CIRJE-F-4

Tail Probabilities of the Maxima of Multilinear Forms and Their Applications

Satoshi Kuriki
The Institute of Statistical Mathematics

Akimichi Takemura
The University of Tokyo

June 1998

Tail probabilities of the maxima of multilinear forms and their applications

Satoshi Kuriki
The Institute of Statistical Mathematics
and
Akimichi Takemura
Faculty of Economics, University of Tokyo

June, 1998

Abstract

Let Z be a k-way array whose $q_1 \times \cdots \times q_k$ elements are independent standard normal variables. For q_i-dimensional vector h_i, $i = 1, \ldots, k$, define a multilinear form of degree k by $(h_1 \otimes \cdots \otimes h_k)' \mathrm{vec}(Z)$. We derive formulas for upper tail probabilities of the maximum of multilinear form with respect to h_i's under the condition $\|h_i\| = 1$ for any i, and of its standardized statistic obtained by dividing by $\|\mathrm{vec}(Z)\|$. We also give formulas for the maximum of symmetric multilinear form $(h_1 \otimes \cdots \otimes h_k)' \mathrm{vec}(\mathrm{sym}(Z))$, where $\mathrm{sym}(Z)$ denotes the symmetrization of Z with respect to indices. These classes of statistics have important applications in testing hypotheses of multivariate analysis such as the analysis of variance of multiway layout data or testing multivariate normality. In order to derive the tail probabilities we employ a geometric approach developed by H. Weyl and J. Sun. Upper and lower bounds for the tail probabilities are given by reexamining the Sun's results. Some numerical examples are given to illustrate the practical usefulness of the obtained formulas including the upper and lower bounds.

Key words: Gaussian field, Johnson-Graybill statistic, Karhunen-Loève expansion, largest eigenvalue, Malkovich-Afifi statistic, multivariate normality, multiway layout, projection pursuit, second fundamental form, volume element, Weyl's tube formula, Wishart distribution.

Abbreviated Title: The maxima of multilinear forms

Affiliation and Address:
Satoshi Kuriki, The Institute of Statistical Mathematics, 4-6-7 Minami-azabu, Minato-ku, Tokyo 106-8569, Japan. Email: kuriki@ism.ac.jp
Akimichi Takemura, Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Email: takemura@e.u-tokyo.ac.jp

1 Introduction

Let $Z = (z_{j_1 \cdots j_k})$, $j_i = 1, \ldots, q_i$, $i = 1, \ldots, k$, be a k-way random array whose components are distributed independently according to the standard normal distribution $N(0,1)$. Let $h_i = (h_{i1}, \ldots, h_{iq_i})' \in R^{q_i}$, $i = 1, \ldots, k$, be coefficient vectors and consider a multilinear form of degree k, or k-form, defined by

$$g_k(h_1, \ldots, h_k; Z) = \sum_{j_1=1}^{q_1} \cdots \sum_{j_k=1}^{q_k} h_{1j_1} \cdots h_{kj_k} z_{j_1 \ldots j_k}.$$

In matrix notation

$$g_k(h_1, \ldots, h_k; Z) = (h_1 \otimes \cdots \otimes h_k)' z,$$

where $\otimes$ denotes the Kronecker product and

$$z = \mathrm{vec}(Z) = (z_{11\ldots 1}, z_{11\ldots 2}, \ldots, z_{q_1 q_2 \ldots q_k})'$$

is the $(\prod_{i=1}^k q_i)$-dimensional column vector consisting of the components of Z by the lexicographic ordering. We consider the maximum of the k-form under the condition $\|h_i\| = 1$ for any i, i.e.,

$$T_k = \max_{\|h_i\|=1,\ \forall i} g_k(h_1, \ldots, h_k; Z), \tag{1.1}$$

and its standardized statistic

$$U_k = T_k / \|z\|. \tag{1.2}$$

Here $\| \cdot \|$ denotes the usual Euclidean norm. Note that $T_k \geq 0$ and $0 \leq U_k \leq 1$ since $\|h_1 \otimes \cdots \otimes h_k\| = \prod_i \|h_i\| = 1$.

By imposing the additional condition that $q_1 = \cdots = q_k$ ($= q$, say), we also consider the symmetric multilinear form of degree k, or symmetric k-form, $g_k(h_1, \ldots, h_k; \mathrm{sym}(Z))$. Here $\mathrm{sym}(Z)$ is the k-way array with $(j_1, \ldots, j_k)$-th component

$$\frac{1}{k!} \sum_{\pi \in P(\{1,\ldots,k\})} z_{j_{\pi(1)} \cdots j_{\pi(k)}},$$

and $P(\{1, \ldots, k\})$ denotes the set of permutations of $\{1, \ldots, k\}$. The corresponding maxima are

$$\tilde{T}_k = \max_{\|h_i\|=1,\ \forall i} g_k(h_1, \ldots, h_k; \mathrm{sym}(Z)), \tag{1.3}$$

and

$$\tilde{U}_k = \tilde{T}_k / \|z\|. \tag{1.4}$$

It can be easily proved that the maximum in (1.3) is attained when $h_1 = \cdots = h_k$ ($= h$, say) holds. Therefore we can write

$$\tilde{T}_k = \max_{\|h\|=1} g_k(h, \ldots, h; \mathrm{sym}(Z)) = \max_{\|h\|=1} \tilde{g}_k(h; Z),$$

where $h = (h_1, \ldots, h_q)' \in R^q$ and

$$\tilde{g}_k(h; Z) = (\underbrace{h \otimes \cdots \otimes h}_{k})' z.$$

Note that $\tilde{T}_k \geq 0$ and $0 \leq \tilde{U}_k \leq 1$ for k odd, and that $-1 \leq \tilde{U}_k \leq 1$ for k even. The primary purpose of this paper is to give some explicit formulas for the tail probabilities for T_k, U_k, $\tilde{T}_k$, and $\tilde{U}_k$. More precisely, we shall give asymptotic series for $P(T_k \geq a)$ and $P(\tilde{T}_k \geq a)$ when a is large, and expressions for $P(U_k \geq a)$ and $P(\tilde{U}_k \geq a)$ for a greater than a suitable constant.

When $k = 2$ the k-way array becomes a $q_1 \times q_2$ random matrix $Z = (z_{j_1 j_2})$, and $T_2 = \max h_1' Z h_2$ is the largest singular value of Z. Therefore T_2^2 is the largest eigenvalue $\lambda_1(ZZ')$ of the $q_1 \times q_1$ matrix ZZ', or $\lambda_1(Z'Z)$ of the $q_2 \times q_2$ matrix $Z'Z$, where ZZ' or $Z'Z$ are distributed according to the Wishart distributions $W_{q_1}(q_2, I_{q_1})$ or $W_{q_2}(q_1, I_{q_2})$, and I_d denotes the identity matrix of order d. The distribution of the largest eigenvalue of the Wishart matrix was extensively studied because of its both practical and theoretical importance. When the expectation parameter matrix is the identity (i.e., the null case), the distribution of the largest eigenvalue can be obtained in principle by integrating out the other eigenvalues in the joint density of eigenvalues (e.g., Chapter 13 of Anderson (1984)). Along this line, some algorithms for evaluating the distribution function were devised. See a survey paper by Pillai (1976). Although this approach enables us to numerically evaluate the distribution functions, it does not yield explicit formula for the marginal distribution of the largest eigenvalue unless $\min(q_1, q_2)$ is small.

The distribution of $U_2^2 = \lambda_1(ZZ')/\mathrm{tr}(ZZ') = \lambda_1(Z'Z)/\mathrm{tr}(Z'Z)$, the largest eigenvalue divided by the trace of the same Wishart matrix, has an important application in the analysis of variance. In the analysis of two-way layout data without replication, Johnson and Graybill (1972) proposed a test statistic for interaction effects, where the null distribution coincides with that of U_2^2. Similarly when we extend their method to three-way layout, or multiway layout of higher order, the distribution of U_k^2, $k \geq 3$, are needed. We summarize these applications in the analysis of multiway layout data in Section 2.1. Davis (1972) gave an algorithm to evaluate the distribution function for U_2^2 numerically, as well as the explicit expressions for $\min(q_1, q_2) = 2, 3$. Using the method by Davis (1972), Schuurmann et al. (1973) provided a table of quantiles.

The maximum of the symmetric 2-form $\tilde{T}_2$ is the largest eigenvalue of the symmetric matrix

$$A = \mathrm{sym}(Z) = (Z + Z')/2, \tag{1.5}$$

whose each diagonal element and upper off-diagonal element are independently distributed as the normal distributions $N(0,1)$ and $N(0, 1/2)$, respectively. It is to be noted that the distribution of A is multivariate symmetric normal distribution, which is the limiting distribution of standardized Wishart matrix as degrees of freedom go to infinity. In Section

2.2 we show that the statistics $\tilde{T}_3$ and $\tilde{T}_4$ arise as the limits of the test statistics for multivariate normality proposed by Malkovich and Afifi (1973).

In order to derive the tail probabilities of the maxima introduced above, we employ a geometric approach. Around sixty years ago, motivated by the work by Hotelling (1939), Weyl (1939) defined the tube in the Euclidean space as well as the unit sphere of general dimension, and derived a formula for the volume of tube. For the history and applications to statistics, see Knowles and Siegmund (1989). More recently, Sun (1993) has developed a general theory of the tail probability of the maximum of Gaussian random field with finite Karhunen-Loève expansion. Sun's theory states that the tail probability is expressed in terms of the geometric quantities which appear as the coefficients of Weyl's tube formula for a manifold defined by the Karhunen-Loève expansion. As we shall see later, evaluation of the tail probabilities for U_k and $\tilde{U}_k$ can be reduced to the evaluation of the volume of tubes. Derivation of the tail probabilities for T_k and $\tilde{T}_k$ are within the scope of Sun (1993). However it is in general difficult to determine the coefficients in Weyl's tube formula. For example, although Sun (1991) discussed the tail probability of pursuit index in exploratory projection pursuit, she could not evaluate the corresponding geometric characteristics explicitly. In our paper, we elucidate the differential geometric structure of corresponding manifolds and determine the geometric characteristics in explicit forms.

Outline of this paper is as follows. In Section 2, applications of the distributions of the maxima introduced in Section 1 are explained. In the testing problems described in Sections 2.1 and 2.2, the distributions for U_k and $\tilde{T}_k$ are required to calculate p-values, respectively. In Section 3, we prepare geometric tools. In our recent paper, Takemura and Kuriki (1997) have developed the distribution theory for the projection onto convex cone. The maximum of Gaussian random field with finite Karhunen-Loève expansion can be regarded as the projection onto nonconvex smooth cone. We summarize the theory by Weyl (1939) and Sun (1993) in a form comparable to Takemura and Kuriki (1997). We also give a theorem to calculate the critical radius, the extreme radius for which Weyl's tube formula is valid. Sections 4 and 5 are devoted to the statistics T_k, U_k, $\tilde{T}_k$, and $\tilde{U}_k$. We elucidate the geometric structures of corresponding manifolds, determine the geometric quantities, and then obtain the formula for the tail probabilities. By giving some numerical examples we also demonstrate that the obtained expressions are practical enough for calculating p-values.

2 Applications to testing hypotheses

In this section we discuss testing problems where the distributions of the maxima introduced in Section 1 are required in calculating their p-values.

2.1 Tests for interaction in multiway layout without replication

Let x_{ij}, $i = 1, \ldots, I$, $j = 1, \ldots, J$, be observed as two-way layout data without replication. For such data Johnson and Graybill (1972) assumed a model:

$$x_{ij} = \alpha_i + \beta_j + g u_i v_j + \varepsilon_{ij}, \tag{2.1}$$

where α_i, β_j, g, u_i, and v_j are unknown parameters and ε_{ij} is a random error distributed independently as $N(0, \sigma^2)$ with σ^2 unknown. They proposed a test for interaction effects, or non-additivity, as a likelihood ratio test for testing $H_0 : g = 0$. They showed that the critical region of the likelihood ratio test is given by

$$\lambda_1(YY')/\mathrm{tr}(YY') > c \tag{2.2}$$

for some constant c, where $Y = (y_{ij})$ is a $I \times J$ matrix with (i, j)-th element

$$y_{ij} = x_{ij} - x_{i\cdot} - x_{\cdot j} + x_{\cdot\cdot}\ ,$$

and $\lambda_1(YY')$ is the largest eigenvalue of YY'. Here the dot means the arithmetic mean with respect to the corresponding subscript, e.g., $x_{i\cdot} = (1/J)\sum_{j=1}^{J} x_{ij}$. Under the null hypothesis $H_0 : g = 0$, the distribution of the likelihood ratio test statistic in (2.2) is shown to be that of U_2^2 in (1.2) with $q_1 = I - 1$, $q_2 = J - 1$.

As an extension of Johnson and Graybill (1972), Kawasaki and Miyakawa (1996) considered the following model in the analysis of three-way layout without replication:

$$x_{ijk} = (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + g u_i v_j w_k + \varepsilon_{ijk},$$

where ε_{ijk} is distributed independently as $N(0, \sigma^2)$, $i = 1, \ldots, I$, $j = 1, \ldots, J$, $k = 1, \ldots, K$. Here as in (2.1) the parameters $(\alpha\beta)_{ij}$, $(\alpha\gamma)_{ik}$, $(\beta\gamma)_{jk}$, g, u_i, v_j, w_k, and σ^2 are unknown. Using this model they proposed a test for the null hypothesis $H_0 : g = 0$. The critical region of the likelihood ratio test is of the form

$$\max_{\|u\|=\|v\|=\|w\|=1} \left(\sum\nolimits_{i,j,k} u_i v_j w_k y_{ijk}\right)^2 \Big/ \sum\nolimits_{i,j,k} y_{ijk}^2 > c, \tag{2.3}$$

where $u = (u_1, \ldots, u_I)'$, $v = (v_1, \ldots, v_J)'$, and $w = (w_1, \ldots, w_K)'$ are unit vectors, and

$$y_{ijk} = x_{ijk} - x_{ij\cdot} - x_{i\cdot k} - x_{\cdot jk} + x_{i\cdot\cdot} + x_{\cdot j\cdot} + x_{\cdot\cdot k} - x_{\cdot\cdot\cdot}$$

is the residual under H_0. The distribution of the test statistic in (2.3) under H_0 is shown to be that of U_3^2 in (1.2) with $q_1 = I - 1$, $q_2 = J - 1$, $q_3 = K - 1$.

In a similar fashion, one can extend this method to multiway layout of higher order. For the k-way layout data without replication we can propose the likelihood ratio test statistics as in (2.3), whose distribution under H_0 is that of U_k^2.

