CIRJE-F-89

Maximum Covariance Difference Test for Equality of Two Covariance Matrices

Akimichi Takemura
University of Tokyo

Satoshi Kuriki
The Institute of Statistical Mathematics

September 2000

Maximum Covariance Difference Test for Equality of Two Covariance Matrices

Akimichi Takemura and Satoshi Kuriki

Abstract. We propose a test of equality of two covariance matrices based on the maximum standardized difference of scalar covariances of two sample covariance matrices. We derive the tail probability of the asymptotic null distribution of the test statistic by the tube method. However the usual formal tube formula has to be suitably modified, because in this case the index set, around which the tube is formed, has zero critical radius.

1. Introduction

Consider the null hypothesis of equality of two covariance matrices

$$H_0 : \Sigma_1 = \Sigma_2$$

of two Wishart populations against the alternative $\Sigma_1 \neq \Sigma_2$. In this paper we propose a test statistic based on maximizing the standardized difference of sample scalar covariances from two populations. Let $p \times p$ ($p \geq 2$) random matrices W_1, W_2 be independently distributed according to Wishart distributions $W_p(n_i, \Sigma_i)$, $i = 1, 2$. Consider the difference of the scalar covariances $a'S_1b - a'S_2b$, where $S_i = W_i/n_i$, $i = 1, 2$, are sample covariance matrices and a, b are $p \times 1$ vectors. Under the null hypothesis $H_0 : \Sigma_1 = \Sigma_2 = \Sigma$, the variance of the difference for fixed a, b is written as

$$\mathrm{Var}(a'S_1b - a'S_2b) = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\left((a'\Sigma a)(b'\Sigma b) + (a'\Sigma b)^2\right). \tag{1.1}$$

See Appendix A. Therefore our proposed statistic is

$$T = \max_{a,b \in R^p} \frac{a'S_1b - a'S_2b}{\sqrt{(1/n_1 + 1/n_2)((a'Sa)(b'Sb) + (a'Sb)^2)}}, \tag{1.2}$$

where S is the pooled sample covariance matrix

$$S = \frac{W_1 + W_2}{n_1 + n_2}.$$

In the case that maximizing vectors a and b in (1.2) coincide, our proposed statistic detects the difference in the scalar variances of two populations. However as discussed in the next section, maximizing vectors a and b may be different. In this case our statistic detects the difference in scalar covariances. In this sense

1991 *Mathematics Subject Classification.* Primary 62H10; Secondary 62H15.

our statistic is different from Roy's maximum-minimum roots test ([**Ro**]), which compares only the scalar variances. In fact we shall show that with overwhelming probability a and b are different for large p (see Table 1 below) or large value of our statistic T (see Proposition 3.2). Therefore the behavior of our statistic is quite different from that of Roy's test. It is clearly different from omnibus type test procedures such as likelihood ratio test or Nagao's trace test ([**Na**]). Comparisons of various test statistics mainly from the viewpoint of power behavior have been given in [**PJ**], [**PA**], [**CP**] among others.

We have described our statistic in the setting of two-sample problem. One-sample version is even simpler. Let a $p \times p$ matrix $W = nS$ be distributed according to Wishart distribution $W_p(n, \Sigma)$. For testing

$$H_0 : \Sigma = I \quad \text{(the identity matrix)}$$

our test statistic is

$$T = \max_{a,b \in R^p} \frac{\sqrt{n} a'(S - I)b}{\sqrt{(a'a)(b'b) + (a'b)^2}}. \tag{1.3}$$

The advantage of maximum type test, including our test statistic and Roy's maximum-minimum roots test, is that when the null hypothesis is rejected, the test statistic itself suggests the direction of departure from the null hypothesis. On the other hand the drawback of the maximum type test is that the asymptotic null distribution is difficult to evaluate, because the limiting distribution is not a χ^2 distribution in general. Recently it has been recognized that the tube formula or the Euler characteristic method provide a general methodology for evaluating the tail probability of the asymptotic null distribution of maximum type statistic in various testing problems in classical multivariate analysis. The tube method can be applied to our proposed statistic as well.

However there is a substantial difficulty in formal application of the tube formula to the present problem. For the application of the tube formula we set up a Gaussian random field $Z(u) = \langle u, Z \rangle,\ u \in M$, (see (3.1) below) corresponding to the asymptotic null distribution of Wishart matrices. The index set M of the Gaussian field is a subset of the $p(p+1)/2 - 1$ dimensional unit sphere $S^{p(p+1)/2-1}$. It turns out that the critical radius of M for our present problem is zero and the formal expansion of the tail probability based on tube formula is only partly valid. In our previous work ([**TK2**]) we have investigated the validity of tube formula and found that the asymptotic expansion is valid up to certain degrees of freedom in the case of zero critical radius. Employing our previous argument it will be shown that for our present problem the asymptotic expansion is valid for $[p/2]$ terms, where $[\cdot]$ denotes the integer part.

The organization of this paper is as follows. In Section 2 we explicitly solve the maximization in (1.2) and (1.3) and give some intuitive meaning of our statistic. Section 3 is the main section of this paper and we investigate geometry of the index set M and evaluate coefficients in the tube formula for M. By studying singularities of M, we will show that the asymptotic expansion based on the tube formula is valid for $[p/2]$ terms. In Section 4 we give some simulation results to confirm the accuracy of approximation by tube formula. From the simulation we see that for the present problem the tube formula approximation is practical, but not remarkably good as in the cases with positive critical radius. Satisfactory approximation is achieved by explicit evaluation of the term of the order, where the formal tube formula breaks down. We also give some simulation results of power behavior of our statistic. Note that the theoretical investigation of the power behavior of our statistic is difficult, because at the present the tube formula is not applicable under contiguous alternative hypotheses.

2. Some simple interpretations of proposed statistic

In this section we explicitly solve the maximization in (1.2) and (1.3). This leads to a clear understanding of the meaning of our proposed test statistic. Throughout this paper for a real symmetric matrix A, we denote the positive (non-negative, non-positive, negative) definiteness by $A > 0$ ($A \geq 0$, $A \leq 0$, $A < 0$).

The maximization is entirely equivalent in (1.2) and (1.3). In (1.2) let

$$Z = \sqrt{\frac{n_1 n_2}{2(n_1+n_2)}} S^{-1/2}(S_1 - S_2)S^{-1/2} \tag{2.1}$$

and replace a by $S^{1/2}a$ and b by $S^{1/2}a$. Then the maximization in (1.2) reduces to the maximization in (1.3), where

$$Z = \sqrt{\frac{n}{2}}(S - I). \tag{2.2}$$

Therefore we consider the maximization problem

$$\max_{a,b \in R^p} \frac{\sqrt{2}a'Zb}{\sqrt{(a'a)(b'b) + (a'b)^2}},$$

where Z is a real symmetric matrix. It is convenient for later discussion to include the factor $1/\sqrt{2}$ in the definition of Z. Since the maximum is always non-negative, we can equivalently solve

$$\max_{a,b \in R^p} \frac{2(a'Zb)^2}{(a'a)(b'b) + (a'b)^2} = \max_{a,b \in R^p} \frac{2a'Zbb'Za}{a'((b'b)I + bb')a},$$

which is a ratio of quadratic forms in a for fixed b. Therefore for fixed b, the maximizing a is proportional to

$$(I + bb'/(b'b))^{-1}Zb = (I - bb'/(2b'b))Zb = Zb - \frac{b'Zb}{2b'b}b$$

and the maximum value is given by

$$\frac{2b'Z(I + bb'/(b'b))^{-1}Zb}{b'b} = \frac{2b'b \cdot b'Z^2b - (b'Zb)^2}{(b'b)^2}. \tag{2.3}$$

