell_n)$ and the eigenvalues of $W \sim W_n(I_n, p)$ are the diagonal elements of the diagonal matrix $D = \mathrm{diag}(d_1, \ldots, d_n)$. Then in probability

(a) $$\lim_{p\to\infty} \left(\frac{Y'Y}{p}\right) = \lim_{p\to\infty} \left(\frac{\mathrm{tr}\Sigma}{p}\right) I_n = a_{10} I_n \ ,$$

(b) $$\lim_{p\to\infty} \left(\frac{1}{p}L\right) = a_{10} I_n \ ,$$

(c) $$\lim_{p\to\infty} \left(\frac{1}{p}D\right) = I_n \ ,$$

(d) $$\lim_{p\to\infty} (H \wedge H') = (a_{20}/a_{10}) I_n \ ,$$

(e) $$\lim_{n\to\infty} \lim_{p\to\infty} \left(\frac{1}{n}\boldsymbol{a}'H'H\boldsymbol{a}\right) = \lim_{p\to\infty} \left(\frac{\boldsymbol{a}' \wedge \boldsymbol{a}}{pa_1}\right)$$

for a non-null vector $\boldsymbol{a} = (a_1, \ldots, a_p)'$ of constants.

For proof, see Srivastava [10].

ACKNOWLEDGEMENTS

This research was supported by Natural Sciences and Engineering Research Council of Canada.

References

[1] Bai, Z. and Saranadasa, H. (1996). Effect of high dimension: by an example of a two sample problem. *Statistica Sinica.,* **6**, 311-329.

[2] Dempster, A.P. (1958). A high dimensional two sample significance test. *Ann. Math. Stat.,* **29**, 995-1010.

[3] Dempster, A.P. (1960). A significance test for the separation of two highly multivariante small samples. *Biometrics,* **16**, 41-50.

[4] Eaton, M.L. (1983). *Multivariate Statistics: A Vector Space Approach.* Wiley, New York.

[5] Fujikoshi, Y., Himeno, T. and Wakaki, H. (2004). Asymptotic results of a high dimensional MANOVA test and power comparison when the dimension is large compared to the sample size. *J. Japan Statist. Soc.,* **34**, 19-26.

[6] Lehmann, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[7] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* Wiley, New York.

[8] Simaika, J. B. (1941). On an optimum property of two important statistical tests. *Biometrika,* **32**, 70-80.

[9] Srivastava, M.S. (2005). Some tests concerning the covariance matrix in high-dimensional data. *J. Japan Stat. Soc.,* **35**, To appear.

[10] Srivastava, M.S. (2004). Multivariate theory for analyzing high dimensional data. *Tech Report #0401,* University of Toronto.

[11] Srivastava, M.S. (2003). Singular Wishart and Multivariate beta distributions. *Ann. Statist.,* **31**, 1537-1560.

[12] Srivastava, M.S. (1971). On fixed-width confidence bounds for regression parameters. *Ann. Math. Statist.,* **42**, 1403-1411.

[13] Srivastava, M.S. and Khatri, C. G. (1979). *An Introduction to Multivariate Statistics.* North-Holland, New York.