2.2 Tests for multivariate normality

Let $x_1, \ldots, x_n \in R^q$ be n independently and identically distributed random vectors. Define the third and fourth sample cumulants with respect to the direction $u \in R^q$, $u \neq 0$, by

$$k_3(u) = \frac{(1/n)\sum_{i=1}^n (u'x_i - u'\bar{x})^3}{(u'Su)^{3/2}}$$

and

$$k_4(u) = \frac{(1/n)\sum_{i=1}^n (u'x_i - u'\bar{x})^4}{(u'Su)^2} - 3,$$

where $\bar{x} = (1/n)\sum_{i=1}^n x_i$ and $S = (1/n)\sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})'$. Motivated by Roy's union intersection principle, Malkovich and Afifi (1973) proposed tests for multivariate normality. The proposed test statistics are

$$b_3 = \max_u k_3(u), \quad b_4^+ = \max_u k_4(u), \quad \text{and} \quad b_4^- = \min_u k_4(u).$$

The hypothesis is rejected when b_3, b_4^+, or $-b_4^-$ are greater than some critical points. Let $y_i(u)$ $(i = 3, 4)$ be real-valued continuous Gaussian fields on S^{q-1} (the unit sphere in R^q) having the mean zero and the covariance function $E[y_i(u)y_i(v)] = 6(u'v)^3$ $(i = 3)$, $24(u'v)^4$ $(i = 4)$. Machado (1983) and Baringhaus and Henze (1991) proved that, under the null hypothesis that the distribution of x_i is q-variate normal, $\sqrt{n}b_3$ converges in distribution to $\max_{u \in S^{q-1}} y_3(u)$, and both $\sqrt{n}b_4^+$ and $-\sqrt{n}b_4^-$ converge in distribution to the common limit $\max_{u \in S^{q-1}} y_4(u)$, as n goes to infinity.

Let z be a q^3 or q^4-dimensional random column vector whose each component is distributed independently as $N(0,1)$. Then it can easily be seen that the Gaussian fields $y_3(u)$ and $y_4(u)$ have representations

$$y_3(u) = \sqrt{6}(u \otimes u \otimes u)'z$$

and

$$y_4(u) = \sqrt{24}(u \otimes u \otimes u \otimes u)'z.$$

This means that $\sqrt{n}b_3$ converges in distribution to $\sqrt{6}\,\tilde{T}_3$, and $\sqrt{n}b_4^+$ and $-\sqrt{n}b_4^-$ converge in distribution to $\sqrt{24}\,\tilde{T}_4$.

3 Distribution of the projection onto nonconvex smooth cone

In this section we summarize geometric tools mainly from Weyl (1939), Sun (1993), and Johansen and Johnstone (1990) in a form suitable for our development. Furthermore by reexamining Sun's derivation of the asymptotic expansion of the tail probability, we give upper and lower bounds for the tail probability $P(T \geq a)$ for the non-standardized maximum (such as T_k or $\tilde{T}_k$ in (1.1) or (1.3)), which are valid for each $a > 0$. We provide our own simplified proofs of these results in Appendix B.

Let $\{x(t) \in R \mid t \in I\}$ be a Gaussian random field such that $E[x(t)] = 0,\ E[x(t)^2] = 1$ with the index set I. We assume that $x(t)$ has a finite Karhunen-Loève expansion:

$$x(t) = \sum_{i=1}^{p} \phi_i(t) z_i = \phi(t)' z, \quad t \in I, \tag{3.1}$$

where $\phi(t) = (\phi_1(t), \dots, \phi_p(t))'$, $z = (z_1, \dots, z_p)'$ and z_i, $i = 1, \dots, p$, are independent standard normal random variables. Note that $E[x(s)x(t)] = \phi(s)'\phi(t)$, and that $\|\phi(t)\| = 1$ since $E[x(t)^2] = 1$. Let

$$M = \phi(I) = \{\phi(t) \mid t \in I\} \subset S^{p-1}.$$

We put some assumptions on M.

Assumption 3.1 *M is a non-bordered compact C^2-submanifold of dimension d in S^{p-1}.*

Define a closed cone $K \subset R^p$ associated with M by

$$K = \bigcup_{c \geq 0} cM = \{c\phi(t) \mid c \geq 0,\ t \in I\}, \tag{3.2}$$

which is smooth except for the origin. For $x \in R^p$ let $x_K \in K$ denote the projection of x onto K:

$$\|x - x_K\| = \min_{y \in K} \|x - y\|.$$

Then

$$\max_{t \in I} x(t) = \max_{u \in M} u'z = \|x_K\|$$

unless $\|x_K\| = 0$. Note that x_K exists since K is closed. x_K may not be unique but $\|x_K\|$ and $\|x - x_K\|$ are uniquely determined. In Takemura and Kuriki (1997) we investigated properties of projection onto a convex cone K. In the case of the convex cone x_K is always uniquely determined and its distribution is nicely characterized as $\bar{\chi}^2$ distribution. By introducing a cone K in (3.2) it becomes clear that the results in this section are closely related to those in Takemura and Kuriki (1997).

For nonconvex K we need to be concerned with the uniqueness of projection x_K. The essential notions are the tube around M and critical radius (critical angle) of M with respect to the geodesic distance of S^{p-1}. Here the geodesic distance between two points $u, v \in S^{p-1}$ is given by $\arccos(u'v)$, which is the lengths of the part of the great circle joining u and v.

For $0 < \theta < \pi$ the tube of geodesic distance θ around M on S^{p-1} is defined by

$$M_\theta = \{v \in S^{p-1} \mid \max_{u \in M} u'v > \cos\theta\}.$$

For each $u \in M$ let $T_u(M)$ denote the tangent space of M at u and $T_u(M)^\perp$ denote its orthogonal complement in R^p. Define a set $C_\theta(u) \subset S^{p-1}$ by

$$C_\theta(u) = \{v \in S^{p-1} \mid u'v > \cos\theta\} \cap T_u(M)^\perp.$$

$C_\theta(u)$ is the set of points v with the geodesic distance less than θ from u and such that the geodesic from u to v is orthogonal to $T_u(M)$ at u. Since M is a closed non-bordered submanifold of S^{p-1} we obviously have

$$M_\theta = \bigcup_{u \in M} C_\theta(u).$$

It is said that M_θ does not have self-overlap if $C_\theta(u)$, $u \in M$, are disjoint. The supremum θ_c of θ for which M_θ does not have self-overlap is called the critical radius (or critical angle) of M:

$$\theta_c = \sup\{\theta \mid M_\theta \text{ does not have self-overlap}\}.$$

Note that the critical radius never exceeds $\pi/2$, which is attained when $M = S^{d'-1} \subset S^{p-1}$, $d' < p$.

For determining the critical radius of M the following lemma (Proposition 4.3 of Johansen and Johnstone (1990)) is very useful. Although Johansen and Johnstone (1990) stated their Proposition 4.3 for the case $\dim M = 1$ only, its statement and proof hold for $\dim M = d > 1$ almost verbatim and we omit the proof.

Lemma 3.1 *The critical radius θ_c of M is given by*

$$\cot^2\theta_c = \sup_{u,v \in M} \frac{1 - u'P_v u}{(1 - u'v)^2} \tag{3.3}$$

where P_v is the orthogonal projection onto the tangent space $T_v(K)$ of K of (3.2) at v.

Remark 3.1 *Let*

$$h(u,v) = \frac{\sqrt{1 - u'P_v u}}{1 - u'v} \tag{3.4}$$

be the square root of the argument of the supremum in (3.3). In Appendix A we show that $h(u,v)$ can be defined also for $u = v$ by taking appropriate supremum as $u \to v$, and the maximum over the compact set $M \times M$ exists and is finite. This implies that the critical radius θ_c is positive under our Assumption 3.1.

Let K_θ denote the cone associated with M_θ:

$$K_\theta = \bigcup_{c \geq 0} cM_\theta.$$

As before K denotes the cone associated with M. If $x \in K_{\theta_c}$ then the projection x_K of x onto K is unique. For $x \in K_{\theta_c}$ write

$$x = x_K + (x - x_K) = ru + sv,$$

where $r = \|x_K\|$, $s = \|x - x_K\|$, and

$$u = x_K/r \in M, \qquad v = (x - x_K)/s \in T_u(K)^{\perp} \cap S^{p-1}.$$

The one-to-one correspondence

$$x \leftrightarrow (r, u, s, v)$$

is of class C^1 and Weyl (1939) derived its Jacobian. We state the Jacobian in the following lemma.

Lemma 3.2 *Let $H(u,v)$ denote the second fundamental form of K at u with respect to the direction $v \in T_u(K)^{\perp} \cap S^{p-1}$. Then*

$$dx = \left| I_{d+1} + \frac{s}{r} H(u,v) \right| r^d dr\, du\, s^{p-d-2} ds\, dv \tag{3.5}$$

where du denotes the volume element of M and dv denotes the volume element of $T_u(K)^{\perp} \cap S^{p-1}$ (the $(p-d-2)$-dimensional unit sphere restricted to the space $T_u(K)^{\perp}$).

A simple proof of this Lemma 3.2 is given in Appendix A of Kuriki and Takemura (1997).

Let $\mathrm{tr}_j H$ denote the j-th trace, i.e., the j-th elementary symmetric function of the eigenvalues of $H = H(u,v)$. Let $\mathrm{tr}_0 H \equiv 1$. Although $T_u(K)$ is of dimension $d+1$, $\mathrm{rank}\, H(u,v) \le d$ since $H(u,v)$ has at least one eigenvalue (principal curvature) equal to 0 with the eigenvector (principal direction) u. Therefore

$$\left| I_{d+1} + \frac{s}{r} H(u,v) \right| r^d = \sum_{e=0}^{d} r^{d-e} s^e \, \mathrm{tr}_e H$$

and (3.5) can alternatively be written as

$$dx = \sum_{e=0}^{d} r^{d-e} s^{p-d-2+e} dr\, ds\, \mathrm{tr}_e H(u,v)\, du\, dv. \tag{3.6}$$

Moreover as shall be explained in Appendix A, the principal curvatures of K at u with respect to the principal directions orthogonal to u coincide with the principal curvatures of M at u. In other words $H(u,v)$ appearing in (3.5) and (3.6) can be replaced with the second fundamental form of M at u with respect to v.

From Lemma 3.2 the volume of M_θ, $\theta \le \theta_c$, is obtained as follows. Let

$$\Omega_d = \mathrm{Vol}(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma(d/2)}$$

denote the total volume of S^{d-1} and let $\bar{B}_{m,n}(a)$ denote the upper tail probability of the beta distribution with parameter (m,n)

$$\bar{B}_{m,n}(a) = \int_a^1 \frac{1}{B(m,n)} \xi^{m-1} (1-\xi)^{n-1} d\xi.$$

Lemma 3.3 *Let* $z \in R^p$ *be distributed according to the standard multivariate normal distribution* $N_p(0, I_p)$ *. For* $0 \le \theta \le \theta_c$

$$\mathrm{Vol}(M_\theta) = \Omega_p \cdot P(z \in K_\theta) = \Omega_p \sum_{\substack{e=0 \\ e:\text{even}}}^{d} w_{d+1-e} \bar{B}_{\frac{1}{2}(d+1-e),\frac{1}{2}(p-d-1+e)}(\cos^2\theta),$$

where

$$w_{d+1-e} = \frac{1}{\Omega_{d+1-e}\Omega_{p-d-1+e}} \int_M \Big[\int_{T_u(K)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv \Big]\, du. \tag{3.7}$$

This formula was given by Weyl (1939). A simple proof is given in Appendix B. Note that w_{d+1-e} corresponds to the weight of $\bar{\chi}^2$ distribution for piecewise smooth cone given in Theorem 2.4 of Takemura and Kuriki (1997).

Now consider the tail probability of standardized maximum statistic. Let $x(t)$ be given as in (3.1) and consider

$$U = \max_{t \in I} \phi(t)'z/\|z\| = \max_{u \in M} u'z/\|z\|. \tag{3.8}$$

Because $z/\|z\|$ has the uniform distribution over S^{p-1}, for $-1 \le a \le 1$

$$P(U \ge a) = \frac{1}{\Omega_p} \mathrm{Vol}(M_\theta), \quad \theta = \theta(a) = \arccos(a).$$

If $a \ge \cos\theta_c$ then $\mathrm{Vol}(M_{\theta(a)})$ is given by Lemma 3.3. For convenience we state this as a lemma.

Lemma 3.4 *For* $a \ge \cos\theta_c$

$$P(U \ge a) = \sum_{\substack{e=0 \\ e:\text{even}}}^{d} w_{d+1-e} \bar{B}_{\frac{1}{2}(d+1-e),\frac{1}{2}(p-d-1+e)}(a^2).$$

Now we consider the non-standardized statistic. Let

$$T = \max_{t \in I} \phi(t)'z = \max_{u \in M} u'z. \tag{3.9}$$

Denote the density and the upper tail probability of χ^2 distribution with m degrees of freedom by $g_m(a)$ and $\bar{G}_m(a)$, respectively. Furthermore for $a, b > 0$ define

$$Q_{m,n}(a,b) = \int_a^\infty g_m(\xi)\,(1 - \bar{G}_n(b\xi))\, d\xi = \bar{G}_m(a) - \int_a^\infty g_m(\xi)\,\bar{G}_n(b\xi)\, d\xi.$$

$Q_{m,n}(a,b)$ can be evaluated by numerical integration. It is also easy to obtain recurrence relations among $Q_{m,n}(a,b)$'s.

Now we can state the following theorem.

Theorem 3.1 *Let* w_{d+1-e} *be given in (3.7). For* $a > 0$

$$Q_L(a) \le P(T \ge a) \le Q_U(a)$$

where

$$Q_L(a) = \sum_{\substack{e=0 \\ e:\text{even}}}^{d} w_{d+1-e} Q_{d+1-e,p-d-1+e}(a^2, \tan^2\theta_c) \tag{3.10}$$

and

$$Q_U(a) = Q_L(a) + \bar{G}_p(a^2(1+\tan^2\theta_c))\Big(1 - \frac{\mathrm{Vol}(M_{\theta_c})}{\Omega_p}\Big). \tag{3.11}$$

Proof is given in Appendix B. Furthermore it is easy to see that

$$Q_U(a) - Q_L(a) \le \bar{G}_p(a^2(1+\tan^2\theta_c)) = o(\bar{G}_1(a^2))$$

and

$$\begin{aligned} \Big|Q_L(a) - \sum_{\substack{e=0 \\ e:\text{even}}}^{d} w_{d+1-e}\bar{G}_{d+1-e}(a^2)\Big| &\le \sum_{\substack{e=0 \\ e:\text{even}}}^{d} |w_{d+1-e}| \int_{a^2}^{\infty} g_{d+1-e}(\xi)\,\bar{G}_{p-d-1+e}(\xi\tan^2\theta_c)\,d\xi \\ &\le \sum_{\substack{e=0 \\ e:\text{even}}}^{d} |w_{d+1-e}|\,\bar{G}_p(a^2(1+\tan^2\theta_c)) = o(\bar{G}_1(a^2)). \end{aligned}$$

As a corollary to Theorem 3.1 we have the following result by Sun (1993):

Corollary 3.1

$$P(T \ge a) = \sum_{\substack{e=0 \\ e:\text{even}}}^{d} w_{d+1-e}\bar{G}_{d+1-e}(a^2) + o(\bar{G}_1(a^2)) \qquad as\ a \to \infty\,. \tag{3.12}$$

Remark 3.2 *Let* $\mathrm{lin}\,K$ *be the intersection of all linear subspaces containing the cone* K. *When* $\mathrm{lin}\,K$ *is a proper subset of* R^p, *there exists a Karhunen-Loève expansion of dimension* $p' = \dim(\mathrm{lin}\,K) < p$, *and* p *in Theorem 3.1 should be replaced with* p' *so as to improve the lower and upper bounds.*

4 The maximum of multilinear form

In this section we derive tail probabilities for the maximum T_k (1.1) of multilinear form of degree k as well as its standardized statistic U_k (1.2) defined in Section 1. Let

$$M_k = \{h_1 \otimes \cdots \otimes h_k \mid h_i \in S^{q_i-1},\ i = 1, \ldots, k\} \tag{4.1}$$

be a manifold of dimension

$$d = \sum_{i=1}^{k}(q_i - 1)$$

in R^p with $p = \prod_{i=1}^k q_i$. Since $\|h_1 \otimes \cdots \otimes h_k\| = \prod_{i=1}^k \|h_i\| = 1$, it holds that $M_k \subset S^{p-1}$. It is easy to check that M_k is a submanifold of S^{p-1} satisfying Assumption 3.1. The statistics T_k and U_k are written as

$$T_k = \max_{u \in M_k} u'z, \qquad U_k = \max_{u \in M_k} u'z/\|z\|,$$

respectively, where z is a p-dimensional column vector distributed as $N_p(0, I_p)$. Then T_k and U_k are of the form of the random variables T and U in (3.9) and (3.8) whose tail probabilities can be derived by virtue of Lemma 3.4, and Theorem 3.1 or Corollary 3.1 of Section 3.