We now maximize this with respect to b under the restriction $b'b = 1$. Note that by diagonalizing Z, we can reduce the problem to the case of diagonal $Z = \mathrm{diag}(l_1, \ldots, l_p)$, $l_1 \geq \cdots \geq l_p$. Write $b = (b_1, \ldots, b_p)'$ and $x_i = b_i^2 \geq 0,\ i = 1, \ldots, p$, with $\sum_{i=1}^p x_i = 1$. Then (2.3) is written as

$$Q = 2\sum x_i l_i^2 - \left(\sum x_i l_i\right)^2.$$

Suppose that $l_1 \geq l_i > 0$. Fixing $x_j, j \neq 1, i$, consider changing x_1, x_i to $x_1 + c, x_i - c$, respectively. Differentiating Q by c we have

$$\frac{1}{2}\frac{\partial Q}{\partial c} = (l_1^2 - l_i^2) - \left(\sum x_j l_j\right)(l_1 - l_i) = (l_1 - l_i)\left(l_1 + l_i - \sum x_j l_j\right).$$

This is non-negative because $l_1 \geq \sum x_j l_j$. Therefore we increase Q by changing x_1, x_i to $x_1 + x_i, 0$, respectively. Similarly if $0 > l_i \geq l_p$ we increase Q by changing x_p, x_i to $x_p + x_i, 0$, respectively. It now follows that for positive definite Z we maximize Q by taking $x_1 = 1, x_2 = \cdots = x_p = 0$. Similarly for negative definite Z we maximize Q by $x_p = 1, x_1 = \cdots = x_{p-1} = 0$. It remains to consider the case $l_1 \geq 0 \geq l_p$. The above argument shows that at the maximum $x_2 = \cdots = x_{p-1} = 0$. Write $x_1 = c,\ x_p = 1 - c$. $\partial Q/\partial c = 2(l_1 - l_p)(l_1 + l_p - cl_1 - (1-c)l_p) = 0$ yields

$$x_1 = c = \frac{l_1}{l_1 - l_p}, \quad x_p = 1 - c = \frac{-l_p}{l_1 - l_p}.$$

a and b are given as

$$\begin{aligned} b &= \left(\frac{\sqrt{l_1}}{\sqrt{l_1 - l_p}}, 0, \ldots, 0, \pm \frac{\sqrt{-l_p}}{\sqrt{l_1 - l_p}} \right)', \\ a &= \left(\frac{\sqrt{l_1}}{\sqrt{l_1 - l_p}}, 0, \ldots, 0, \mp \frac{\sqrt{-l_p}}{\sqrt{l_1 - l_p}} \right)'. \end{aligned} \tag{2.4}$$

(see Appendix A for a.) Summarizing the above derivation in terms of original Z, we have proved the following theorem.

THEOREM 2.1. *Define Z by (2.1) for the two-sample problem or by (2.2) for the one-sample problem. Write the spectral decomposition of Z as $Z = \sum_{i=1}^p l_i h_i h_i'$, $l_1 \geq \cdots \geq l_p$. Then the value of the statistic T in (1.2) or (1.3) is given by*

$$T = \begin{cases} l_1, & \text{if} \quad Z > 0, \\ \sqrt{l_1^2 + l_p^2}, & \text{if} \quad l_1 \geq 0 \geq l_p, \\ -l_p, & \text{if} \quad Z < 0. \end{cases}$$

The set of maximizing vectors $\{S^{1/2}a, S^{1/2}b\}$ for the two-sample problem or $\{a, b\}$ for the one-sample problem is given by

$$\{S^{1/2}a, S^{1/2}b\} \text{ or } \{a, b\} = \begin{cases} \{h_1, h_1\}, & \text{if} \quad Z > 0, \\ \left\{ \frac{\sqrt{l_1}h_1 + \sqrt{-l_p}h_p}{\sqrt{l_1 - l_p}}, \frac{\sqrt{l_1}h_1 - \sqrt{-l_p}h_p}{\sqrt{l_1 - l_p}} \right\}, & \text{if} \quad l_1 \geq 0 \geq l_p, \\ \{h_p, -h_p\}, & \text{if} \quad Z < 0. \end{cases}$$

Note that for the case of non-negative definite or non-positive definite Z, $a = \pm b$ and our proposed statistic T corresponds to maximized difference in scalar variances. However for the case $l_1 > 0 > l_p$, a and b are not parallel and T corresponds to difference in scalar covariances. The situation is clearly understood by considering the simple case of one-sample problem with $p = 2$. In Figure 1 we have depicted the concentration ellipse $x'S^{-1}x = 1$ for 3 cases of Theorem 2.1. In Case 1, $S - I_2$ is positive definite. In this case $a = b$ and we are looking at the maximum variance. Similarly in Case 3, where $S - I_2$ is negative definite, we are looking at the minimum variance. Case 2 depicts the intermediate case. For simplicity consider

$$S = \begin{pmatrix} \mathrm{Var}(x_1) & \mathrm{Cov}(x_1, x_2) \\ \mathrm{Cov}(x_1, x_2) & \mathrm{Var}(x_2) \end{pmatrix} = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}, \quad r \neq 0.$$

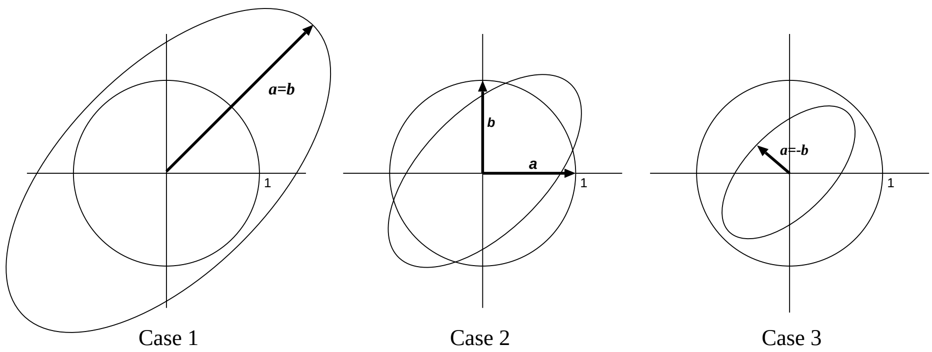

FIGURE 1. Three cases of maximization.

Then $a = (1,0)'$, $b = (0,1)'$ and

$$T = |\operatorname{Cov}(x_1, x_2)| = |r|.$$

In this case our statistic T detects the deviation of the covariance $\operatorname{Cov}(x_1, x_2)$ from 0.

3. Evaluation of tail probability of asymptotic null distribution by tube formula

In this section we evaluate the tail probability of the asymptotic null distribution of our statistic using the tube formula for general dimension. For the case $p = 2$, the exact form of the asymptotic null distribution is easy to evaluate and given in Appendix D.

Considering the asymptotic distribution of Z in (2.1) or (2.2), in this section let Z distributed according to the standard symmetric matrix normal distribution, i.e., the elements z_{ij}, $1 \le i \le j \le p$, of Z are independent normal random variables with mean 0 and variance

$$\operatorname{Var}(z_{ij}) = \begin{cases} 1, & \textit{if} \quad i = j, \\ 1/2, & \textit{if} \quad i < j. \end{cases}$$

Then by the invariance principle the asymptotic null distribution of T in (1.2) or (1.3) is written as

$$\lim_{n\to\infty} P(T > x) = P\left(\max_{a,b \in R^p} \frac{\sqrt{2}a'Zb}{\sqrt{(a'a)(b'b) + (a'b)^2}} > x \right),$$

where $n = \min(n_1, n_2)$ for the two-sample problem. Therefore the problem of evaluating the asymptotic null distribution is reduced to the evaluation of the distribution of the maximum of a Gaussian random field, for which the tube formula ([**Su**], see also [**KT1**], [**TK1**]) can be employed.

The sample space $\mathcal{S}$ of the standard symmetric matrix normal variable Z is the set of $p \times p$ real symmetric matrices with the inner product

$$\langle A, B \rangle = \operatorname{tr} AB, \quad A, B \in \mathcal{S}.$$

By the correspondence

$$A = (a_{ij}) \leftrightarrow (a_{11}, \ldots, a_{pp}, \sqrt{2}a_{12}, \ldots, \sqrt{2}a_{p-1,p})$$

$\mathcal{S}$ can be identified with $R^{p(p+1)/2}$ as is done in [**KT2**]. Let $K \subset \mathcal{S}$ denote the cone

$$K = \{ab' + ba' \mid a, b \in R^p\}$$

and define

$$M = K \cap S^{p(p+1)/2-1} = \left\{ \frac{ab' + ba'}{\|ab' + ba'\|} \mid a, b \in R^p \right\},$$

where

$$\|ab' + ba'\| = \langle ab' + ba', ab' + ba' \rangle^{1/2} = \sqrt{2((a'a)(b'b) + (a'b)^2)}.$$

Then

$$\max_{a,b \in R^p} \frac{\sqrt{2}a'Zb}{\sqrt{(a'a)(b'b) + (a'b)^2}} = \max_{u \in M} \langle u, Z \rangle \tag{3.1}$$

and our problem is reduced to the canonical form suitable for application of the tube formula.