In Section 4.1 we determine the geometric quantities of M_k . We first determine the tangent space $T_u(M_k)$ of the manifold M_k at each point $u \in M_k$, and obtain the metric (first fundamental form) $G(u) = (g_{ij}(u))$ and the volume element du at u . Second we determine the orthogonal complement $T_u^\perp = T_u(K_k)^\perp$ of the tangent space of $K_k = \bigcup_{c \geq 0} cM_k$ at $u \in M_k$. Then the second fundamental form $H(u, v)$ of M_k at $u \in M_k$ with respect to the direction $v \in T_u(K_k)^\perp$ is obtained.

In Section 4.2 the coefficient w_{d+1-e} in (3.7) for M_k shall be given. We perform double integration of the generalized trace $\mathrm{tr}_e H(u, v)$ of the second fundamental form with respect to the volume element measure dv over $T_u(K_k)^\perp \cap S^{p-1}$ followed by the integration with respect to the volume element du over M_k . By dividing the result by $\Omega_{d+1-e}\Omega_{p-d-1+e}$ we obtain w_{d+1-e} .

In addition to w_{d+1-e}, we have to know the critical radius θ_c of M_k which is required by Lemma 3.4 and Theorem 3.1. Calculation of θ_c by virtue of Lemma 3.1 is given in Section 4.3.

We present some numerical examples in Section 4.4 to show the accuracy of the obtained formulas.

4.1 Volume element and second fundamental form

We begin by determining the geometric quantities of M_k . We introduce a local coordinate system to make calculations simple. Let $t_i = (t_{i1}, \ldots, t_{i,q_i-1})'$ be a local coordinate system of S^{q_i-1} so that $h_i \in S^{q_i-1}$ has a representation $h_i = h_i(t_i)$. Then $u = h_1 \otimes \cdots \otimes h_k \in M_k$ has a local representation $u = \phi(t)$, where

$$\phi(t) = h_1(t_1) \otimes \cdots \otimes h_k(t_k)$$

with parameter $t = (t_1', \ldots, t_k')'$ of dimension $d = \sum_{i=1}^k (q_i - 1)$.

Taking a derivative of $\phi(t)$ with respect to t_{ia}, we have

$$\frac{\partial \phi}{\partial t_{ia}} = h_1 \otimes \cdots \otimes h_{i-1} \otimes \frac{\partial h_i}{\partial t_{ia}} \otimes h_{i+1} \otimes \cdots \otimes h_k.$$

The tangent space $T_u(M_k)$ at $u = \phi(t)$ is spanned by

$$\Big\{ \frac{\partial \phi}{\partial t_{ia}} \in R^p \ \Big| \ i = 1, \ldots, k, \ a = 1, \ldots, q_i - 1 \Big\},$$

and $T_u(K_k)$ is spanned by $T_u(M_k)$ and u. The (ia, jb)-th element of the metric $G = G(u)$ at u is given by

$$\Big(\frac{\partial \phi}{\partial t_{ia}}\Big)'\frac{\partial \phi}{\partial t_{jb}} = \delta_{ij}\Big(\frac{\partial h_i}{\partial t_{ia}}\Big)'\frac{\partial h_i}{\partial t_{ib}} = \delta_{ij}\bar{g}_{i,ab}, \tag{4.2}$$

where δ_{ij} is the Kronecker's delta and

$$\bar{g}_{i,ab} = \Big(\frac{\partial h_i}{\partial t_{ia}}\Big)'\frac{\partial h_i}{\partial t_{ib}}$$

is the (a, b)-th element of the metric $\bar{G}_i$ of S^{q_i-1} at $h_i = h_i(t_i)$. Therefore the metric of M_k is given by $G = \mathrm{diag}(\bar{G}_1, \ldots, \bar{G}_k)$ with $\bar{G}_i = (\bar{g}_{i,ab})$ a $(q_i - 1) \times (q_i - 1)$ matrix. The volume element at u is

$$du = |G|^{\frac{1}{2}} \prod_{i=1}^{k}\prod_{a=1}^{q_i-1} dt_{ia} = \prod_{i=1}^{k}\Big\{|\bar{G}_i|^{\frac{1}{2}}\prod_{a=1}^{q_i-1} dt_{ia}\Big\},$$

which is a product of the volume elements of S^{q_i-1}, $i = 1, \ldots, k$.

Lemma 4.1 *The volume element of* M_k *at* $u = h_1 \otimes \cdots \otimes h_k$ *is given by* $du = \prod_{i=1}^{k} dS^{q_i-1}$, *where* dS^{q_i-1} *denotes the volume element of* S^{q_i-1} *at* h_i.

We need to be careful about the fact that M_k and $S^{q_1-1} \times \cdots \times S^{q_k-1}$ are not one-to-one. Indeed $h_1 \otimes \cdots \otimes h_k$ is invariant under an even number of sign changes $h_i \mapsto -h_i$. The multiplicity of the map $g_k : S^{q_1-1} \times \cdots \times S^{q_k-1} \to M_k$ is 2^{k-1}, since the signs of $h_1, \ldots, h_{k-1}$ can be arbitrarily chosen.

Noting this fact, we have the following.

Corollary 4.1 *The total volume of* M_k *is*

$$\mathrm{Vol}(M_k) = \int_{M_k} du = \frac{1}{2^{k-1}}\prod_{i=1}^{k}\int_{S^{q_i-1}} dS^{q_i-1} = \frac{1}{2^{k-1}}\prod_{i=1}^{k}\Omega_{q_i}.$$

Let H_i be a $q_i \times (q_i - 1)$ matrix such that (h_i, H_i) is $q_i \times q_i$ orthogonal. Let

$$\frac{\partial \phi}{\partial t_i} = \Big(\frac{\partial \phi}{\partial t_{i1}}, \ldots, \frac{\partial \phi}{\partial t_{i,q_i-1}}\Big)$$

be a $p \times (q_i - 1)$ matrix, and let

$$\frac{\partial h_i}{\partial t_i} = \Big(\frac{\partial h_i}{\partial t_{i1}}, \ldots, \frac{\partial h_i}{\partial t_{i,q_i-1}}\Big)$$

be a $q_i \times (q_i - 1)$ matrix. Then the columns of two $p \times (q_i - 1)$ matrices

$$B_i = h_1 \otimes \cdots \otimes h_{i-1} \otimes H_i \otimes h_{i+1} \otimes \cdots \otimes h_k$$

and

$$\frac{\partial \phi}{\partial t_i} = h_1 \otimes \cdots \otimes h_{i-1} \otimes \frac{\partial h_i}{\partial t_i} \otimes h_{i+1} \otimes \cdots \otimes h_k$$

span the same space, since $h_i'(\partial h_i/\partial t_i) = 0$ and $\mathrm{rank}\,(\partial h_i/\partial t_i) = q_i - 1$.

Any vector orthogonal to $u = h_1 \otimes \cdots \otimes h_k$ and the column spaces of B_i, $i = 1, \ldots, k$, can be written as

$$\begin{aligned} v \;=\; & (H_1 \otimes H_2 \otimes h_3 \otimes \cdots \otimes h_k)\, e_{12} + (H_1 \otimes h_2 \otimes H_3 \otimes h_4 \otimes \cdots \otimes h_k)\, e_{13} \\ & \quad + \cdots + (h_1 \otimes \cdots \otimes h_{k-2} \otimes H_{k-1} \otimes H_k)\, e_{k-1,k} \\ & + (H_1 \otimes H_2 \otimes H_3 \otimes h_4 \otimes \cdots \otimes h_k)\, e_{123} + \cdots \\ & + \cdots \\ & + (H_1 \otimes H_2 \otimes \cdots \otimes H_k)\, e_{12\cdots k}, \end{aligned} \tag{4.3}$$

where e's are column vectors of appropriate sizes, e.g., e_{12} is $(q_1-1)(q_2-1) \times 1$, e_{123} is $(q_1-1)(q_2-1)(q_3-1) \times 1$, $e_{12\cdots k}$ is $\prod_{i=1}^k (q_i - 1) \times 1$. Then the set of vectors v in (4.3) form the linear subspace $T_u(K_k)^\perp$ whose dimension is $p - d - 1 = \prod_{i=1}^k q_i - \sum_{i=1}^k (q_i - 1) - 1$.

Now taking a second derivative we have

$$\begin{aligned} & \frac{\partial^2 \phi}{\partial t_{ia} \partial t_{jb}} \\ & = h_1 \otimes \cdots \otimes h_{i-1} \otimes \frac{\partial^2 h_i}{\partial t_{ia} \partial t_{jb}} \otimes h_{i+1} \otimes \cdots \otimes h_k && \text{if } i = j, \\ & = h_1 \otimes \cdots \otimes h_{i-1} \otimes \frac{\partial h_i}{\partial t_{ia}} \otimes h_{i+1} \otimes \cdots \otimes h_{j-1} \otimes \frac{\partial h_j}{\partial t_{jb}} \otimes h_{j+1} \otimes \cdots \otimes h_k && \text{if } i < j. \end{aligned}$$

Then for v in (4.3)

$$\begin{aligned} v' \frac{\partial^2 \phi}{\partial t_{ia} \partial t_{jb}} &= 0 && \text{if } i = j, \\ &= e_{ij}' \Big(H_i' \frac{\partial h_i}{\partial t_{ia}} \otimes H_j' \frac{\partial h_j}{\partial t_{jb}} \Big) && \text{if } i < j. \end{aligned}$$

For $i < j$ let E_{ij} be the $(q_i - 1) \times (q_j - 1)$ matrix defined by $\mathrm{vec}(E_{ij}) = e_{ij}$. There exists a $(q_i - 1) \times (q_i - 1)$ nonsingular matrix F_i such that

$$\frac{\partial h_i}{\partial t_i} = H_i F_i.$$

Then the $d \times d$ $(d = \sum_{i=1}^k (q_i - 1))$ matrix with (ia, jb)-th element $v'(\partial^2 \phi / \partial t_{ia} \partial t_{jb})$ is written as a block matrix with (i, j)-th block

$$\begin{cases} O & \text{if } i = j \ , \\ F_i' E_{ij} F_j & \text{if } i < j \ , \\ F_i' E_{ij}' F_j & \text{if } i > j \ , \end{cases}$$

$i, j = 1, \ldots, k$.

On the other hand, as we have seen in (4.2), the metric G of M_k is written as a diagonal block matrix with (i,i)-th block $F_i'F_i$, $i = 1, \ldots, k$. This implies the following lemma.

Lemma 4.2 *In an appropriate coordinate system, the second fundamental form of M_k at u with respect to the direction v in (4.3) is written as*

$$H(u,v) = -\begin{pmatrix} O & E_{12} & E_{13} & \cdots & E_{1k} \\ E_{12}' & O & E_{23} & \cdots & E_{2k} \\ E_{13}' & E_{23}' & O & \cdots & E_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_{1k}' & E_{2k}' & E_{3k}' & \cdots & O \end{pmatrix}. \tag{4.4}$$

4.2 Derivation of the coefficient w_{d+1-e}

For fixed $u \in M_k$ we evaluate the integral

$$\int_{T_u(K_k)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv, \tag{4.5}$$

where dv is the volume element of $T_u(K_k)^\perp \cap S^{p-1}$, the unit sphere restricted to $T_u(K_k)^\perp$. We introduce a random variable and replace the integration with an expectation.

Let $y \in R^p$ be a singular Gaussian vector distributed as $N_p(0, P_u^\perp)$, where $P_u^\perp$ is the orthogonal projection matrix onto the linear subspace $T_u(K_k)^\perp$. Then $r = \|y\|$ and $v = y/\|y\|$ are independently distributed. r^2 has χ^2 distribution with $p-d-1$ degrees of freedom and v has the uniform distribution over $T_u(K_k)^\perp \cap S^{p-1}$. Since $H(u,v)$ is linear in v, we have

$$\begin{aligned} E[\mathrm{tr}_e H(u,y)] &= E[\mathrm{tr}_e H(u,rv)] = E[r^e \mathrm{tr}_e H(u,v)] \\ &= E[r^e] \cdot E[\mathrm{tr}_e H(u,v)] \\ &= E[(\chi^2_{p-d-1})^{e/2}] \cdot \frac{1}{\Omega_{p-d-1}} \int \mathrm{tr}_e H(u,v)\, dv, \end{aligned}$$

where

$$E[(\chi^2_{p-d-1})^{e/2}] = 2^{e/2} \frac{\Gamma(\frac{1}{2}(p-d-1+e))}{\Gamma(\frac{1}{2}(p-d-1))}.$$

Hence we have a representation of the integral (4.5) as

$$\int_{T_u(K_k)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv = \frac{\Omega_{p-d-1}}{E[(\chi^2_{p-d-1})^{e/2}]} \cdot E[\mathrm{tr}_e H(u,y)]. \tag{4.6}$$

Note that the random vector y can be written as $Q_u \bar{y}$, where Q_u is a $p \times (p-d-1)$ matrix such that $Q_u Q_u' = P_u^\perp$ and $\bar{y}$ is a $(p-d-1)$-dimensional random vector distributed as $N_{p-d-1}(0, I_{p-d-1})$.

Now we return to our problem of multilinear from of degree k. As we saw, $T_u(K_k)^\perp$ is spanned by the vectors of the form of v in (4.3). In this parameterization

$$\|v\|^2 = \sum_{1\le i<j\le k} \|e_{ij}\|^2 + \sum_{1\le i<j<l\le k} \|e_{ijl}\|^2 + \cdots + \|e_{12\cdots k}\|^2,$$

which means that elements of the vectors

$$e_{ij}\ (i<j),\ e_{ijl}\ (i<j<l), \ldots,\ e_{12\cdots k}$$

form an orthonormal basis of $T_u(K_k)^\perp$. If we assume that every element of these vectors $e_{ij}, e_{ijl}, \ldots, e_{12\cdots k}$ is independently distributed as $N(0,1)$, then v has the distribution $N_p(0, P_u^\perp)$. Therefore the problem is reduced to evaluating the expectation $E[\mathrm{tr}_e H]$ with $H = H(u,v)$ in (4.4), where each component of E_{ij} $(i<j)$ is independently distributed as $N(0,1)$.

The expectation $E[\mathrm{tr}_e H]$ can be represented in terms of the following combinatorial quantities. Put $d_i = q_i - 1$, $i = 1,\ldots,k$, and $d = \sum_{i=1}^k d_i = \sum_{i=1}^k (q_i - 1)$. Let

$$A_i = \Big\{ \sum_{j=1}^{i-1} d_j + 1, \sum_{j=1}^{i-1} d_j + 2, \ldots, \sum_{j=1}^{i} d_j \Big\}, \quad i = 1,\ldots,k. \tag{4.7}$$

$A_1,\ldots,A_k$ form a partition of $\{1,2,\ldots,d\}$. The cardinality of A_i is $d_i = \mathrm{card}(A_i)$. Let a map $\tau : \{1,\ldots,d\} \to \{1,\ldots,k\}$ be defined by $\tau(a) = i$ for $a \in A_i$, i.e., $\tau(a)$ indicates the set which contains a.

Consider a set of m pairings

$$\{(a_1,a_2),\ldots,(a_{2m-1},a_{2m}) \mid a_1 < a_3 < \cdots < a_{2m-1},\ a_1 < a_2, \ldots, a_{2m-1} < a_{2m}\} \tag{4.8}$$

such that

(i) $2m$ indices $a_1, a_2, \ldots, a_{2m}$ are distinct elements of $\{1,2,\ldots,d\}$.

(ii) For each pairing in (4.8), say (a_{2l-1}, a_{2l}), a_{2l-1} and a_{2l} do not belong to the same set of (4.7), i.e., $\tau(a_{2l-1}) \ne \tau(a_{2l})$, $l = 1,\ldots,m$.

Furthermore let $n_k(d_1,d_2,\ldots,d_k;m)$ denote the total number of sets of m pairings satisfying (i) and (ii). Now we can state the following lemma.

Lemma 4.3 *Let y be distributed as $N_p(0, P_u^\perp)$. Then*

$$\begin{aligned} E[\mathrm{tr}_e H(u,y)] &= (-1)^{e/2} n_k(d_1,d_2,\ldots,d_k;e/2) && \textit{for } e \textit{ even}, \\ &= 0 && \textit{for } e \textit{ odd}. \end{aligned}$$

Proof. Note first that the generalized trace $\mathrm{tr}_e H$ of H is written as

$$\mathrm{tr}_e H = \sum_{\substack{A\subset\{1,\ldots,d\} \\ \mathrm{card}(A)=e}} \det H[A],$$

where $H[A]$ with $A = \{1 \leq a_1 < \cdots < a_e \leq d\}$ denotes the $e \times e$ submatrix of H formed by deleting all but columns and rows of H numbered $a_1, \ldots, a_e$ (Muirhead (1982), Appendix A7). Consider the termwise expectation

$$E[\det H[A]] = \sum_{\pi \in P(A)} \operatorname{sgn}(\pi)\, E[h_{a_1\pi(a_1)} \cdots h_{a_e\pi(a_e)}] \tag{4.9}$$

where $P(A)$ is the set of permutations of the elements of A.

Since $H = (h_{ab})_{1 \leq a,b \leq d}$ is a symmetric random matrix whose diagonal and upper off-diagonal elements are zero mean independent random variables (maybe a constant 0), $E[h_{a_1\pi(a_1)} \cdots h_{a_e\pi(a_e)}] = 0$ unless e is even and $\pi(a) \neq a$, $\pi(\pi(a)) = a$, $\forall a$. In this case $\operatorname{sgn}(\pi) = (-1)^{e/2}$, and by relabeling the indices of a's, non-vanishing terms in (4.9) can be written uniquely in the form of

$$(-1)^{e/2} E[h^2_{a_1a_2} h^2_{a_3a_4} \cdots h^2_{a_{e-1}a_e}]$$

with $a_1 < a_3 < \cdots < a_{m-1}$, $a_1 < a_2, \ldots, a_{e-1} < a_e$. Moreover $h_{a_{2l-1}a_{2l}} = 0$ for $\tau(a_{2l-1}) = \tau(a_{2l})$, since $H[A_i] = O$ ($d_i \times d_i$ matrix consisting of 0's), $i = 1, \ldots, k$.

Therefore for e even we have

$$E[\operatorname{tr}_e H(u, y)] = (-1)^{e/2} \sum_{\substack{A \subset \{1,\ldots,d\} \\ \operatorname{card}(A) = e}} \sum_{\substack{a_1 < a_3 < \cdots < a_{e-1} \\ a_1 < a_2, \ldots, a_{2e-1} < a_{2e} \\ \tau(a_{2l-1}) \neq \tau(a_{2l}),\ \forall l}} E[h^2_{a_1a_2} h^2_{a_3a_4} \cdots h^2_{a_{e-1}a_e}].$$

Since the expectation in the right hand side is 1, and the summation is taken over all sets of $m = e/2$ pairings (4.8) satisfying (i) and (ii), we prove the lemma. ∎

For $k = 2$ and 3, n_k can be written as follows.

Lemma 4.4 *For $d_i \geq 0$, $m \geq 0$,*

$$\begin{aligned} n_2(d_1, d_2; m) &= \frac{d_1!\, d_2!}{(d_1 - m)!\, (d_2 - m)!\, m!} && \textit{if } m \leq \min(d_1, d_2), \\ &= 0 && \textit{otherwise}, \end{aligned} \tag{4.10}$$

and

$$n_3(d_1, d_2, d_3; m) = \sum_{\substack{l_1 + l_2 + l_3 = m \\ l_i \geq \max(m - d_i, 0),\ \forall i}} \prod_{i=1}^{3} \frac{d_i!}{l_i!\, (l_i - m + d_i)!}, \tag{4.11}$$

where the summation is taken over triplets (l_1, l_2, l_3) of integers such that $l_1 + l_2 + l_3 = m$ and $l_i \geq 0$, $l_i - m + d_i \geq 0$, $i = 1, 2, 3$. Put $n_3(d_1, d_2, d_3; m) = 0$ if there does not exist a feasible triplet (l_1, l_2, l_3).

Proof. Consider the case $k = 2$. $n_2(d_1, d_2; m)$ is the total number of m pairings of the form

$$\{(b_1, c_1), \ldots, (b_m, c_m)\}, \quad b_1, \ldots, b_m \in A_1,\ c_1, \ldots, c_m \in A_2.$$

There are $\binom{d_1}{m}$ ways of choosing m elements from $A_1 = \{1, \ldots, d_1\}$ and there are $\binom{d_2}{m}$ ways of choosing m elements from $A_2 = \{d_1 + 1, \ldots, d\}$. Furthermore there are $m!$ ways of forming pairs of the $2m$ chosen elements. Therefore

$$n_2(d_1, d_2; m) = \binom{d_1}{m}\binom{d_2}{m} m!.$$

This proves (4.10).

Consider the case $k = 3$. Fix $l_1, l_2, l_3 \geq 0$ such that $l_1 + l_2 + l_3 = m$. Choose two subsets B_{12} and B_{13} of $A_1 = \{1, \ldots, d_1\}$ such that $\text{card}(B_{12}) = l_3$, $\text{card}(B_{13}) = l_2$, $B_{12} \cap B_{13} = \emptyset$. Similarly choose $B_{21} \subset A_2 = \{d_1 + 1, \ldots, d_1 + d_2\}$ and $B_{23} \subset A_2$ such that $\text{card}(B_{21}) = l_3$, $\text{card}(B_{23}) = l_1$, $B_{21} \cap B_{23} = \emptyset$; choose $B_{31} \subset A_3 = \{d_1 + d_2 + 1, \ldots, d\}$ and $B_{32} \subset A_3$ such that $\text{card}(B_{31}) = l_2$, $\text{card}(B_{32}) = l_1$, $B_{31} \cap B_{32} = \emptyset$. There are $l_3!$ ways of making l_3 pairings between B_{12} and B_{21}. Similarly there are $l_2!$ ways of making l_2 pairings between B_{13} and B_{31}, and $l_1!$ ways of making l_1 pairings between B_{23} and B_{32}. Then for fixed l_1, l_2, l_3 there are

$$\binom{d_1}{l_2,\ l_3,\ d_1 - l_2 - l_3}\binom{d_2}{l_1,\ l_3,\ d_2 - l_1 - l_3}\binom{d_3}{l_1,\ l_2,\ d_3 - l_1 - l_2} l_1!\, l_2!\, l_3!$$

ways of making m pairings of the form (4.8) satisfying (i) and (ii). Taking summation for feasible triplets (l_1, l_2, l_3) proves (4.11). ∎

For $k \neq 2, 3$, the following recurrence formula is useful for calculating $n_k(d_1, \ldots, d_k; m)$. Since $n_k(d_1, \ldots, d_k; m)$ is symmetric in $d_1, \ldots, d_k$, we can restrict our attention to $d_1 \geq \cdots \geq d_k$.

Lemma 4.5 *For $k \geq 2$, $d_1 \geq \cdots \geq d_k \geq 0$, and $m \geq 0$, it holds*

$$\begin{aligned}
&n_k(d_1, \ldots, d_k; m) \\
&\quad = 1 && \textit{if } m = 0, \\
&\quad = 0 && \textit{if } m > 0,\ d_k = 0,\ k = 2, \\
&\quad = n_{k-1}(d_1, \ldots, d_{k-1}; m) && \textit{if } m > 0,\ d_k = 0,\ k \geq 3, \\
&\quad = n_k(d_1 - 1, d_2, \ldots, d_k; m) \\
&\qquad + \sum_{j=2}^{k} d_j\, n_k(d_1 - 1, d_2, \ldots, d_{j-1}, d_j - 1, d_{j+1}, \ldots, d_k; m - 1) \\
&&& \textit{otherwise.}
\end{aligned}$$

Here in the last expression the arguments of n_k should be reordered so that $d_1 \geq d_2 \geq \cdots \geq d_k \geq 0$ (if necessary). For example, if $d_2 > d_1 - 1 \geq d_3$, $n_k(d_1 - 1, d_2, \ldots, d_k; m)$ should be replaced with $n_k(d_2, d_1 - 1, d_3, \ldots, d_k; m)$.

Proof. Consider the first element '1' of $A_1 = \{1, \ldots, d_1\}$. Among $n_k(d_1, \ldots, d_k; m)$ possible ways of m pairings, there are $n_k(d_1 - 1, d_2, \ldots, d_k; m)$ ways where '1' does not appear in the m pairings; while there are

$$d_j \times n_k\,(d_1 - 1, d_2, \ldots, d_{j-1}, d_j - 1, d_{j+1}, \ldots, d_k; m - 1)$$

ways where '1' makes a pairing with one element of A_j. Then the last equation of the recurrence formula follows when $m \geq 1$ and $d_k \geq 1$. The other three equations are obvious boundary conditions for the recursion. ∎

Now we proceed to integrating (4.5) with respect to du:

$$\int_{M_k} \Big[\int_{T_u(K_k)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u, v)\, dv \Big]\, du.$$

As we already saw, the integrand does not depend on u. Therefore the integration with respect to du over M_k reduces to multiplying by a constant $\int_{M_k} du = \mathrm{Vol}(M_k)$ obtained in Corollary 4.1.

Then from (4.6) the coefficient in (3.7) for M_k is

$$\begin{aligned} w_{d+1-e} &= \frac{1}{\Omega_{d+1-e}\Omega_{p-d-1+e}} \cdot \mathrm{Vol}(M_k) \cdot \frac{\Omega_{p-d-1}}{E[(\chi^2_{p-d-1})^{e/2}]} E[\mathrm{tr}_e H(u, y)] \\ &= \frac{\mathrm{Vol}(M_k)}{\Omega_{d+1}} \cdot \frac{\Gamma(\frac{1}{2}(d + 1 - e))}{2^{e/2}\,\Gamma(\frac{1}{2}(d + 1))} E[\mathrm{tr}_e H(u, y)]. \end{aligned}$$

Summarizing the above calculations, we obtain the following theorem.

Theorem 4.1 *The nonzero coefficient w_{d+1-e} in (3.7) for the tail probabilities of T_k in (1.1) and U_k in (1.2) is given by*

$$w_{d+1-e} = \frac{\pi^{(k-1)/2}}{\prod_{i=1}^k \Gamma(\frac{1}{2}q_i)} \Big(-\frac{1}{2}\Big)^{e/2} \Gamma\Big(\frac{1}{2}(d + 1 - e)\Big)\, n_k(q_1 - 1, \ldots, q_k - 1; e/2),$$

$e = 0, 2, \ldots, [d/2] \times 2$, *where* $d = \sum_{i=1}^k (q_i - 1)$, *and* $n_k(q_1 - 1, \ldots, q_k - 1; e/2)$ *is given by Lemmas 4.4 or 4.5.*

Corollary 4.2 *Let W be a $q \times q$ matrix distributed according to the Wishart distribution $W_q(\nu, I_q)$, and let $\lambda_1(W)$ be the largest eigenvalue of W. Then the tail probability $P(\lambda_1(W) \geq x)$ is given by (3.12) with $a = \sqrt{x}$,*

$$\begin{aligned} w_{d+1-e} &= w_{q+\nu-1-e} \\ &= \frac{\sqrt{\pi}\,\Gamma(q)\,\Gamma(\nu)}{\Gamma(\frac{1}{2}q)\,\Gamma(\frac{1}{2}\nu)} \Big(-\frac{1}{2}\Big)^{e/2} \frac{\Gamma(\frac{1}{2}(q + \nu - 1) - e/2)}{\Gamma(q - e/2)\,\Gamma(\nu - e/2)\,(e/2)!} \end{aligned} \qquad (4.12)$$

for $e/2 = 0, 1, \ldots, \min(q, \nu) - 1$ *; the other* w_{d+1-e}*'s are 0.*

Note that we can assume $\nu \geq q$ without loss of generality. Indeed (4.12) is symmetric in ν and q.

Remark 4.1 *It can be proved that there is a relation on the coefficients $w_{q+\nu-1-e}$ in (4.12): For $\nu > q-1$ (ν is not necessarily integer)*

$$\sum_{j=0}^{q-1} w_{q+\nu-1-2j} = \begin{cases} 1 & \text{if } q \text{ is odd,} \\ 0 & \text{if } q \text{ is even.} \end{cases} \tag{4.13}$$

Noting this relation and that $\bar{G}_m(x) = 2g_m(x) + \bar{G}_{m-2}(x)$, we can rewrite the formula for the tail probability of $\lambda_1(W)$ given by Corollary 4.2 as

$$P(\lambda_1(W) \geq x) \quad \sim \quad \sum_{j=0}^{q-2} \bar{w}_{q+\nu-1-2j}\, g_{q+\nu-1-2j}(x) + \begin{cases} \bar{G}_{\nu-q+1}(x) & \text{if } q \text{ is odd} \\ 0 & \text{if } q \text{ is even} \end{cases} \tag{4.14}$$

where

$$\bar{w}_{q+\nu-1-2j} = 2\sum_{i=0}^{j} w_{q+\nu-1-2i}.$$

This is another expression of the formula by Hanumara and Thompson (1968). They concluded that this formula is accurate enough for calculating significance levels and made a table of quantiles based on it. Although they did not state any mathematical properties, we now know that (4.14) is justified as an asymptotic expansion as x goes to infinity.

A proof of (4.13) shall be given in Appendix D.

4.3 Critical radius

In this subsection we obtain the concrete value of the critical radius θ_c of the manifold M_k in (4.1) by virtue of Lemma 3.1.