We first determine the global geometry of K and M.

PROPOSITION 3.1. Let K_2 denote the set of $p \times p$ real symmetric matrices of rank less than or equal to 2. Let $\mathcal{S}_+$ and $\mathcal{S}_-$ denote the set of positive definite matrices and negative definite matrices, respectively. Then

$$K = K_2 \cap \mathcal{S}_+^C \cap \mathcal{S}_-^C,$$

where A^C denotes the complement of A.

PROOF. If $a \propto b$ then $ab' + ba' \propto aa'$ is of rank 0 or 1. If a and b are linearly independent then

$$ab' + ba' = \frac{1}{2}\left((a+b)(a+b)' - (a-b)(a-b)'\right).$$

Here $a + b$ and $a - b$ are linearly independent. By Sylvester's law $ab' + ba'$ has 1 positive and 1 negative characteristic root. Furthermore note that all matrices of rank 2 with 1 positive and 1 negative characteristic root can be written as $ab' + ba'$ again by Sylvester's law. This proves the proposition. □

From this proposition we see that $M = K \cap S^{p(p+1)/2-1}$ is a manifold with boundary and the boundary consists of matrices of rank 1:

$$\partial M = \{A \in M \mid \operatorname{rank} A = 1\} = \{hh' \mid h \in S^{p-1}\} \cup \{-hh' \mid h \in S^{p-1}\}.$$

Note that each of the two components of ∂M forms a manifold of dimension $p - 1$. We shall show that except for $p = 2$ the boundary of M has a singularity in the sense that at $u \in \partial M$ the tangent cone (supporting cone) $S_u(M)$ of M is not convex. Because of this singularity the critical radius of M is 0 for $p \geq 3$ as discussed in [**TK2**]. For the case $p = 2$, $\mathcal{S}$ is identified with R^3 and we can fully describe K and M. For illustrative purpose this is done in Appendix B. The relative interior M^O of M consists of matrices with one positive and one negative root and forms a manifold of dimension $2p - 2$. In the following we denote the smallest cone containing M^O and ∂K by $K^O = \bigcup_{c>0} cM^O$, $\partial K = \bigcup_{c>0} c\,\partial M$, respectively.

Let $\bar{G}_m$ denote the upper probability of χ^2 distribution with m degrees of freedom. From the standard argument of the tube formula, we know that the order

TABLE 1. Probability of positive definiteness of Z.

p	$P(Z>0)$
2	0.1464466
3	0.0249209
4	0.0024567
5	0.0001401
6	0.0000046
7	8.8×10^{-8}

of the tail probability of the projection of Z onto K^O, is of the order $O(\bar{G}_{2p-1})$, i.e.,

$$P\left(\sup_{u\in M^O}\langle u, Z\rangle > t\right) = P\left(\max_{u\in M}\langle u, Z\rangle > t\right) = O\big(\bar{G}_{2p-1}(t^2)\big) \quad \text{as} \quad t\to\infty. \tag{3.2}$$

Here $2p-1$ is the dimension of K^O. Similarly the order of the tail probability of the projection of Z onto ∂K is of the order $O(\bar{G}_p)$:

$$P\left(\max_{u\in\partial M}\langle u, Z\rangle > t\right) = O\big(\bar{G}_p(t^2)\big) = O\big(t^{-(p-1)}\bar{G}_{2p-1}(t^2)\big) \quad \text{as} \quad t\to\infty. \tag{3.3}$$

Note that (3.2) corresponds to the second case of Theorem 2.1 and (3.3) corresponds to the cases 1,3 of Theorem 2.1. Therefore simply by counting the dimensions of K^O and ∂K we have the following proposition.

PROPOSITION 3.2. *Under the asymptotic null distribution*

$$P(Z>0 \text{ or } Z<0 \mid T>t) = O\big(t^{-(p-1)}\big) \quad \text{as} \quad t\to\infty.$$

This proposition shows that as far as small P-values are concerned, the case of $a=\pm b$ in the maximization of (1.2) or (1.3) can be ignored. The statement of this proposition will be strengthened in Proposition 3.8

Actually the unconditional probability of the case $a=\pm b$ becomes very small as p becomes large. Table 1 lists the probabilities of Case 1, i.e., the probability of Z being positive definite, for dimensions $p=2,3,\ldots,7$. We see that for $p\geq 4$, the unconditional probability is negligible at the usual significance levels of 5% or 1%. The probabilities in Table 1 were calculated by a recurrence formula given in Lemma 3.5 below.

We now determine various differential geometric quantities of K and M. Actually most of the calculations have been done in [**KT2**]. The local geometry of ∂M is given by the following proposition. Since the two components of ∂M are entirely similar, we only consider the non-negative definite part of ∂M. Since the relation between the geometries of M and K is trivial we only state results in terms of K.

PROPOSITION 3.3. *Let $y=h_1h_1'\in\partial M$. Choose $h_2,\ldots,h_p$ such that $\{h_1,\ldots,h_p\}$ is an orthonormal basis of R^p. The tangent cone $S_y(K)$ of K at y is given by*

$$S_y(K) = \operatorname{span}\left\{h_1h_j' + h_jh_1',\ j=1,\ldots,p\right\} \oplus \{-qq' \mid h_1'q=0\},$$

where $\oplus$ denotes the orthogonal sum. The convex hull $C(S_y(K))$ of $S_y(K)$ is

$$C(S_y(K)) = \operatorname{span}\left\{h_1h_j' + h_jh_1',\ j=1,\ldots,p\right\} \oplus \{A\leq 0 \mid h_1'Ah_1=0\}$$

and the normal cone $N_y(K) = C(S_y(K))^*$ is

$$N_y(K) = C(S_y(K))^* = \{A \geq 0 \mid h_1' A h_1 = 0\}\,.$$

The non-zero characteristic roots of the second fundamental form of K at y with respect to $A \in N_y(K)$ is given by $\{-l_j,\ j = 2, \ldots, p\}$, where $l_2, \ldots, l_p$ are characteristic roots of A.

Proof. All the results are given in Section 2 of [**KT2**] except for the description of the tangent cone $S_y(K)$. Differentiating $y = h_1 h_1'$ with respect to h_1 gives the first term of $S_y(K)$. Now consider $x = h_1 h_1' - \epsilon q q' \in K^O$, $h_1' q = 0$, $\epsilon > 0$, in a neighborhood of y. This gives the second term of $S_y(K)$. $\square$

Note that at $y \in \partial M$, $S_y(K)$ and its convex hull $C(S_y(K))$ are different and $S_y(K)$ is not convex. As discussed in [**TK2**], this implies that the critical radius of M is zero and the tube formula applied to M^O is not entirely valid. We shall discuss this point below in more detail.

M^O and K^O are smooth and we summarize results on M^O and K^O from [**KT2**].

Proposition 3.4. Let $y \in M^O$ and write the spectral decomposition of y as $y = l_1 h_1 h_1' + l_p h_p h_p'$, $l_1 > 0 > l_p$, $l_1^2 + l_p^2 = 1$. Choose $h_2, \ldots, h_{p-1}$ such that $\{h_1, \ldots, h_p\}$ is an orthonormal basis of R^p. The tangent space of $S_y(K)$ of K at y is

$$S_y(K) = \operatorname{span}\left\{h_i h_j' + h_j h_i',\ i = 1, p,\ j = 1, 2, \ldots, p\right\}.$$

The normal space $N_y(K)$ of K at y is

$$N_y(K) = S_y(K)^* = \operatorname{span}\left\{h_i h_j' + h_j h_i',\ 2 \leq i \leq j \leq p-1\right\}.$$

The non-zero characteristic roots of the second fundamental form of K at y with respect to $A \in N_y(K)$ is given by

$$\left\{-\frac{l_j}{l_1}, -\frac{l_j}{l_p},\ j = 2, \ldots, p-1\right\},$$

where $l_2, \ldots, l_{p-1}$ are characteristic roots of A.