Fix a point $v = h_1 \otimes \cdots \otimes h_k \in M_k$ with $h_i \in S^{q_i-1}$. Let H_i, $i = 1, \ldots, k$, be $q_i \times (q_i - 1)$ matrices such that (h_i, H_i) is $q_i \times q_i$ orthogonal. Let $K_k = \bigcup_{c \geq 0} cM_k$ be the cone associated with M_k. The tangent space $T_v(K_k)$ at v is spanned by $v = h_1 \otimes \cdots \otimes h_k$ and the column spaces of

$$B_i = h_1 \otimes \cdots \otimes h_{i-1} \otimes H_i \otimes h_{i+1} \otimes \cdots \otimes h_k, \quad i = 1, \ldots, k.$$

Then the orthogonal projection matrix onto $T_v(K_k)$ is given by

$$\begin{aligned} P_v &= vv' + \sum_{i=1}^{k} B_i B_i' \\ &= \sum_{i=1}^{k} h_1 h_1' \otimes \cdots \otimes h_{i-1} h_{i-1}' \otimes I_{q_i} \otimes h_{i+1} h_{i+1}' \otimes \cdots \otimes h_k h_k' \\ &\quad -(k-1) h_1 h_1' \otimes \cdots \otimes h_k h_k'. \end{aligned}$$

Let $\tilde{v} = \tilde{h}_1 \otimes \cdots \otimes \tilde{h}_k \in M_k$. Then $\tilde{v}'v = \prod_{i=1}^k (\tilde{h}_i' h_i)$ and

$$\tilde{v}' P_v \tilde{v} = \sum_{i=1}^k \prod_{j \neq i} (\tilde{h}_j' h_j)^2 - (k-1) \prod_{i=1}^k (\tilde{h}_i' h_i)^2.$$

Note that both $\tilde{v}' P_v \tilde{v}$ and $\tilde{v}'v$ depend on $\tilde{v}$ and v through $\tilde{h}_i' h_i = x_i$ (say) which takes values $-1 \leq x_i \leq 1$. Then by Lemma 3.1

$$\cot^2 \theta_c = \sup_{\tilde{v}, v \in M} \frac{1 - \tilde{v}' P_v \tilde{v}}{(1 - \tilde{v}'v)^2} = \sup_{-1 < x_i < 1, \ \forall i} \frac{1 - \sum_i \prod_{j \neq i} x_j^2 + (k-1) \prod_i x_i^2}{(1 - \prod_i x_i)^2}.$$

Here we take the supremum by two steps: First, take the supremum under the restriction that $\prod_i x_i$ ($= y$, say) is fixed. Second, take the supremum with respect to $-1 < y < 1$. By the relation between the arithmetic and geometric means, we have

$$\sum_{i=1}^k \prod_{j \neq i} x_i^2 \geq k \Big(\prod_{i=1}^k \prod_{j \neq i} x_j^2 \Big)^{1/k} = k \, |y|^{2(k-1)/k},$$

where the equality holds if and only if $x_1^2 = \cdots = x_k^2$. Then we have

$$\cot^2 \theta_c = \sup_{-1 < y < 1} \frac{1 - k \, |y|^{2(k-1)/k} + (k-1) y^2}{(1-y)^2}. \tag{4.15}$$

Note that in (4.15) we can restrict y to be nonnegative. Here we give a lemma, whose proof is given in Appendix C.

Lemma 4.6

$$\sup_{0 \leq z < 1} \frac{1 - k z^{2(k-1)} + (k-1) z^{2k}}{(1 - z^k)^2} = \frac{2(k-1)}{k}, \tag{4.16}$$

where the supremum is attained when $z \uparrow 1$.

Then by making a change of variable $y = z^k$ in (4.15), we have by Lemma 4.6 that $\cot^2 \theta_c = 2(k-1)/k$.

Theorem 4.2 *The critical radius* θ_c *of* M_k *in (4.1) is given by*

$$\cos^2 \theta_c = \frac{2k-2}{3k-2}, \quad k \geq 2.$$

4.4 Examples

4.4.1 The maximum of 2-form (3×3)

Consider the statistic T_2 in (1.1) with $q_1 = 3, \ q_2 = 3$. Then T_2 is the square root of the largest eigenvalue of the Wishart matrix $W_3(3, I_3)$. Then $p = q_1 q_2 = 9$ and $d = q_1 + q_2 - 2 = 4$. The approximate tail probability for T_2 is given by

$$P(T_2 \geq x) \sim 3 \bar{G}_5(x^2) - 4 \bar{G}_3(x^2) + 2 \bar{G}_1(x^2). \tag{4.17}$$

Since the critical radius θ_c is given by $\tan^2\theta_c = 1$, the lower bound is

$$Q_L(x) = 3Q_{5,4}(x^2, 1) - 4Q_{3,6}(x^2, 1) + 2Q_{1,8}(x^2, 1). \tag{4.18}$$

Let M_c denote the tube of distance θ_c around M_k. The upper bound is

$$Q_U(x) = Q_L(x) + \bar{G}_9(2x^2)\,(1 - \mathrm{Vol}(M_c)/\Omega_9), \tag{4.19}$$

where $\mathrm{Vol}(M_c)/\Omega_9 \doteq 0.990$.

In Figure 4.1 the approximate tail probability by (4.17) as well as the exact tail probability calculated by the method of Pfaffian (e.g. Pillai (1976)) are plotted by solid lines. (A monotonically decreasing curve corresponds to the exact value.) The lower and upper bounds by (4.18) and (4.19) are plotted by broken lines. The exact value and the upper bound are too close to be distinguished. We can conclude in this case that the approximation formula by asymptotic expansion is sufficiently accurate.

Also recalling that the value of $\mathrm{Vol}(M_c)/\Omega_p$ is the maximum p-value which can be calculated by Lemma 3.4, we can also conclude that Lemma 3.4 provides a practical method for calculating p-values of U_2^2.

4.4.2 The maximum of 3-form $(2 \times 2 \times 2)$

As another example we consider the statistic T_3 in (1.1) with $q_1 = q_2 = q_3 = 2$. Then $p = \prod_i q_i = 8$ and $d = \sum_i (q_i - 1) = 3$. Since $n_3(1,1,1;0) = 1$ and $n_3(1,1,1;2/2) = 3$, we have $w_4 = \pi$, $w_2 = -3\pi/2$, and the other w's are 0. Therefore we have

$$P(T_3 \geq x) \sim \pi \bar{G}_4(x^2) - (3\pi/2)\bar{G}_2(x^2). \tag{4.20}$$

By Theorem 4.2 the critical radius θ_c of M_3 in (4.1) is given by $\cos^2\theta_c = 4/7$. Then $\tan^2\theta_c = 3/4$ and the lower and upper bounds for $P(T_3 \geq x)$ are given by

$$Q_L(x) = \pi Q_{4,4}(x^2, 3/4) - (3\pi/2)Q_{2,6}(x^2, 3/4) \tag{4.21}$$

and

$$Q_U(x) = Q_L(x) + \bar{G}_8(7x^2/4)\,(1 - \mathrm{Vol}(M_c)/\Omega_8), \tag{4.22}$$

where

$$\mathrm{Vol}(M_c)/\Omega_8 = \pi \bar{B}_{2,2}(4/7) - (3\pi/2)\bar{B}_{1,3}(4/7) \doteq 0.866.$$

These three functions (4.20), (4.21), and (4.22) are plotted in Figure 4.2. A solid line indicates the approximate value by (4.20), while two broken lines indicate the lower and upper bounds by (4.21) and (4.22). Differently from the case of Wishart matrix, the exact distribution of T_3 is not known. However, we see that the asymptotic expansion (4.20) gives a fairly good approximation because it is located in the narrow band between Q_L and Q_U.

Also $\mathrm{Vol}(M_c)/\Omega_8$ is adequately large and in this case Lemma 3.4 is practical enough for calculating p-values of U_3^2.

5 The maximum of symmetric multilinear form

In this section we obtain tail probabilities for the maximum $\tilde{T}_k$ (1.3) of symmetric multilinear form of degree k and its standardized statistic $\tilde{U}_k$ (1.4) introduced in Section 1. The construction of this section is the same as that of Section 4. In Section 5.1 the geometric quantities such as the volume element and the second fundamental form are determined. The coefficient w_{d+1-e} in (3.7) is given in Section 5.2. Section 5.3 is devoted to the calculation of critical radius. A numerical example is illustrated in the last Section 5.4.

5.1 Volume element and second fundamental form

Let

$$\tilde{M}_k = \{\underbrace{h \otimes \cdots \otimes h}_{k} \mid h \in S^{q-1}\} \tag{5.1}$$

be a manifold of dimension

$$d = q - 1$$

in S^{p-1} with $p = q^k$. Throughout Section 5, we use d and $q-1$ interchangeably. As is the manifold M_k in (4.1), it is easy to check that $\tilde{M}_k$ is a submanifold of S^{p-1} satisfying Assumption 3.1. The statistics $\tilde{T}_k$ and $\tilde{U}_k$ are written as

$$\tilde{T}_k = \max_{u \in \tilde{M}_k} u'z, \qquad \tilde{U}_k = \max_{u \in \tilde{M}_k} u'z/\|z\|,$$

respectively, where z is a p-dimensional column vector distributed as $N_p(0, I_p)$. Here it is to be noted that the representation $(h \otimes \cdots \otimes h)'z$ is not of minimal dimension. $\tilde{M}_k$ or its associated cone $\tilde{K}_k = \bigcup_{c \geq 0} c\tilde{M}_k$ is degenerated. It is easily proved that

$$\dim \mathrm{lin}(\tilde{K}_k) = \binom{q+k-1}{k}$$

(see, e.g., Takemura (1993)). As stated in Remark 3.2 we have to be careful that the $p = q^k$ appearing in Theorem 3.1 should be replaced with $p' = \binom{q+k-1}{k}$.

First of all, we introduce a local coordinate system for the sake of convenience of calculation. Let $t = (t_1, \ldots, t_{q-1})'$ be a local coordinate system of S^{q-1} so that $h \in S^{q-1}$ has a representation $h = h(t)$. Then $u = h \otimes \cdots \otimes h \in \tilde{M}_k$ has a local representation $u = \varphi(t)$ where

$$\varphi(t) = \underbrace{h(t) \otimes \cdots \otimes h(t)}_{k}.$$

Taking a derivative of $\varphi(t)$ with respect to t_i, we have

$$\frac{\partial \varphi}{\partial t_i} = \sum_{l=1}^{k} \underbrace{h \otimes \cdots \otimes h}_{l-1} \otimes \frac{\partial h}{\partial t_i} \otimes \underbrace{h \otimes \cdots \otimes h}_{k-l}.$$

The tangent space $T_u(\tilde{M}_k)$ at $u = \varphi(t)$ is spanned by

$$\left\{\frac{\partial\varphi}{\partial t_i} \in R^p \;\middle|\; i = 1, \ldots, d\right\}.$$

The tangent space $T_u(\tilde{K}_k)$ of $\tilde{K}_k$ is spanned by $T_u(\tilde{M}_k)$ and u. The (i,j)-th element of the metric $G = G(u)$ at u is given by

$$\left(\frac{\partial\varphi}{\partial t_i}\right)' \frac{\partial\varphi}{\partial t_j} = k\left(\frac{\partial h}{\partial t_i}\right)' \frac{\partial h}{\partial t_j} = k\bar{g}_{ij}, \tag{5.2}$$

where

$$\bar{g}_{ij} = \left(\frac{\partial h}{\partial t_i}\right)' \frac{\partial h}{\partial t_j}$$

is the (i,j)-th element of the metric $\bar{G}$ of S^{q-1} at $h = h(t)$. Therefore we have $G = k\bar{G}$, and hence the volume element at u is

$$du = |G|^{\frac{1}{2}} \prod_{i=1}^{q-1} dt_i = k^{\frac{1}{2}(q-1)} |\bar{G}|^{\frac{1}{2}} \prod_{i=1}^{q-1} dt_i.$$

Lemma 5.1 *The volume element of* $\tilde{M}_k$ *at* $u = \underbrace{h \otimes \cdots \otimes h}_{k}$ *is given by* $du = k^{\frac{1}{2}(q-1)} dS^{q-1}$, *where* dS^{q-1} *denotes the volume element of* S^{q-1} *at* h.

As in the case of M_k, $\tilde{M}_k$ and S^{p-1} are not necessarily one-to-one. When k is even, $\underbrace{h \otimes \cdots \otimes h}_{k}$ is invariant under the sign change $h \mapsto -h$ and hence the multiplicity of the map $\tilde{g}_k : S^{q-1} \to \tilde{M}_k$ is 2. On the other hand when k is odd, it is easy to show that $\tilde{M}_k$ and S^{q-1} are one-to-one.

Noting this fact, we have the following.

Corollary 5.1 *The total volume of* $\tilde{M}_k$ *is*

$$\mathrm{Vol}(\tilde{M}_k) = \int_{\tilde{M}_k} du = \frac{k^{\frac{1}{2}(q-1)}}{2^{(k-1) \bmod 2}} \int_{S^{q-1}} dS^{q-1} = \frac{k^{\frac{1}{2}(q-1)}}{2^{(k-1) \bmod 2}} \Omega_q.$$

Here, $(k-1) \bmod 2 = 1$ *if* k *is even,* $= 0$ *if* k *is odd.*

Let H be a $q \times (q-1)$ matrix such that (h, H) is $q \times q$ orthogonal. Using H any vector $v \in R^p$ orthogonal to $u = \varphi(t)$ is written as

$$\begin{aligned} v \;=\; & (H \otimes \underbrace{h \otimes \cdots \otimes h}_{k-1})\, e_1 + (h \otimes H \otimes \underbrace{h \otimes \cdots \otimes h}_{k-2})\, e_2 \\ & + \cdots + (\underbrace{h \otimes \cdots \otimes h}_{k-1} \otimes H)\, e_k \\ & + (H \otimes H \otimes \underbrace{h \otimes \cdots \otimes h}_{k-2})\, e_{12} + (H \otimes h \otimes H \otimes \underbrace{h \otimes \cdots \otimes h}_{k-3})\, e_{13} \end{aligned}$$

$$
\begin{aligned}
&+\cdots+(\underbrace{h\otimes\cdots\otimes h}_{k-2}\otimes H\otimes H)\,e_{k-1,k}\\
&+(H\otimes H\otimes H\otimes\underbrace{h\otimes\cdots\otimes h}_{k-3})\,e_{123}+\cdots\\
&+\cdots\\
&+(\underbrace{H\otimes H\otimes\cdots\otimes H}_{k})\,e_{12\cdots k},
\end{aligned}
\tag{5.3}
$$

where $e_{i_1\cdots i_l}$ $(1\le i_1<\cdots<i_l\le k)$ is a $(q-1)^l\times 1$ column vector.

Suppose that $v\in T_u(\tilde{K}_k)^\perp$. Then it has to be

$$
v'\frac{\partial\varphi}{\partial t_i}=\sum_{l=1}^{k}e_l'H'\frac{\partial h}{\partial t_i}=0.
$$

Since the columns of a $q\times(q-1)$ matrix

$$
\frac{\partial h}{\partial t}=\Big(\frac{\partial h}{\partial t_1},\ldots,\frac{\partial h}{\partial t_{q-1}}\Big)
$$

is of rank $q-1$ whose columns are orthogonal to h, it holds that $\sum_{l=1}^{k}e_l=0$.

Conversely, v in (5.3) with $\sum_{l=1}^{k}e_l=0$ belongs to the linear subspace $T_u(\tilde{K}_k)^\perp$ of dimension $p-d-1=q^k-q$.