Given the above geometrical quantities we can now employ the tube formula for approximating the tail probability of asymptotic null distribution of our statistic T. Calculations of the tube formula for this case is very similar to those in [**KT2**] and in Section 3.2 of [**TK2**].

Let $l_1 \geq \cdots \geq l_p$ denote the characteristic roots of Z. The exact tail probability of T is given

$$\begin{aligned} P(T > t) &= 2P\left(l_1 > t, Z > 0\right) + P\left(l_1^2 + l_p^2 > t^2,\ l_1 > 0 > l_p\right) \\ &= 2F_1(t) + F_2(t) \quad \text{(say)}. \end{aligned} \tag{3.4}$$

The joint density function of $l_1, \ldots, l_p$ is given by

$$\begin{aligned} f(l_1, \ldots, l_p) &= d(p) \exp\left(-\frac{1}{2}\sum_{i=1}^p l_i^2\right) \prod_{1 \leq i < j \leq p} (l_i - l_j) \\ &= d(p) \exp\left(-\frac{1}{2}\sum_{i=1}^p l_i^2\right) \det\left(l_i^{p-j}\right)_{1 \leq i,j \leq p}\,, \end{aligned} \tag{3.5}$$

where

$$d(p) = \frac{1}{2^{p/2}\prod_{i=1}^{p}\Gamma(i/2)}. \tag{3.6}$$

Let

$$\begin{aligned} B_1(t) &= \{l_1 > t,\ \infty > l_2 > \cdots > l_p > 0\}, \\ B_2(t) &= \left\{l_1^2 + l_p^2 > t^2,\ l_1 > 0 > l_p,\ \infty > l_2 > \cdots > l_{p-1} > -\infty\right\}. \end{aligned}$$

Following the derivation in [**KT2**] and the example of Section 3.2 of [**TK2**], it is shown that the tube formula approximation to $F_1(t)$ and $F_2(t)$ is obtained by ignoring the order constraint $l_1 \geq l_2$ and $l_{p-1} \geq l_p$ in integrating the joint density:

$$\begin{aligned} P(T > t) \simeq \tilde{P}(T > t) &= \left\{2\int_{B_1(t)} + \int_{B_2(t)}\right\} f(l_1,\dots,l_p)\, dl_1 \dots dl_p \\ &= 2\,\tilde{F}_1(t) + \tilde{F}_2(t). \end{aligned} \tag{3.7}$$

Now we employ the techniques of [**Ku**] for evaluating $\tilde{F}_1(t)$ and $\tilde{F}_2(t)$. Define

$$U_k(q_1,\dots,q_k) = \underset{\infty > l_1 > \cdots > l_k > 0}{\int\cdots\int} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} l_i^2\right) \det\left(l_i^{q_j}\right)_{1\leq i,j\leq k} \prod_{i=1}^{k} dl_i, \tag{3.8}$$

$$V_k(q_1,\dots,q_k) = \underset{\infty > l_1 > \cdots > l_k > -\infty}{\int\cdots\int} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} l_i^2\right) \det\left(l_i^{q_j}\right)_{1\leq i,j\leq k} \prod_{i=1}^{k} dl_i, \tag{3.9}$$

where $q_1,\dots,q_k$ are non-negative integers. U_k was introduced in Section 2 of [**Ku**] and can be evaluated by the following recurrence formula.

LEMMA 3.5 (Theorem 2.2 of [**Ku**]). *$U_k(q_1,\dots,q_k)$ satisfies the following recurrence relation:*

$$\begin{aligned} &U_k(q_1,\dots,q_k) \\ &\quad = (-1)^{k-1} U_{k-1}(q_2,\dots,q_k) I(q_1 = 1) \\ &\qquad + (q_1 - 1) U_k(q_1 - 2, q_2,\dots,q_k) \\ &\qquad + 2\sum_{j=2}^{k} (-1)^j \frac{1}{2^{\frac{1}{2}(q_1+q_j)}} U_1(q_1 + q_j - 1) U_{k-2}(q_2,\dots,q_{j-1},q_{j+1},\dots,q_k) \end{aligned}$$

$(k \geq 2,\ q_1 \geq 1)$, *with the initial condition*

$$\begin{aligned} U_1(q) &= I(q = 1) + (q-1)U_1(q-2) \qquad (q \geq 1), \\ U_1(0) &= \sqrt{\pi/2}, \end{aligned}$$

where $I(\cdot)$ denotes the indicator function.

Similarly $V_k(q_1,\dots,q_k)$ can be evaluated by the the following recurrence formula. The proof is entirely the same as Theorem 2.2 of [**Ku**] and omitted.

LEMMA 3.6. *$V_k(q_1,\dots,q_k)$ satisfies the following recurrence relation:*

$$\begin{aligned} &V_k(q_1,\dots,q_k) \\ &\quad = (q_1 - 1) V_k(q_1 - 2, q_2,\dots,q_k) \\ &\qquad + 2\sum_{j=2}^{k} (-1)^j \frac{1}{2^{\frac{1}{2}(q_1+q_j)}} V_1(q_1 + q_j - 1) V_{k-2}(q_2,\dots,q_{j-1},q_{j+1},\dots,q_k) \end{aligned}$$

$(k \geq 2,\ q_1 \geq 1)$, *with the initial condition*

$$\begin{aligned} V_1(q) &= (q-1)V_1(q-2) \qquad (q \geq 1), \\ V_1(0) &= \sqrt{2\pi}. \end{aligned}$$

Using U_k and V_k the tube formula approximations $\tilde{F}_1(t), \tilde{F}_2(t)$ to $F_1(t), F_2(t)$ are evaluated as follows.

PROPOSITION 3.7. For $k = 1, \ldots, p$, let

$$\begin{aligned} \tau_{p-k+1} = d(p)(-1)^{k-1} 2^{(p-k-1)/2}\, \Gamma\left(\frac{p-k+1}{2}\right) \\ \times U_{p-1}(p-1, \ldots, p-k+1, p-k-1, \ldots, 0) \end{aligned}$$

and for $k = 3, 5, \ldots, 2p-1$, let

(3.10)

$$\begin{aligned} \omega_{2p-k+2} = d(p) \sum_{1 \leq l < m \leq p,\, l+m=k} & (-1)^{l-1} 2^{(2p-k)/2} \\ \times \Gamma\left(\frac{p-l+1}{2}\right) & \Gamma\left(\frac{p-m+1}{2}\right) \\ \times V_{p-2}(p-1, \ldots, p-l+1, & p-l-1, \ldots, p-m+1, p-m-1, \ldots, 0), \end{aligned}$$

where $d(p)$ is given in (3.6) and U_k, V_k are given in (3.8), (3.9). Then $\tilde{F}_1(t), \tilde{F}_2(t)$ in the formal tube formula approximation $\tilde{P}(T > t) = 2\tilde{F}_1(t) + \tilde{F}_2(t)$ to the tail probability $P(T > t)$ are written as

$$\begin{aligned} \tilde{F}_1(t) &= \sum_{k=1}^{p} \tau_{p-k+1} \bar{G}_{p-k+1}(t^2), \\ \tilde{F}_2(t) &= \sum_{k=3,\, k:\text{odd}}^{2p-1} \omega_{2p-k+2}\, \bar{G}_{2p-k+2}(t^2). \end{aligned}$$

PROOF. Consider the determinant term $\det(l_i^{p-j})_{1 \leq i,j \leq p}$ in the joint density (3.5) of the characteristic roots of Z. For $\tilde{F}_1$ expand this determinant as

$$\begin{aligned} \det\bigl(l_i^{p-j}\bigr)_{1 \leq i,j \leq p} &= \sum_{k=1}^{p} (-1)^{k-1} l_1^{p-k} \det \begin{pmatrix} l_2^{p-1} & \cdots & l_2^{p-k+1} & l_2^{p-k-1} & \cdots & 1 \\ \vdots & & \vdots & \vdots & & \vdots \\ l_p^{p-1} & \cdots & l_p^{p-k+1} & l_p^{p-k-1} & \cdots & 1 \end{pmatrix} \\ &= \sum_{k=1}^{p} (-1)^{k-1} l_1^{p-k} \det\bigl(l_{i+1}^{p-j-I(k \leq j)}\bigr)_{1 \leq i,j \leq p-1}. \end{aligned}$$

Now l_1 and $l_2, \ldots, l_p$ can be separately integrated out. Integration with respect to l_1 gives

$$\int_t^\infty l_1^{p-k} e^{-l_1^2/2}\, dl_1 = 2^{(p-k-1)/2}\, \Gamma\left(\frac{p-k+1}{2}\right) \bar{G}_{p-k+1}(t^2).$$

Integration with respect to $l_2, \ldots, l_p$ gives U_{p-1}. This proves the expansion of $\tilde{F}_1$.