Now taking a second derivative we have

$$
\begin{aligned}
\frac{\partial^2\varphi}{\partial t_i\partial t_j}
&=\sum_{l=1}^{k}\underbrace{h\otimes\cdots\otimes h}_{l-1}\otimes\frac{\partial^2 h}{\partial t_i\partial t_j}\otimes\underbrace{h\otimes\cdots\otimes h}_{k-l}\\
&\quad+\sum_{1\le l<m\le k}\underbrace{h\otimes\cdots\otimes h}_{l-1}\otimes\frac{\partial h}{\partial t_i}\otimes\underbrace{h\otimes\cdots\otimes h}_{m-l-1}\otimes\frac{\partial h}{\partial t_j}\otimes\underbrace{h\otimes\cdots\otimes h}_{k-m}\\
&\quad+\sum_{1\le l<m\le k}\underbrace{h\otimes\cdots\otimes h}_{l-1}\otimes\frac{\partial h}{\partial t_j}\otimes\underbrace{h\otimes\cdots\otimes h}_{m-l-1}\otimes\frac{\partial h}{\partial t_i}\otimes\underbrace{h\otimes\cdots\otimes h}_{k-m}
\end{aligned}
$$

Then for v in (5.3) with $\sum_{l=1}^{k}e_l=0$

$$
v'\frac{\partial^2\varphi}{\partial t_i\partial t_j}=\sum_{1\le l<m\le k}e_{lm}'\Big(H'\frac{\partial h}{\partial t_i}\otimes H'\frac{\partial h}{\partial t_j}+H'\frac{\partial h}{\partial t_j}\otimes H'\frac{\partial h}{\partial t_i}\Big).
$$

For $l<m$ let E_{lm} be the $(q-1)\times(q-1)$ matrix defined by $\mathrm{vec}(E_{lm})=e_{lm}$. There exists a $(q-1)\times(q-1)$ nonsingular matrix F such that

$$
\frac{\partial h}{\partial t}=HF.
$$

Then $v'(\partial^2\varphi/\partial t_i\partial t_j)$ is shown to be the (i,j)-th element of

$$
F'\Big\{\sum_{1\le l<m\le k}(E_{lm}+E_{lm}')\Big\}F.
$$

On the other hand, as we have seen in (5.2), the metric G of $\check{M}_k$ is written as $k\,F'F$. Therefore we have the following lemma.

Lemma 5.2 *In an appropriate coordinate system, the second fundamental form of* $\tilde{M}_k$ *at* u *with respect to the direction* v *in (5.3) with* $\sum_{l=1}^k e_l = 0$ *is written as*

$$H(u,v) = -\frac{1}{k} \sum_{1 \le l < m \le k} (E_{lm} + E'_{lm}). \tag{5.4}$$

5.2 Derivation of the coefficient w_{d+1-e}

Now let us proceed to the evaluation of the integral

$$\int_{T_u(\tilde{K}_k)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv. \tag{5.5}$$

As in Section 4 we calculate this integral through taking an expectation.

Let R_k be a $k \times (k-1)$ matrix such that

$$R'_k R_k = I_{k-1} \quad \text{and} \quad 1'_k R_k = 0,$$

where 1_k is a $k \times 1$ vector consisting of 1's. Then $q \times 1$ vectors $e_1, \ldots, e_k$ satisfying $\sum_{l=1}^k e_l = 0$ can be reparameterized as

$$(e_1, \ldots, e_k) = (\bar{e}_1, \ldots, \bar{e}_{k-1}) R'_k,$$

where $\bar{e}_i$ is $(q-1) \times 1$. Using this parameterization, the squared norm of v with $\sum_{l=1}^k e_l = 0$ is written as

$$\|v\|^2 = \sum_{1 \le i \le k-1} \|\bar{e}_i\|^2 + \sum_{1 \le i < j \le k} \|e_{ij}\|^2 + \sum_{1 \le i < j < l \le k} \|e_{ijl}\|^2 + \cdots + \|e_{12\cdots k}\|^2,$$

which means that elements of the vectors

$$\bar{e}_i,\ e_{ij}\ (i<j),\ e_{ijl}\ (i<j<l), \ldots,\ e_{12\cdots k} \tag{5.6}$$

form an orthonormal basis of $T_u(\tilde{K}_k)^\perp$. We consider that every element of these vectors (5.6) is independent random variables distributed as $N(0,1)$ and take the expectation $E[\mathrm{tr}_e H(u,v)]$. Then the integral (5.5) can be evaluated by

$$\int_{T_u(\tilde{K}_k)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv = \frac{\Omega_{p-d-1}}{E[(\chi^2_{p-d-1})^{e/2}]} \cdot E[\mathrm{tr}_e H(u,v)].$$

Rewrite $H(u,v)$ in (5.4) as

$$H(u,v) = \sqrt{\frac{2(k-1)}{k}} C_d,$$

where

$$C_d = -\frac{1}{\sqrt{2(k-1)k}} \sum_{1 \le l < m \le k} (E_{lm} + E'_{lm}).$$

We have assumed that each component of E_{lm} $(l < m)$ is independently distributed as $N(0,1)$. $C_d = (c_{ij})$ is a $d \times d$ symmetric random matrix whose diagonal element c_{ii} and upper off-diagonal element c_{ij} $(i < j)$ are distributed independently as $N(0,1)$ and $N(0,1/2)$, respectively.

Consider

$$E[\mathrm{tr}_e H] = \sum_{\substack{A \subset \{1,\dots,d\} \\ \mathrm{card}(A)=e}} E[\det\, H[A]].$$

For e odd $E[\mathrm{tr}_e H] = 0$ holds because any central moment of odd degrees is 0. Now suppose that e is even. Since $H[A]$ is equivalent in distribution to $\sqrt{2(k-1)/k}\, C_e$, we have

$$E[\mathrm{tr}_e H] = \binom{d}{e} \Bigl\{ \frac{2(k-1)}{k} \Bigr\}^{e/2} E[\det\, C_e]. \tag{5.7}$$

Here for $C_e = (c_{ij})$

$$E[\det\, C_e] = \sum_{\pi \in P(\{1,\dots,e\})} \mathrm{sgn}(\pi) E[c_{1\pi(1)} c_{2\pi(2)} \cdots c_{e\pi(e)}]. \tag{5.8}$$

The expectation of the right hand side of (5.8) above does not vanish if and only if $\pi(i) \neq i$ and $\pi(\pi(i)) = i$ for any i. In this case $\mathrm{sgn}(\pi) = (-1)^{e/2}$, and non-vanishing terms of the right hand side of (5.8) can be written uniquely in the form

$$(-1)^{e/2} E[c_{i_1 i_2}^2 c_{i_3 i_4}^2 \cdots c_{i_{e-1} i_e}^2]$$

with $i_1 < i_3 < \cdots < i_{e-1}$, $i_1 < i_2, \dots, i_{e-1} < i_e$. Counting the number of ways of forming $e/2$ pairings from $\{1, 2, \dots, e\}$

$$\{(i_1, i_2), (i_3, i_4), \dots, (i_{e-1}, i_e) \mid i_1 < i_3 < \cdots < i_{e-1},\ i_1 < i_2, \dots, i_{e-1} < i_e\},$$

we have for e even that

$$E[\det\, C_e] = (-1)^{e/2} \frac{e!}{2^{e/2}(e/2)!} (1/2)^{e/2}.$$

Hence from (5.7)

$$E[\mathrm{tr}_e H] = \Bigl(-\frac{k-1}{2k} \Bigr)^{e/2} \frac{d!}{(d-e)!\,(e/2)!}.$$

Now it remains to evaluate the integral

$$\int_{\tilde{M}_k} \Bigl[\int_{T_u(\tilde{K}_k)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv \Bigr] du.$$

As in the case of multilinear form in Section 4, the integrand does not depend on u, and the integration with respect to du over $\tilde{M}_k$ reduces to multiplying by a constant $\int_{\tilde{M}_k} du = \mathrm{Vol}(\tilde{M}_k)$ obtained in Corollary 5.1. Then the coefficient in (3.7) is given by

$$\begin{aligned} w_{d+1-e} &= \frac{1}{\Omega_{d+1-e}\Omega_{p-d-1+e}} \cdot \mathrm{Vol}(\tilde{M}_k) \cdot \frac{\Omega_{p-d-1}}{E[(\chi^2_{p-d-1})^{e/2}]} E[\mathrm{tr}_e H(u,v)] \\ &= \frac{\mathrm{Vol}(\tilde{M}_k)}{\Omega_{d+1}} \cdot \frac{\Gamma(\frac{1}{2}(d+1-e))}{2^{e/2}\,\Gamma(\frac{1}{2}(d+1))} E[\mathrm{tr}_e H(u,v)]. \end{aligned}$$

Theorem 5.1 *The nonzero coefficient* w_{d+1-e} *in (3.7) for the tail probabilities of* $\tilde{T}_k$ *in (1.3) and* $\tilde{U}_k$ *in (1.4) is given by*

$$w_{d+1-e} = w_{q-e} = \frac{k^{\frac{1}{2}(q-1)}}{2^{(k-1) \bmod 2}} \cdot \Big(-\frac{k-1}{4k}\Big)^{e/2} \frac{\Gamma(\frac{1}{2}(q-e))\,\Gamma(q)}{\Gamma(\frac{1}{2}q)\,\Gamma(q-e)\,(e/2)!},$$

$e = 0, 2, \ldots, [(q-1)/2] \times 2$.

5.3 Critical radius

In this subsection we obtain the concrete value of the critical radius $\tilde{\theta}_c$ of the manifold $\tilde{M}_k$ in (5.1) by virtue of Lemma 3.1.

Fix a point $v = h \otimes \cdots \otimes h \in \tilde{M}_k$ with $h \in S^{q-1}$. Let H be a $q \times (q-1)$ matrix such that (h, H) is $q \times q$ orthogonal. Let $\tilde{K}_k = \bigcup_{c \geq 0} c\tilde{M}_k$ be the cone associated with $\tilde{M}_k$. Then the tangent space $T_v(\tilde{K}_k)$ at v is spanned by $v = h \otimes \cdots \otimes h$ and the column spaces of

$$B = \sum_{l=1}^{k} \underbrace{h \otimes \cdots \otimes h}_{l-1} \otimes H \otimes \underbrace{h \otimes \cdots \otimes h}_{k-l}.$$

The orthogonal projection matrix onto $T_v(\tilde{K}_k)$ is easily shown to be

$$P_v = vv' + \frac{1}{k} BB'.$$

Let $\tilde{v} = \tilde{h} \otimes \cdots \otimes \tilde{h} \in \tilde{M}_k$. Then $\tilde{v}'v = (\tilde{h}'h)^k$, $B'\tilde{v} = k(\tilde{h}'h)^{k-1} H'\tilde{h}$, and

$$\begin{aligned} \tilde{v}' P_v \tilde{v} &= (\tilde{v}'v)^2 + \frac{1}{k}(B'\tilde{v})'(B'\tilde{v}) \\ &= (\tilde{h}'h)^{2k} + k(\tilde{h}'h)^{2(k-1)} \tilde{h}' H H' \tilde{h} \\ &= k(\tilde{h}'h)^{2(k-1)} - (k-1)(\tilde{h}'h)^{2k}. \end{aligned}$$

Put $x = \tilde{h}'h$. Then by Lemma 3.1 we have

$$\cot^2 \theta_c = \sup_{\tilde{v}, v \in \tilde{M}_k} \frac{1 - \tilde{v}' P_v \tilde{v}}{(1 - \tilde{v}'v)^2} = \sup_{-1 < x < 1} \frac{1 - kx^{2(k-1)} + (k-1)x^{2k}}{(1 - x^k)^2} = \frac{2(k-1)}{k}.$$

The last equality follows from Lemma 4.6.

Theorem 5.2 *The critical radius* $\tilde{\theta}_c$ *of* $\tilde{M}_k$ *in (5.1) is given by*

$$\cos^2 \tilde{\theta}_c = \frac{2k-2}{3k-2}, \quad k \geq 2.$$

5.4 An example: the largest eigenvalue of 4×4 symmetric normal distribution

Let A be distributed as $q\times q$ multivariate symmetric normal distribution in (1.5). The asymptotic series for the tail probability of the largest eigenvalue $\tilde{T}_2=\lambda_1(A)$ is given by Theorem 5.1. For example, for $q=4$ we have

$$P(\tilde{T}_2\geq x)\sim\sqrt{2}\bar{G}_4(x^2)-\frac{3}{2\sqrt{2}}\bar{G}_2(x^2). \tag{5.9}$$

In this case the exact distribution function can be given by

$$\begin{aligned} P(\tilde{T}_2\leq x) &= \Phi(\sqrt{2}x)^2-\frac{\sqrt{\pi}}{2}(2x^2+1)\phi(x)\Phi(x)\Phi(\sqrt{2}x) \\ &\quad -\sqrt{2}x\phi(\sqrt{2}x)\Phi(\sqrt{2}x)-\frac{1}{2}x\phi(\sqrt{3}x)\Phi(x)-\frac{1}{\sqrt{2\pi}}\phi(2x), \end{aligned} \tag{5.10}$$

where $\Phi(x)$ and $\phi(x)$ denote the cumulative distribution function and the density function of $N(0,1)$, respectively (Kuriki (1993)).

In Figure 5.1 the approximate tail probability (5.9) as well as the corresponding exact value by (5.10) are plotted by solid lines. (A monotonically decreasing curve corresponds to the exact value.) We can conclude that the asymptotic formula (5.9) is accurate enough when the tail probability is around 0.3 or less.

The lower and upper bounds Q_L and Q_U are also plotted by broken lines in Figure 5.1. Note that $p'=q(q+1)/2=10$ is used instead of $p=q^2=16$. Unfortunately, unlike Figures 4.1 and 4.2, the bounds given by Q_L and Q_U are not very satisfactory. In this case $\mathrm{Vol}(\tilde{M}_c)/\Omega_{10}\doteq 0.376$ is not close to 1 and

$$Q_U(x)-Q_L(x)=\bar{G}_{10}(2x^2)\,(1-\mathrm{Vol}(\tilde{M}_c)/\Omega_{10})$$

are relatively large.

A Critical radius and local radius of curvature

Here we investigate the relation between the global critical radius and the local radius of curvature. In Section 3 we considered the tube of $M\subset S^{p-1}$ with respect to the geodesic distance of S^{p-1} . For clarity and completeness of argument we first consider the tube in R^p with respect to the ordinary Euclidean distance. It will be shown that geodesic curvature of M is closely related to the curvature of the cone $K=\bigcup_{c\geq 0}cM$.

Let N be a non-bordered compact C^2-submanifold of dimension d in R^p . The tube around N with radius ρ is defined as

$$N_\rho=\{y\mid \|y-y_N\|<\rho\},$$

where y_N is the projection of y onto N. As in Section 3 for $x \in N$ we define

$$C_\rho(x) = \{x + y \mid y \in T_x(N)^\perp, \ \|y\| < \rho\}$$

where $T_x(N)^\perp$ denotes the orthogonal complement of the tangent space of N at x. Then $N_\rho = \bigcup_{x\in N} C_\rho(x)$. It is said that N_ρ does not have self-overlap if $C_\rho(x), \ x \in N$, are disjoint. The critical radius ρ_c of N is defined as

$$\rho_c = \sup\{\rho \mid N_\rho \text{ does not have self-overlap}\}.$$

Note that if $N \subset S^{p-1}$ then $N_\rho \cap S^{p-1}$ is a tube of N with respect to the geodesic distance of S^{p-1}. The problem is that N_ρ may have self-overlap in R^p even if $N_\rho \cap S^{p-1}$ does not have self-overlap in S^{p-1}. For this reason we make distinction between tube with respect to Euclidean distance and tube with respect to the geodesic distance on S^{p-1}.

The following lemma (Proposition 4.1 of Johansen and Johnstone (1990)) holds for the case $\dim N = d > 1$. We omit the proof because of the same reason as Lemma 3.1.