For $\tilde{F}_2$, expand the determinant $\det(l_i^{p-j})_{1\le i,j\le p}$ as

$$\det\bigl(l_i^{p-j}\bigr)_{1\le i,j\le p} = \sum_{1\le l<m\le p} (-1)^{l+m+p-1} \times \det\begin{pmatrix} l_1^{p-l} & l_1^{p-m} \\ l_p^{p-l} & l_p^{p-m}\end{pmatrix} \det\bigl(l_{i+1}^{p-j-I(l\le j)-I(m-1\le j)}\bigr)_{1\le i,j\le p-2}.$$

Now l_1, l_p and $l_2,\dots,l_{p-1}$ can be separately integrated out. Integration with respect to l_1, l_p gives

$$\int_{l_1^2+l_p^2>t^2,\, l_1>0>l_p} \left(l_1^{p-l} l_p^{p-m} - l_1^{p-m} l_p^{p-l}\right) e^{-(l_1^2+l_p^2)/2}\, dl_1 dl_p$$
$$= \{(-1)^{p-m} - (-1)^{p-l}\}\, 2^{(2p-l-m-2)/2} \times \Gamma\left(\frac{p-l+1}{2}\right)\Gamma\left(\frac{p-m+1}{2}\right)\bar{G}_{2p-l-m+2}(t^2).$$

Note that

$$(-1)^{p-m} - (-1)^{p-l} = \begin{cases} 0 & \textit{if } m+l \textit{ is even}, \\ 2(-1)^{p-m} & \textit{if } m+l \textit{ is odd}. \end{cases}$$

Integration with respect to $l_2,\dots,l_{p-1}$ gives V_{p-2}. Letting $k=m+l$ and rearranging terms according to the value of k gives the expansion of $\tilde{F}_2$. □

In Appendix C we list $\tilde{F}_1(t), \tilde{F}_2(t)$ for $p=2,\dots,5$.

It is of considerable interest to explicitly evaluate the leading term of the tube formula approximations for general p, because the leading term of the tube formula is always valid as shown in [**TK2**].

PROPOSITION 3.8. The order of $F_1(t)$ and $F_2(t)$ (as $t\to\infty$) in (3.4) is the same as the order of $\tilde{F}_1(t)$ and $\tilde{F}_2(t)$, respectively, and given by

$$F_1(t) \sim 2^{(p-3)/2}\, P(Z_{p-1}>0)\, \bar{G}_p(t^2), \tag{3.11}$$

$$F_2(t) \sim 2^{p-5/2}\, \bar{G}_{2p-1}(t^2), \tag{3.12}$$

where $P(Z_{p-1}>0)$ denotes the probability that $(p-1)\times(p-1)$ symmetric matrix normal random variable is positive definite.

PROOF. The leading term of $\tilde{F}_1(t)$ or $F_1(t)$ is

$$d(p)\, 2^{(p-2)/2}\, \Gamma(p/2)\, U_{p-1}(p-2,\dots,0)\, \bar{G}_p(t^2).$$

Now $U_{p-1}(p-2,\dots,0) = P(Z_{p-1}>0)/d(p-1)$, $d(p)/d(p-1) = 1/\{2^{1/2}\,\Gamma(p/2)\}$. This proves (3.11). Next, the leading term of $\tilde{F}_2(t)$ or $F_2(t)$ is

$$d(p)\, 2^{(2p-3)/2}\, \Gamma(p/2)\, \Gamma((p-1)/2)\, V_{p-2}(p-3,\dots,0)\, \bar{G}_{2p-1}(t^2).$$

Now $V_{p-2}(p-3,\dots,0) = 1/d(p-2)$, $d(p)/d(p-2) = 1/\{2\,\Gamma(p/2)\Gamma((p-1)/2)\}$. This proves (3.12). □

Proposition 3.8 provides a more precise statement of Proposition 3.2.

If M has positive critical radius, then the tube formula approximation is valid in the sense that in (3.7), $P(T>t) = \tilde{P}(T>t)(1+R(t))$, where the remainder term $R(t)$ is of exponentially small order in t as $t\to\infty$. An exponential bound for the remainder term $R(t)$ in terms of the critical radius was given in [**KT1**].

However in our case M has zero critical radius and the tube formula is only partly valid. From [**TK2**] we see that the terms of the tube formula $\tilde{F}_2$ for K^O is valid for the degrees of freedom greater than $p = \dim \partial K$. In other words, the error $F_2(t) - \tilde{F}_2(t)$ is of the same order as the order of $F_1(t)$ or $\tilde{F}_1(t)$. Therefore if we use the tube formula approximation $\tilde{F}_2(t)$ for $F_2(t)$, then $F_1(t)$ or $\tilde{F}_1(t)$ is of no use. We state this fact as the following theorem.

THEOREM 3.9. *Under the asymptotic null distribution, the tail probability of our statistic T of (1.2) or (1.3) is evaluated as*

$$P(T > t) = \begin{cases} \omega_{2p-1}\bar{G}_{2p-1}(t^2) + \omega_{2p-3}\bar{G}_{2p-3}(t^2) + \cdots + \omega_{p+2}\bar{G}_{p+2}(t^2) \\ \qquad\qquad +R_p(t), \qquad p : \text{odd}, \\ \omega_{2p-1}\bar{G}_{2p-1}(t^2) + \omega_{2p-3}\bar{G}_{2p-3}(t^2) + \cdots + \omega_{p+1}\bar{G}_{p+1}(t^2) \\ \qquad\qquad +R_p(t), \qquad p : \text{even}, \end{cases}$$

where ω_m is defined in (3.10), the order of the remainder term $R_p(t)$ is $O(\bar{G}_p(t^2))$.

It is of interest to investigate the term of order $O(\bar{G}_p(t^2))$ in $F_2(t) - \tilde{F}_2(t)$. Since the formal tube formula breaks down at this order, we need to examine the multiple integral in $F_2(t)$ and $\tilde{F}_2(t)$ much more closely.

For $p \geq 3$ define

$$\eta_p = -d(p)\, 2^{p/2}\, \Gamma\left(\frac{p}{2}\right) \int_{\substack{0 > l_p > l_{p-1} \\ \infty > l_2 > \cdots > l_{p-1}}} \exp\left(-\frac{1}{2}\sum_{i=2}^{p-1} l_i^2\right) \prod_{2 \leq i < j \leq p} (l_i - l_j) \prod_{i=2}^{p} dl_i. \tag{3.13}$$

Then we have the following proposition.

PROPOSITION 3.10. Let $p \geq 3$. As $t \to \infty$

$$F_2(t) - \tilde{F}_2(t) = \eta_p\, \bar{G}_p(t^2) + o\big(\bar{G}_p(t^2)\big), \tag{3.14}$$

where $d(p)$ is given in (3.6).

We give an outline of a proof of Proposition 3.10 in Appendix E. Combining (3.14) with the leading term of $F_1(t)$ we can strengthen Theorem 3.9 as follows.