Lemma A.1 *The critical radius ρ_c of N is given by*

$$\rho_c = \inf_{x,y\in N} \frac{\|x-y\|^2}{2\|P_y^\perp(x-y)\|}, \tag{A.1}$$

where $P_y^\perp$ is the orthogonal projection onto the orthogonal complement of the tangent space $T_y(N)$ of N at y.

Here we discuss the property of

$$h(x,y) = \frac{2\|P_y^\perp(x-y)\|}{\|x-y\|^2} \tag{A.2}$$

appearing in (A.1). Since $P_y^\perp$ is continuous in y, $h(x,y)$ is continuous on $\{(x,y) \in N \times N \mid x \neq y\}$. Then we investigate the behavior of $h(x,y)$ as $\|x-y\| \to 0$. Since we are considering local property of N we can take d-dimensional local coordinates $t = (t^1, \ldots, t^d)$ and express x, y in terms of t. For the sake of convenience we use the Einstein convention of indices.

Write $y = \phi(t)$ and $x = \phi(t + dt)$. Then

$$\|x - y\| = \|\phi(t+dt) - \phi(t)\|^2 = g_{ij}dt^i dt^j + o(\|dt\|^2),$$

where

$$g_{ij} = \Big(\frac{\partial\phi}{\partial t^i}\Big)' \frac{\partial\phi}{\partial t^j}, \quad i,j = 1,\ldots,d,$$

are the elements of the first fundamental form at $y = \phi(t)$. On the other hand

$$\begin{aligned} P_y^\perp(\phi(t+dt) - \phi(t)) &= P_y^\perp \frac{\partial\phi}{\partial t^i}dt^i + \frac{1}{2}P_y^\perp \frac{\partial^2\phi}{\partial t^i \partial t^j}dt^i dt^j + o(\|dt\|^2) \\ &= \frac{1}{2}P_y^\perp \frac{\partial^2\phi}{\partial t^i \partial t^j}dt^i dt^j + o(\|dt\|^2) \end{aligned}$$

and

$$2\|P_y^{\perp}(\phi(t+dt)-\phi(t))\| = \left\|P_y^{\perp}\frac{\partial^2\phi}{\partial t^i \partial t^j}dt^i dt^j\right\| + o(\|dt\|^2).$$

Let

$$w^* \propto -P_y^{\perp}\frac{\partial^2\phi}{\partial t^i \partial t^j}dt^i dt^j$$

such that $\|w^*\| = 1$. (If the right hand side is the zero vector, let $w^* = 0$.) Then

$$2\|P_y^{\perp}(\phi(t+dt)-\phi(t))\| = H_{ij}(w^*)dt^i dt^j + o(\|dt\|^2),$$

where

$$H_{ij}(w) = -w'\frac{\partial^2\phi}{\partial t^i \partial t^j}, \quad i,j = 1,\ldots,d.$$

Therefore we have

$$h(x,y) = \frac{H_{ij}(w^*)dt^i dt^j}{g_{ij}dt^i dt^j} + o(\|dt\|^2).$$

The $d \times d$ matrix with (i,j)-th element $H_i^j(w) = H_{ik}(w)g^{kj}$ is called the second fundamental form of N at y with respect to the direction w . The eigenvalues of the second fundamental form are called the principal curvatures and their associated eigenvectors are called principal directions. Note that w^* depends on dt through the direction $dt/\|dt\|$.

Fix dt . Then

$$\begin{aligned} H_{ij}(w^*)dt^i dt^j &= -w^{*\prime}\frac{\partial^2\phi}{\partial t^i \partial t^j}dt^i dt^j = -w^{*\prime}P_y^{\perp}\frac{\partial^2\phi}{\partial t^i \partial t^j}dt^i dt^j \\ &= \max_{\|w\|=1}\Big(-w'P_y^{\perp}\frac{\partial^2\phi}{\partial t^i \partial t^j}dt^i dt^j\Big) \\ &= \max_{w\in T_y(N)^{\perp},\ \|w\|=1}\Big(-w'\frac{\partial^2\phi}{\partial t^i \partial t^j}dt^i dt^j\Big) \\ &= \max_{w\in T_y(N)^{\perp}\cap S^{p-1}} H_{ij}(w)dt^i dt^j. \end{aligned}$$

Taking maximum with respect to the direction $dt/\|dt\|$ we have

$$\begin{aligned} \limsup_{x\to y} h(x,y) &= \max_{\|dt\|=1}\frac{H_{ij}(w^*)dt^i dt^j}{g_{ij}dt^i dt^j} \\ &= \max_{w\in T_y(N)^{\perp}\cap S^{p-1}}\ \max_{\|dt\|=1}\frac{H_{ij}(w)dt^i dt^j}{g_{ij}dt^i dt^j} \\ &= \max_{w\in T_y(N)^{\perp}\cap S^{p-1}} |\lambda_{\max}(w)|, \end{aligned}$$

where $|\lambda_{\max}(w)|$ denotes the principal curvature having the largest absolute value. $1/|\lambda_{\max}(w)|$ is the local radius of curvature at y with respect the direction $\pm w$.

Write

$$h(y,y) = \limsup_{x\to y} h(x,y)$$

so that $h(x,y)$ is defined and finite for all $(x,y) \in N \times N$. By continuity of the radius of curvature it is easy to see that as $x, z \to y$

$$h(y,y) = \limsup_{x,\, z \to y} h(x,z).$$

Now by simple compactness argument h attains a finite maximum over $N \times N$. To prove this let $(x_i, y_i),\ i = 1, 2, \dots,$ be a sequence of points of $N \times N$ such that $h(x_i, y_i) \uparrow \bar{h} = \sup_{x,y \in N} h(x,y)$. By compactness we can assume without loss of generality that $(x_i, y_i) \to (x_0, y_0)$. If $x_0 \neq y_0$ then $h(x_0, y_0) = \bar{h}$ by continuity. If $x_0 = y_0$ then $h(x_0, y_0) = \limsup_{(x,y)\to(x_0,y_0)} h(x,y) \geq \lim_{i\to\infty} h(x_i, y_i) = \bar{h}$. However obviously $h(x_0, y_0) \leq \bar{h}$. This proves that h attains a finite maximum over $N \times N$, and hence the critical radius ρ_c is positive under our assumptions.

So far we considered the tube with respect to the Euclidean distance. We proceed to discuss the tube in the unit sphere S^{p-1} with respect to the geodesic distance. $h(u,v)$ in (3.4) can be written as

$$h(u,v) = \frac{\sqrt{1 - u' P_v u}}{1 - u'v} = \frac{2\|P_v^{\perp}(u - v)\|}{\|u - v\|^2},$$

which is identical to $h(x,y)$ in (A.2) with N replaced with K except that u is restricted to $M \subset S^{p-1}$. However as $u \to v$, $(u-v)/\|u-v\|$ becomes orthogonal to v. On the other hand since K is a cone, one of the principal direction of K at v is v itself and other principal directions are orthogonal to v. Therefore the calculation involving the second fundamental form of M at $v \in M$ can be replaced with the calculation of second fundamental form of K at $v \in K$. In particular $h(v,v) = \limsup_{u \to v} h(u,v)$ is similarly defined and $h(u,v)$ attains a finite maximum over $M \times M$. This proves the claims of Remark 3.1.

B Proof of Lemma 3.3 and Theorem 3.1

Let z be distributed as $N_p(0, I_p)$, and let $r = \|z_K\|$, $s = \|z - z_K\|$. By (3.6)

$$\begin{aligned} P(z \in K_\theta) &= P(s < r \tan\theta) \\ &= \frac{1}{(2\pi)^{p/2}} \sum_{e=0}^{d} \iint_{0 \leq s < r\tan\theta} e^{-\frac{1}{2}(r^2+s^2)} r^{d-e} s^{p-d-2+e} dr ds \\ &\qquad\qquad \times \int_M \Big[\int_{T_u(K)^{\perp} \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv \Big]\, du. \end{aligned}$$

By a simple change of variables we obtain

$$\begin{aligned} &\iint_{0 \leq s < r\tan\theta} e^{-\frac{1}{2}(r^2+s^2)} r^{d-e} s^{p-d-2+e}\, dr\, ds \\ &\qquad = \bar{B}_{\frac{1}{2}(d+1-e), \frac{1}{2}(p-d-1+e)}(\cos^2\theta) \cdot 2^{p/2-2}\, \Gamma\Big(\frac{d+1-e}{2}\Big)\, \Gamma\Big(\frac{p-d-1+e}{2}\Big). \end{aligned}$$

Note that

$$\int_{T_u(K)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv = 0$$

if e is odd since $\mathrm{tr}_e H(u,v)$ is an odd degree polynomial in the elements of v. This proves Lemma 3.3.

Now we proceed to the proof of Theorem 3.1.

$$P(T \geq a) = P(T \geq a,\ z \in K_{\theta_c}) + P(T \geq a,\ z \notin K_{\theta_c}).$$

We bound the second term on the right hand side from above. Note that the projection z_K always exists and we can write

$$z = r\frac{z_K}{\|z_K\|} + s\frac{z - z_K}{\|z - z_K\|}$$

and $z \in K_{\theta_c}$ if and only if

$$s < r \tan \theta_c.$$

Since $r = \max(T, 0)$, we have for $z \notin K_{\theta_c}$ and $T \geq 0$

$$\|z\|^2 = r^2 + s^2 \geq r^2(1 + \tan^2 \theta_c) \geq T^2(1 + \tan^2 \theta_c).$$

Therefore for $a > 0$

$$\begin{aligned} P(T \geq a,\ z \notin K_{\theta_c}) &\leq P(\|z\|^2 \geq a^2(1 + \tan^2 \theta_c),\ z \notin K_{\theta_c}) \\ &= \bar{G}_p(a^2(1 + \tan^2 \theta_c))\, P(z \notin K_{\theta_c}) \end{aligned}$$

and

$$P(T \geq a) \leq P(T \geq a, z \in K_{\theta_c}) + \bar{G}_p(a^2(1 + \tan^2 \theta_c))\, P(z \notin K_{\theta_c}).$$

Furthermore

$$P(T \geq a) \geq P(T \geq a, z \in K_{\theta_c}).$$

Therefore it remains to show that $P(T \geq a, z \in K_{\theta_c})$ for $a > 0$ can be written as $Q_L(a)$ of (3.10). Now

$$\begin{aligned} P(T \geq a, z \in K_{\theta_c}) &= \frac{1}{(2\pi)^{p/2}} \sum_{\substack{e=0 \\ e:\text{even}}}^{d} \int\int_{\substack{a \leq r < \infty \\ 0 \leq s < r \tan\theta}} e^{-\frac{1}{2}(r^2+s^2)} r^{d-e} s^{p-d-2+e} dr\, ds \\ &\qquad \times \int_M \Big[\int_{T_u(K)^\perp \cap S^{p-1}} \mathrm{tr}_e H(u,v)\, dv \Big]\, du. \end{aligned}$$

Integrating the right hand side with respect to s first we see that $P(T \geq a, z \in K_{\theta_c}) = Q_L(a)$. This proves the theorem.

C Proof of Lemma 4.6

Let $f(z) = 1 - kz^{2(k-1)} + (k-1)z^{2k}$ and $g(z) = (1-z^k)^2$ be the numerator and denominator of the argument of the supremum in (4.16). When $k=2$, $f(z) \equiv g(z)$ and the statement holds trivially. Consider the case $k \geq 3$. We claim that

$$\frac{d}{dz}\Big(\frac{f(z)}{g(z)}\Big) > 0 \quad \text{for } 0 < z < 1\,. \tag{C.1}$$

In fact, simple calculation yields that

$$\frac{d}{dz}\Big(\frac{f(z)}{g(z)}\Big) = \frac{2k(1-z^k)z^{k-1}}{g(z)^2} \cdot h(z),$$

where

$$\begin{aligned} h(z) &= 1-(k-1)z^{k-2}+(k-1)z^k - z^{2k-2} \\ &= (1-z^2)\{1+z^2+\cdots+(z^2)^{k-2}-(k-1)z^{k-2}\}. \end{aligned}$$

By the convexity of the map $\xi \mapsto (z^2)^\xi$, we have

$$\frac{1+z^2+\cdots+(z^2)^{k-2}}{k-1} \geq (z^2)^{\frac{0+1+\cdots+(k-2)}{k-1}} = |z|^{k-2},$$

and the equality holds if and only if $|z| = 1$. Therefore $h(z) > 0$ for $|z| < 1$, which implies (C.1). Therefore we have the supremum in (4.16) as

$$\sup_{0\leq z<1}\frac{f(z)}{g(z)} = \lim_{z\uparrow 1}\frac{f(z)}{g(z)} = \lim_{z\uparrow 1}\frac{\frac{d^2}{dz^2}f(z)}{\frac{d^2}{dz^2}g(z)} = \frac{2(k-1)}{k}.$$

D Proof of (4.13)

In order to prove (4.13) we prepare a lemma.

Lemma D.1 *For $a > 0$ and $n = 0, 1, 2, \ldots,$ define*

$$I_n(a) = \sum_{k=0}^{n}(-1)^{n-k}\,2^k\,\frac{\Gamma(\frac{a}{2}+k)}{\Gamma(a+k)}\binom{n}{k}$$

and

$$J_n(a) = \sum_{k=0}^{n}(-1)^{n-k}\,2^k\,\frac{\Gamma(\frac{a}{2}+k+1)}{\Gamma(a+k+1)}\binom{n}{k}.$$

Then

$$I_n(a) = \begin{cases} c_n(a) & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd} \end{cases} \tag{D.1}$$

and

$$J_n(a) = \begin{cases} \frac{1}{2}\, c_n(a) & \text{if } n \text{ is even} \\ \frac{1}{2}\, c_{n+1}(a) & \text{if } n \text{ is odd} \end{cases} \tag{D.2}$$

where

$$c_n(a) = 2^n \frac{\Gamma(\frac{n+1}{2})\,\Gamma(\frac{n+a}{2})}{\sqrt{\pi}\,\Gamma(n+a)} = \frac{\Gamma(\frac{n+1}{2})}{2^{a-1}\,\Gamma(\frac{n+a+1}{2})}.$$

Proof. We use induction on n. The claims (D.1) and (D.2) are easily checked for $n = 0, 1$. Assume that they are true for $n-1$ and n.

Making a use of the identity $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$, we have

$$\begin{aligned} I_{n+1}(a) &= \sum_{k=0}^{n+1} (-1)^{n+1-k}\, 2^k \frac{\Gamma(\frac{a}{2}+k)}{\Gamma(a+k)} \left\{ \binom{n}{k} + \binom{n}{k-1} \right\} \\ &= -I_n(a) + 2J_n(a) \\ &= \begin{cases} 0 & \text{if } n \text{ is even} \\ c_{n+1}(a) & \text{if } n \text{ is odd.} \end{cases} \end{aligned} \tag{D.3}$$

Similarly we have

$$J_{n+1}(a) = -J_n(a) + 2\sum_{k=0}^{n} (-1)^{n-k}\, 2^k \frac{\Gamma(\frac{a}{2}+k+2)}{\Gamma(a+k+2)} \binom{n}{k}.$$

Noting that

$$\frac{\Gamma(\frac{a}{2}+k+2)}{\Gamma(a+k+2)}\binom{n}{k} = \frac{\Gamma(\frac{a+2}{2}+k)}{\Gamma((a+2)+k)} \left\{ \frac{a+2}{2}\binom{n}{k} + n\binom{n-1}{k-1} \right\},$$

we have

$$\begin{aligned} J_{n+1}(a) &= -J_n(a) + (a+2)\, I_n(a+2) + 4n\, J_{n-1}(a+2) \\ &= \begin{cases} \frac{1}{2}\, c_{n+2}(a) & \text{if } n \text{ is even} \\ \frac{1}{2}\, c_{n+1}(a) & \text{if } n \text{ is odd.} \end{cases} \end{aligned} \tag{D.4}$$

(D.3) and (D.4) imply that (D.1) and (D.2) hold for any $a > 0$ and $n \geq 2$. The proof is completed. ∎

The relation (4.13) is equivalent with (D.1) with $n = q-1$, $k = q-1-j$, and $a = \nu - q + 1$.