THEOREM 3.11. *The term of order $\bar{G}_p(t^2)$ in the remainder term $R_p(t)$ of Theorem 3.9 is evaluated as*

(3.15)

$$R_p(t) = \begin{cases} \{\omega_p + \eta_p + 2^{(p-1)/2}\, P(Z_{p-1} > 0)\}\bar{G}_p(t^2) + o\big(\bar{G}_p(t^2)\big), & p : \text{odd}, \\ \{\eta_p + 2^{(p-1)/2}\, P(Z_{p-1} > 0)\}\bar{G}_p(t^2) + o\big(\bar{G}_p(t^2)\big), & p : \text{even}, \end{cases}$$

where ω_m is defined in (3.10).

This theorem recovers the term of order $\bar{G}_p(t^2)$ in the formal tube formula approximation for $P(T > t)$ in Theorem 3.9. However from a practical viewpoint there is a difficulty in applying (3.15) for large p, because exact evaluation of the constant η_p for large p seems to be difficult. For $p = 3, 4, 5$, the values of η_p in (3.13) are evaluated as

$$\eta_3 = \frac{1}{2\sqrt{2}}, \qquad \eta_4 = \frac{3}{2} + \frac{4}{3\pi}, \qquad \eta_5 = \frac{135}{16\sqrt{2}} + \frac{25}{8\pi} - \frac{81}{8\sqrt{2}\pi}\tan^{-1}\left(\frac{1}{\sqrt{2}}\right).$$

Evaluation of η_5 is already quite laborious.

Table 2. Limiting power behavior of three tests.

$p = 4$, Level of significance = 5%

$(\psi_1,\psi_2,\psi_3,\psi_4)$	LR	Roy	TK
(2, 2, 0, 0)	0.435	0.436	0.323
(2, 0, 0, –2)	0.435	0.352	0.451
(1.5, 1.5, –1.5, –1.5)	0.490	0.332	0.446
(2, 0, –1, –1)	0.323	0.261	0.332

4. Some simulation results

In this section we present some simulation results concerning our proposed statistic.

First we investigate accuracy of our tail probability approximation to the asymptotic null distribution by simulation. We do not consider approximation of (finite degrees of freedom) Wishart distribution by the symmetric matrix normal distribution, because this approximation is not specific to our statistic.

Figure 2 shows the tube formula approximation for the dimensions $p = 3, 4, 5$. In Figure 2 "simulated" (solid line) shows the simulated true tail probability of our statistic based on simulation of size 100,000. "main term" (dotted line) is the approximation using only the main term (3.12). "Thm3.9" and "Thm3.11" (dashed lines) are based on the approximation in Theorems 3.9 and 3.11. Note that for $p = 3$ the approximation in Theorem 3.9 consists of only the main term. From the figures we see that the approximation by Theorem 3.9 is practical for P-values below 5% range but is not as good as our previous studies given in [**KT1**]. This suggests that the case of zero critical radius poses some difficulty from the numerical viewpoint as well. On the other hand the addition of the d.f. p term in Theorem 3.11 greatly improves the approximation and is very satisfactory.

We now very briefly investigate the power behavior of our statistic. We only compare the first order limiting power behavior of our statistic against the likelihood ratio test and Roy's maximum-minimum roots test. For more extensive power comparison of existing test procedures against two-sided alternatives see [**CP**]. [**PJ**] and [**PA**] give detailed power comparisons of existing tests against one-sided alternatives.

As the limit of contiguous alternatives, let Z be distributed according to the symmetric matrix normal distribution with non-zero mean matrix Ψ. The covariance structure of Z is the same as the null case. In the first order, the likelihood ratio test and other omnibus type test procedures are equivalent and have non-central χ^2 distribution with $p(p+1)/2$ degrees of freedom and non-centrality parameter $\mathrm{tr}\,\Psi^2$. Roy's maximum-minimum roots test is $T_R = \max(l_1, -l_p)$. The significance level is taken to be 5%. We obtain the upper 5 percentile of Roy's maximum-minimum roots test by generating 100,000 T_R's under the asymptotic null distribution. The upper 5 percentile of our statistic was already obtained by the simulation study of the previous paragraph. Then the power is computed by generating T_R and our static 100,000 times each under the limiting alternative distribution and by counting the number of times these statistics exceed the 5 percentiles.

Note that Ψ can be assumed to be diagonal $\mathrm{diag}(\psi_1, \ldots, \psi_p)$ without loss of generality. We only present the results for the case of $p = 4$. In Table 2, TK stands

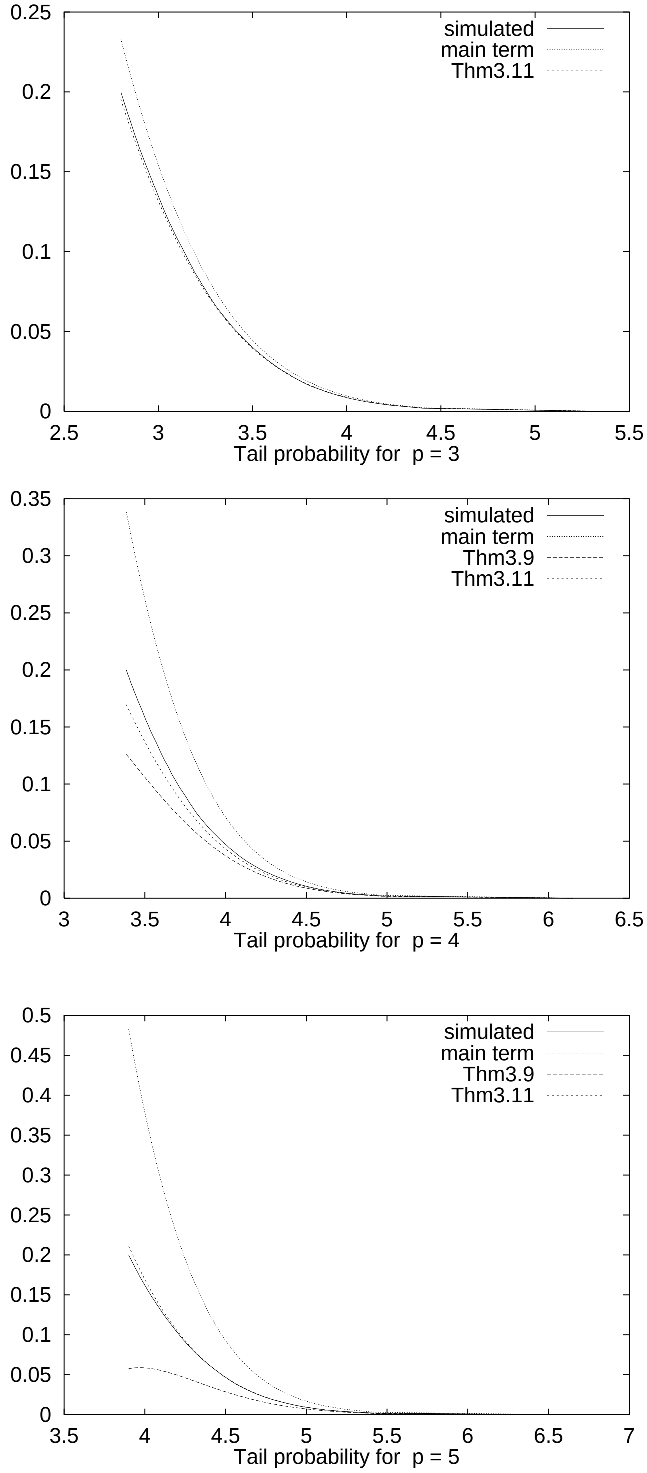

FIGURE 2. Tube formula approximation for $p = 3, 4, 5$.

for our proposed statistic. The entries of the table are the values of the asymptotic power.

From Table 2 our statistic tends to do better than other procedures when Ψ has one large root and one small (i.e., negative but large in absolute value) root. On the other hand Roy's test is good when Ψ is positive definite or negative definite. The likelihood ratio test seems to be a reasonable overall test.

Appendix A. Derivation of some formulae

Derivation of (1.1).

Note that $\mathrm{Var}(a'S_1b - a'S_2b) = \mathrm{Var}(a'W_1b)/n_1^2 + \mathrm{Var}(a'W_2b)/n_2^2$. Let $W = (w_{ij})$ be distributed according to Wishart distribution $W_p(n, \Sigma)$, $\Sigma = (\sigma_{ij})$. It suffices to show that $\mathrm{Var}(a'Wb) = n(a'\Sigma a)(b'\Sigma b) + n(a'\Sigma b)^2$. Using the fact that $\mathrm{Cov}(w_{ij}, w_{kl}) = n(\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk})$ we have

$$
\begin{aligned}
\mathrm{Var}(a'Wb) &= \mathrm{Cov}\Big(\sum_{i,j} a_i b_j w_{ij}, \sum_{k,l} a_k b_l w_{kl}\Big) \\
&= n \sum_{i,j,k,l} a_i a_k b_j b_l (\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}) \\
&= n\Big(\sum_{i,k} a_i a_k \sigma_{ik}\Big)\Big(\sum_{j,l} b_j b_l \sigma_{jl}\Big) + n\Big(\sum_{i,l} a_i b_l \sigma_{il}\Big)\Big(\sum_{j,k} b_j a_k \sigma_{jk}\Big) \\
&= n(a'\Sigma a)(b'\Sigma b) + n(a'\Sigma b)^2.
\end{aligned}
$$

Derivation of (2.4).

The factor $1/\sqrt{l_1 - l_p}$ is only for normalization. Therefore we can take

$$b = (\sqrt{l_1}, 0, \ldots, 0, \pm\sqrt{-l_p})'.$$

Then

$$
\begin{aligned}
a &= Zb - \frac{b'Zb}{2b'b} b \\
&= (l_1\sqrt{l_1}, 0, \ldots, 0, \pm l_p\sqrt{-l_p})' - \frac{l_1^2 - l_p^2}{2(l_1 - l_p)}(\sqrt{l_1}, 0, \ldots, 0, \pm\sqrt{-l_p})' \\
&= \frac{l_1 - l_p}{2}(\sqrt{l_1}, 0, \ldots, 0, \mp\sqrt{-l_p})'.
\end{aligned}
$$

Normalizing this we obtain a in (2.4).

Appendix B. The cone K for $p = 2$

For $p = 2$ the sample space $\mathcal{S}$ is identified with R^3 and almost all $Z \in \mathcal{S}$ has rank 2. Therefore we only need to identify $\mathcal{S}_+$ and $\mathcal{S}_-$. For real symmetric $Z = (z_{ij})_{1\le i,j\le 2}$ write $u = z_{11}, v = z_{22}, w = \sqrt{2}z_{12}$. Then $Z > 0$ if and only if

$$u > 0, \quad 2uv - w^2 > 0.$$

If we make 45 degrees rotation in (u, v) plane and define $s = (u + v)/\sqrt{2}$, $t = (u - v)/\sqrt{2}$. Then $2uv = s^2 - t^2$ and

$$\mathcal{S}_+ = \{(s, t, w) \mid s + t > 0,\ s^2 > t^2 + w^2\}$$

is a circular cone. $\mathcal{S}_- = -\mathcal{S}_+$ is the circular cone opposite to $\mathcal{S}_+$. Therefore K is the cone obtained by subtracting two circular cones from R^3. We see that for $p=2$, $M = K \cap S^2$ is a 2-dimensional manifold with smooth boundaries.

Appendix C. Explicit form of tube formula approximations for small dimensions

Here we give a brief list of tube formula approximations of Proposition 3.7. These following formulas were obtained by the applying the recurrence relations of Lemmas 3.5 and 3.6.

$p=2:$

$$\begin{aligned}
\tilde{F}_1 &= \frac{1}{2\sqrt{2}}(\bar{G}_2 - \bar{G}_1), \\
\tilde{F}_2 &= \frac{1}{\sqrt{2}}\bar{G}_3.
\end{aligned}$$

$p=3:$

$$\begin{aligned}
\tilde{F}_1 &= \left(-\frac{1}{2\sqrt{2}} + \frac{1}{2}\right)\bar{G}_3 - \frac{1}{\sqrt{2\pi}}\bar{G}_2 + \left(\frac{1}{2\sqrt{2}} + \frac{1}{4}\right)\bar{G}_1, \\
\tilde{F}_2 &= \sqrt{2}(\bar{G}_5 - \bar{G}_3).
\end{aligned}$$

$p=4:$

$$\begin{aligned}
\tilde{F}_1 &= \left(\frac{1}{2\sqrt{2}} - \frac{1}{\pi}\right)\bar{G}_4 + \left(\frac{1}{4\sqrt{2}} - \frac{1}{4}\right)\bar{G}_3 + \left(-\frac{3}{8\sqrt{2}} + \frac{1}{\pi}\right)\bar{G}_2 \\
&\quad + \left(-\frac{3}{8\sqrt{2}} + \frac{1}{4}\right)\bar{G}_1, \\
\tilde{F}_2 &= 2\sqrt{2}\bar{G}_7 - \frac{7}{\sqrt{2}}\bar{G}_5 + \frac{7}{2\sqrt{2}}\bar{G}_3.
\end{aligned}$$

$p=5:$

$$\begin{aligned}
\tilde{F}_1 &= \left(-\frac{1}{4\sqrt{2}} - \frac{1}{\pi} + \frac{1}{2}\right)\bar{G}_5 + \left(-\frac{1}{\sqrt{2\pi}} + \frac{2}{3\pi}\right)\bar{G}_4 + \left(\frac{1}{8\sqrt{2}} + \frac{4}{3\pi} - \frac{1}{2}\right)\bar{G}_3 \\
&\quad + \left(\frac{11}{12\sqrt{2\pi}} - \frac{2}{3\pi}\right)\bar{G}_2 + \left(\frac{1}{8\sqrt{2}} - \frac{2}{3\pi} + \frac{1}{8}\right)\bar{G}_1, \\
\tilde{F}_2 &= 4\sqrt{2}\bar{G}_9 - 9\sqrt{2}\bar{G}_7 + \frac{15}{\sqrt{2}}\bar{G}_5 - \frac{5}{\sqrt{2}}\bar{G}_3.
\end{aligned}$$

Appendix D. Exact asymptotic tail probability for $p = 2$

For the case $p=2$, we can easily evaluate the exact tail probability of the asymptotic null distribution. Note that for $p=2$, $\langle Z, Z\rangle = l_1^2 + l_2^2$ and the event $\{l_1 > 0 > l_2\}$ are independent. Furthermore $P(l_1 > 0 > l_2) = 1/\sqrt{2}$. Therefore for $p=2$

$$F_2(t) = \frac{1}{\sqrt{2}}\bar{G}_3(t^2).$$

Now we evaluate $F_1(t) = \int_{l_1 > t,\, l_1 > l_2 > 0} f(l_1, l_2)\, dl_1 dl_2$. Integrating with respect to l_2 first then with respect to l_1 gives

$$\int_0^{l_1} f(l_1, l_2)\, dl_2 = \frac{l_1}{\sqrt{2}} e^{-l_1^2/2}\left(\Phi(l_1) - \frac{1}{2}\right) - \frac{1}{2\sqrt{\pi}} e^{-l_1^2/2} + \frac{1}{2\sqrt{\pi}} e^{-l_1^2},$$

$$\int_{l_1=t}^{\infty}\int_{l_2=0}^{l_1} f(l_1,l_2)\,dl_2dl_1 = \frac{e^{-t^2/2}}{2\sqrt{2}} - \frac{1-\Phi(t)}{\sqrt{2}} - \frac{1-\Phi(t)}{\sqrt{2}}e^{-t^2/2} + (1-\Phi(\sqrt{2}t)),$$

where Φ is the cumulative distribution function of the standard normal distribution. Noting that $\bar{G}_2(t^2) = e^{-t^2/2}$, $\bar{G}_1(t^2) = 2(1-\Phi(t))$ we obtain

$$P(T>t) = \frac{1}{\sqrt{2}}\bar{G}_3(t^2) + \frac{1}{\sqrt{2}}\bar{G}_2(t^2) - \frac{1}{\sqrt{2}}\bar{G}_1(t^2) - \frac{1}{\sqrt{2}}\bar{G}_2(t^2)\bar{G}_1(t^2) + \bar{G}_1(2t^2).$$

Note that this exact formula coincides with the tube formula $\tilde{F}_2+2\tilde{F}_1$ for $p=2$ given in Appendix C, except for remainder terms which is negligible with exponentially small order. This should be the case because M has positive critical radius for $p=2$.