References

[1] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. John-Wiley & Sons, New York.

[2] Baringhaus, L. and Henze, N. (1991). Limit distributions for measures of multivariate skewness and kurtosis based on projections. *J. Multivariate Anal.*, **38**, 51–69.

[3] Davis, A. W. (1972). On the ratios of the individual latent roots to the trace of a Wishart matrix. *J. Multivariate Anal.*, **2**, 440–443.

[4] Hanumara, R. C. and Thompson, Jr., W. A. (1968). Percentage points of the extreme roots of a Wishart matrix. *Biometrika*, **55**, 505–512.

[5] Hotelling, H. (1939). Tubes and spheres in n-spaces, and a class of statistical problems. *Am. J. Math.*, **61**, 440–460.

[6] Johansen, S. and Johnstone, I. (1990). Hotelling's theorem on the volume of tubes: Some illustrations in simultaneous inference and data analysis. *Ann. Statist.*, **18**, 652–684.

[7] Johnson, D. E. and Graybill, F. A. (1972). An analysis of a two-way model with interaction and no replication. *J. Amer. Statist. Assoc.*, **67**, 862–868.

[8] Kawasaki, H. and Miyakawa, M. (1996). A test of three-factor interaction in a three-way layout without replication. *Quality, JSQC*, **26**, 97–108 (in Japanese).

[9] Knowles, M. and Siegmund, D. (1989). On Hotelling's approach to testing for a nonlinear parameter in regression. *Internat. Statist. Rev.*, **57**, 205–220.

[10] Kuriki, S. (1993). Likelihood ratio tests in multivariate variance components models. Doctoral dissertation, Faculty of Engineering, Univ. of Tokyo.

[11] Kuriki, S. and Takemura, A. (1997). James-Stein type estimator by shrinkage to closed convex set with smooth boundary. *Discussion Paper 97-F-22*, Faculty of Economics, Univ. of Tokyo (submitted for publication).

[12] Machado, S. G. (1983). Two statistics for testing for multivariate normality. *Biometrika*, **70**, 713–718.

[13] Malkovich, J. F. and Afifi, A. A. (1973). On tests for multivariate normality. *J. Amer. Statist. Assoc.*, **68**, 176–179.

[14] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. John-Wiley & Sons, New York.

[15] Pillai, K. C. S. (1976). Distributions of the characteristic roots in multivariate analysis: Part I, Null distribution. *Canad. J. Statist. Sec. A & B*, **4**, 157–184.

[16] Schuurmann, F. J., Krishnaiah, P. R., and Chattopadhyay, A. K. (1973). On the distribution of the ratios of the extreme roots to the trace of the Wishart matrix. *J. Multivariate Anal.*, **3**, 445–453.

[17] Sun, J. (1991). Significance levels in exploratory projection pursuit. *Biometrika*, **78**, 759–769.

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

[19] Takemura, A. (1993). Maximally orthogonally invariant higher order moments and their application to testing elliptically-contouredness. In *Statistical Science & Data Analysis* (Editors K. Matsushita et al.), 225–235, VSP, Utrecht.

[20] Takemura, A. and Kuriki, S. (1997). Weights of $\bar{\chi}^2$ distribution for smooth or piecewise smooth cone alternatives. *Ann. Statist.*, **25**, 2368–2387.

[21] Weyl, H. (1939). On the volume of tubes. *Am. J. Math.*, **61**, 461–472.

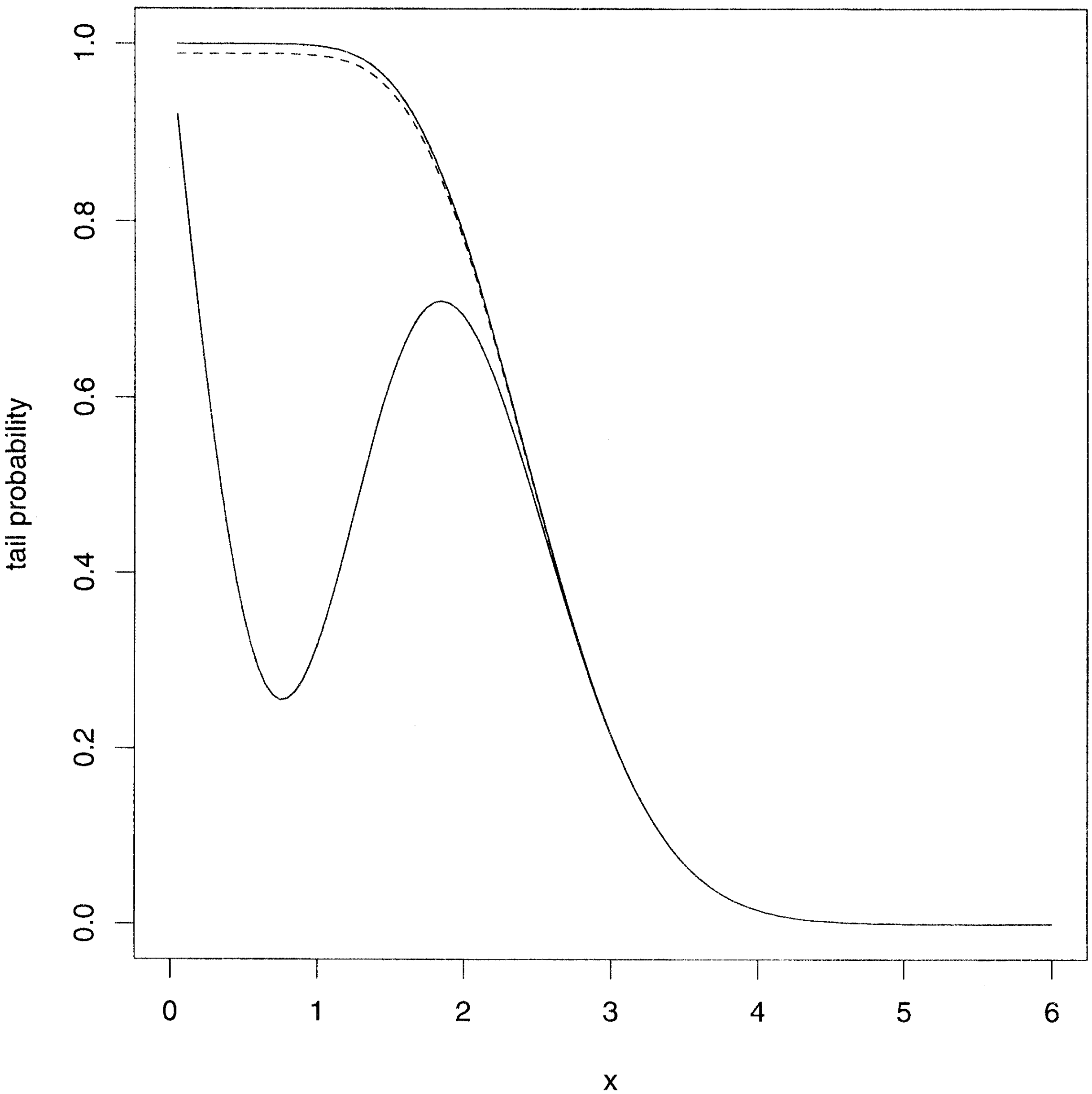

Figure 4.1 2-form (3×3)

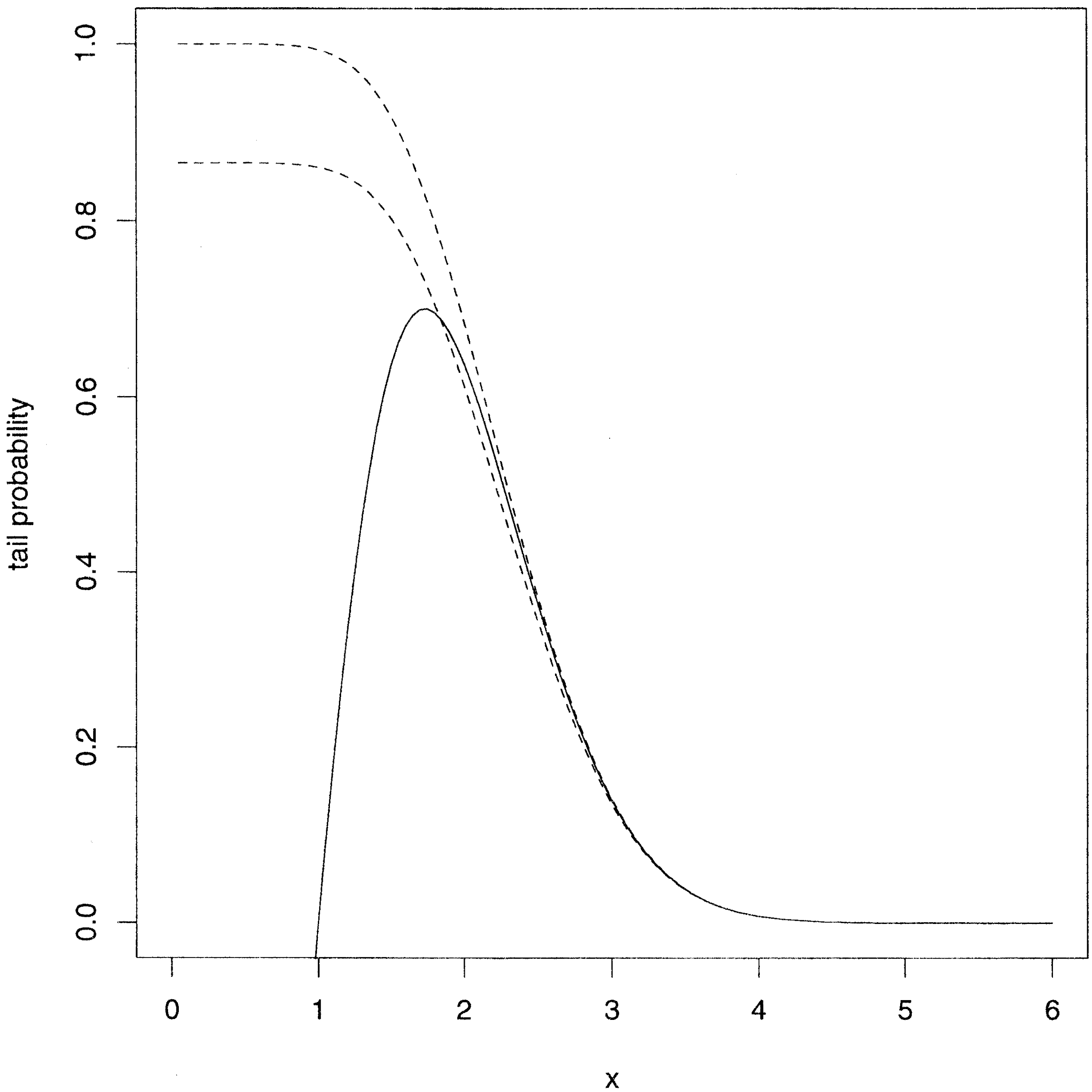

Figure 4.2 3-form $(2 \times 2 \times 2)$

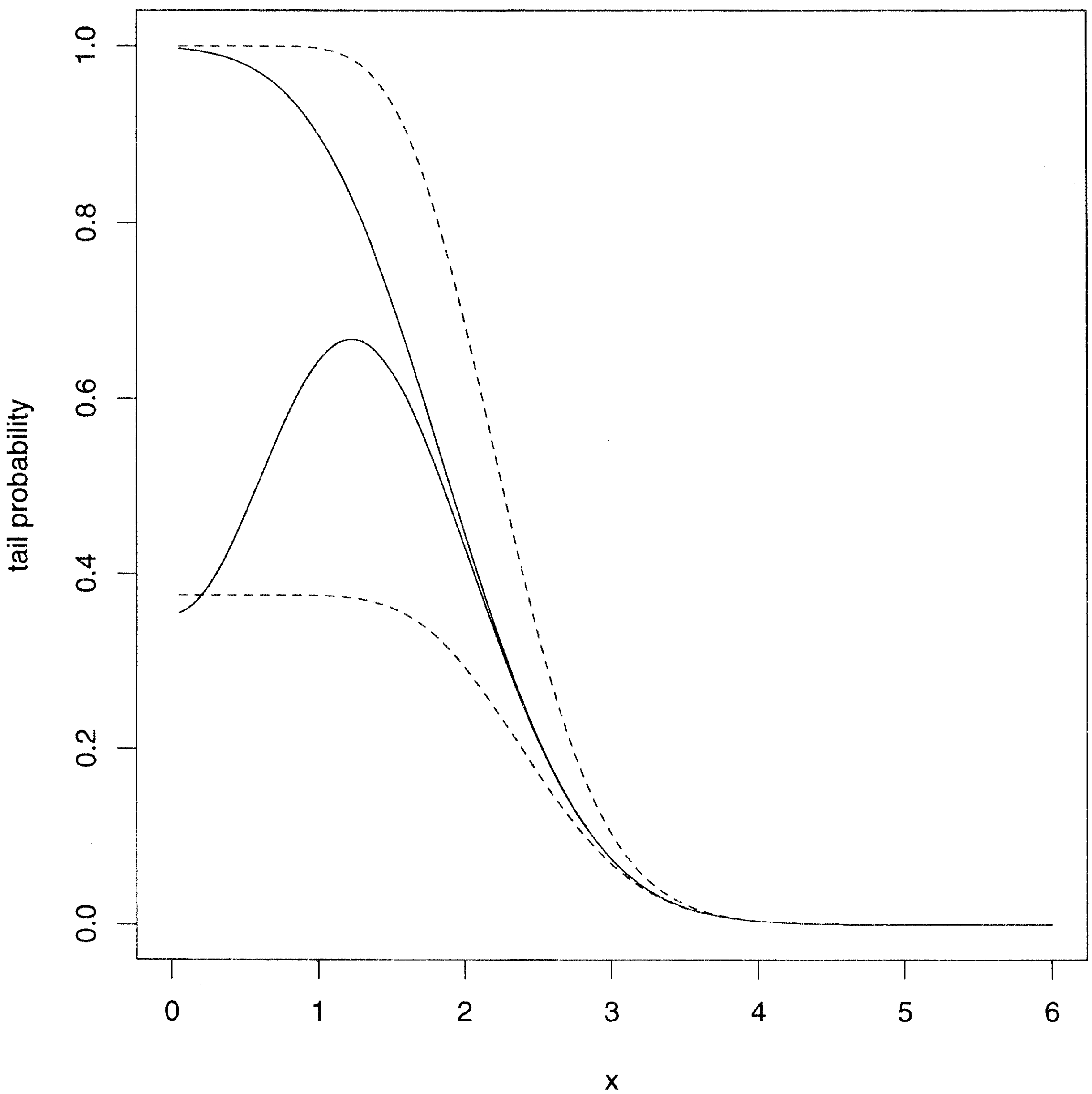

Figure 5.1 Symmetric 2-form (4×4)