Appendix E. Outline of a Proof of Proposition 3.10

Here we give an outline of a proof of Proposition 3.10. Since full justification of all the steps of the approximations takes too much space, we omit some justifications of neglecting remainder terms. A complete memo of justifications can be obtained from the authors upon request.

Note that the difference between $\tilde{F}_2$ and F_2 is due to the difference of the ranges of integration and we can write

$$\text{(E.1)}\quad \frac{1}{d(p)}\left(\tilde{F}_2(t) - F_2(t)\right) = \int\limits_{\substack{l_1>0>l_p\\ l_2>\cdots>l_{p-1}\\ l_1^2+l_p^2>t^2\\ l_2>l_1 \text{ or } l_{p-1}<l_p}} \exp\left(-\frac{1}{2}\sum_{i=1}^p l_i^2\right)\det\left(l_i^{p-j}\right)\prod_{i=1}^p dl_i$$

$$= \int\limits_{\substack{l_1>0>l_p\\ l_2>\cdots>l_{p-1}\\ l_1^2+l_p^2>t^2\\ l_2>l_1}} + \int\limits_{\substack{l_1>0>l_p\\ l_2>\cdots>l_{p-1}\\ l_1^2+l_p^2>t^2\\ l_{p-1}<l_p}} - \int\limits_{\substack{l_1>0>l_p\\ l_2>\cdots>l_{p-1}\\ l_1^2+l_p^2>t^2\\ l_2>l_1 \text{ and } l_{p-1}<l_p}} .$$

It can be easily shown that the third term on the right hand side of (E.1) is negligible with exponentially small order. Intuitively this is the case because the probability of both l_2 and l_{p-1} being greater than t in absolute value should be exponentially smaller than the probability that just one of them is greater than t in absolute value. By symmetry the first term and the second term on the right hand side are equal. Therefore

$$\frac{1}{d(p)}\left(\tilde{F}_2(t) - F_2(t)\right) \sim 2\int\limits_{\substack{l_1>0>l_p\\ l_2>\cdots>l_{p-1}\\ l_1^2+l_p^2>t^2\\ l_{p-1}<l_p}} \exp\left(-\frac{1}{2}\sum_{i=1}^p l_i^2\right)\det\left(l_i^{p-j}\right)\prod_{i=1}^p dl_i.$$

We now argue that the main contribution to this integral comes from the region where $|l_p|/l_1$ is close to 0. In fact consider the range $-\sqrt{t} > l_p$. Fix $0<\epsilon<1$. Since $l_{p-1}^2 > l_p^2 > t$ on the range of integration we have

$$\exp\left(-\frac{1}{2}l_{p-1}^2\right) \le \exp\left(-\frac{1-\epsilon}{2}t\right)\exp\left(-\frac{\epsilon}{2}l_{p-1}^2\right).$$

From this it easily follows that there exists C_ϵ such that

$$\int\limits_{\substack{l_1>0,-\sqrt{t}>l_p\\ l_2>\cdots>l_{p-1}\\ l_1^2+l_p^2>t^2\\ l_{p-1}<l_p}} \exp\left(-\frac{1}{2}\sum_{i=1}^p l_i^2\right)\left|\det\left(l_i^{p-j}\right)\right|\prod_{i=1}^p dl_i \le C_\epsilon \exp\left(-\frac{1-\epsilon}{2}t\right)\bar{G}_1(t^2)$$

for all $t > 0$. Therefore this range of integration is negligible with exponentially small order. Similarly we can show that the range $l_i^2 > t^2,\ i = 2,\dots,p-1$, can be ignored.

It follows that

$$\frac{1}{d(p)}\left(\tilde{F}_2(t) - F_2(t)\right) \sim 2 \int\limits_{\substack{l_1>0>l_p\ge-\sqrt{t}\\ \sqrt{t}\ge l_2>\cdots>l_{p-1}\\ l_1^2+l_p^2>t^2\\ -\sqrt{t}\le l_{p-1}<l_p}} \exp\left(-\frac{1}{2}\sum_{i=1}^p l_i^2\right)\det\left(l_i^{p-j}\right)\prod_{i=1}^p dl_i.$$

Note that on this range of integration l_1 is large:

$$l_1 > \sqrt{t^2 - l_p^2} \ge \sqrt{t^2 - t}.$$

This implies that on our range of integration the main term of $\det(l_i^{p-j})$ is the term with the highest degree in l_1, i.e.,

$$\det\left(l_i^{p-j}\right) \sim l_1^{p-1}\prod_{2\le i<j\le p}(l_i - l_j).$$

Now we integrate l_1 out. Using

$$\int_{\sqrt{t^2-l_p^2}}^\infty l_1^{p-1}\exp\left(-\frac{1}{2}l_1^2\right)dl_1 \sim t^{p-2}\exp\left(-\frac{t^2}{2}\right)\exp\left(\frac{l_p^2}{2}\right)$$

we have

$$\frac{1}{d(p)}\left(\tilde{F}_2(t) - F_2(t)\right) \sim 2t^{p-2}\exp\left(-\frac{t^2}{2}\right)$$
$$\times \int\limits_{\substack{0>l_p\ge-\sqrt{t}\\ \sqrt{t}\ge l_2>\cdots>l_{p-1}\\ -\sqrt{t}\le l_{p-1}<l_p}} \exp\left(-\frac{1}{2}\sum_{i=2}^{p-1} l_i^2\right)\prod_{2\le i<j\le p}(l_i - l_j)\prod_{i=2}^p dl_i.$$

Proposition 3.10 follows from this because $2t^{p-2}e^{-t^2/2} \sim \Gamma(p/2)2^{p/2}\bar{G}_p(t^2)$.

References

[CP] S. S. Chu and K. C. S. Pillai, Power comparison of two-sided tests of equality of two covariance matrices based on six criteria, *Ann. Inst. Stat. Math.* **31** (1979), 182–205.

[Ku] S. Kuriki, One-sided test for the equality of two covariance matrices, *Ann. Statist.* **21** (1993), 1379–1384.

[KT1] S. Kuriki and A. Takemura, Tail probabilities of the maxima of multilinear forms and their applications, Discussion Paper CIRJE-F-4, Faculty of Economics, Univ. of Tokyo (1998). Submitted for publication.

[KT2] ———, Some geometry of the cone of nonnegative definite matrices and weights of associated $\bar{\chi}^2$ distribution, *Ann. Inst. Stat. Math.* **52** (2000), 1–14.

[Na] H. Nagao, On some test criteria for covariance matrix, *Ann. Statist.* **1** (1973), 700–709.

[PA] K. C. S. Pillai and S. Al-Ani, Power comparisons of tests of equality of two covariance matrices based on individual characteristic roots, *J. Amer. Statist. Assoc.* **65** (1970), 438–446.

[PJ] K. C. S. Pillai and K. Jayachandran, Power comparisons of tests of equality of two covariance matrices based on four criteria, *Biometrika* **55** (1968), 335–342.

[Ro] S. N. Roy, On a heuristic method of test construction and its use in multivariate analysis, *Ann. Math. Statist.* **24** (1953), 220–238.

[Su] J. Sun, Tail probabilities of the maxima of Gaussian random fields, *Ann. Probab.* **21** (1993), 34–71.

[TK1] A. Takemura and S. Kuriki, Maximum of Gaussian field on piecewise smooth domain: Equivalence of tube method and Euler characteristic method, Discussion Paper CIRJE-F-54, Faculty of Economics, Univ. of Tokyo (1999). Submitted for publication.

[TK2] ———, Tail probability via tube formula and Euler characteristic method when critical radius is zero, Discussion Paper CIRJE-F-59, Faculty of Economics, Univ. of Tokyo (1999). Submitted for publication.

Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

E-mail address: `takemura@e.u-tokyo.ac.jp`

The Institute of Statistical Mathematics, 4-6-7 Minami-azabu, Minato-ku, Tokyo 106-8569, Japan.

E-mail address: `kuriki@ism.ac.jp`