Application of tube formula to distributional problems in multiway layouts

Satoshi Kuriki
Institute of Statistical Mathematics
and
Akimichi Takemura
Graduate School of Information Science and Technology
University of Tokyo

October, 2001

Abstract

Recently, an integral geometric method called the tube method has been actively developed. The tube method gives us a powerful tool to tackle problems where conventional matrix theory such as the singular value decomposition cannot be applied. The aim of this paper is to survey several recent applications of the tube method to distributional problems in multiway layouts mainly by the authors. Null distributions of test statistics of the following four testing problems are discussed: (i) A test for interaction in three-way layout based on the three-way analogue of the largest singular value. (ii) A test for multivariate normality by searching a nonnormal direction proposed by Malkovich and Afifi. (iii) Testing independence in ordered categorical data by maximizing row and column scores under order restriction. (iv) Detecting a change point in two-way layout with ordinal factors.

Key words: change point, Gaussian random field, integral geometry, multilinear model, ordered categorical data, testing multivariate normality.

1 Introduction.

Standard results of matrix theory play a major role in conventional multivariate statistical analysis or categorical data analysis. In particular when data are summarized as a two-way table, many methods of data analysis have been based on the singular value decomposition of the data matrix. However data may not be summarized as a single matrix. In a factorial design or cross-classification, data are usually obtained as a multiway layout. Moreover even when data are summarized as a two-way table, usual matrix theory is not necessarily applicable. For example, in a two-way cross-classified table with ordinal row and column categories, ordinary matrix methods invariant with respect to permutations of rows or columns are not suitable.

Recently, an integral geometric method called the tube method has been actively developed. This method originates from Hotelling [10] and Weyl [32]. Sun [27] showed how the tube method can be used for deriving distributions of maxima of Gaussian random fields. A closely related technique is the Euler characteristic method developed mainly by Adler and Worsley. A detailed review of the Euler characteristic method and its relation to the tube method is given in Adler [1]. The Euler characteristic method has been extensively used for analyzing brain image data (e.g., Worsley [33]). In Takemura and Kuriki [30] we have established the equivalence of the two methods using an extended form of the Morse theorem (see also Kuriki and Takemura [15]).

Initially we have been investigating the coefficients of the $\bar{\chi}^2$ distribution appearing in order restricted inference (Kuriki [13], Takemura and Kuriki [28], Kuriki and Takemura [16]) from geometric viewpoint. During this investigation we have recognized that the tube method leads to approximation of the upper tail probability of maximum type statistics in the form of linear combination of χ^2 distributions, which can be regarded as a generalization of the $\bar{\chi}^2$ distribution. From this view point we applied the tube method to many distributional problems in conventional multivariate analysis, where matrix theory cannot be applied (Kuriki [14], Kuriki and Takemura [15], [17], [18], Takemura and Kuriki [30], [31]).

In this paper we review several applications of the tube method to distributional problems in multiway layouts by the authors and by Ninomiya [23], [24]. In Section 2 we give a brief introduction to the tube method. In Sections 3–6 we review four testing problems: (i) A test for interaction in three-way layout based on the three-way analogue of the largest singular value. (ii) A test for multivariate normality by searching a nonnormal direction proposed by Malkovich and Afifi [20]. (iii) Testing independence in ordered categorical data by maximizing row and column scores under order restriction. (iv) Detecting a change point in two-way layout with ordinal factors.

2 A brief introduction to the tube method.

Let $z = (z_1, \ldots, z_n)'$ be an n-dimensional vector consisting of independent standard random variables $N(0, 1)$. Let S^{n-1} be the unit sphere in R^n, and let $M \subset S^{n-1}$ be a C^2-submanifold of dimension $d = \dim M$ with piecewise smooth boundaries. Define a random field with index set M by

$$Z(u) = u'z = \sum_{i=1}^{n} u_i z_i, \qquad u = (u_1, \ldots, u_n)' \in M.$$

$Z(u)$ is the length of the orthogonal projection of z onto the half line joining the origin and $u \in M$ (see Figure 1).

`--- Figure 1 around here ---`

Note that $Z(u)$ is a continuous Gaussian random field on M with mean 0, variance 1, and covariance function $\mathrm{cov}[Z(u), Z(v)] = u'v$. We also define the standardized random field

$$Y(u) = u'z/\|z\|, \qquad u \in M,$$

where $\|z\| = \sqrt{z'z}$.

Consider distributions of the maxima of $Z(u)$ and $Y(u)$:

$$T = \max_{u \in M} Z(u), \qquad U = \max_{u \in M} Y(u). \tag{2.1}$$

It is in general difficult to derive exact distribution functions of T and U. However, approximations for upper tail probabilities

$$P(T \geq x), \quad x \uparrow \infty, \qquad \text{and} \qquad P(U \geq x), \quad x \uparrow 1,$$

are obtained by the tube method explained below.

The subset of S^{n-1} consisting of points whose geodesic distance from $M \subset S^{n-1}$ is less or equal to θ,

$$M_\theta = \Big\{ v \in S^{n-1} \mid \min_{u \in M} \cos^{-1}(u'v) \leq \theta \Big\},$$

is called the *tube* around M with radius θ (see Figure 2).

`--- Figure 2 around here ---`

The spherical projection point of $v \in M_\theta$ onto M, i.e., the point which attains the minimum $\min_{u \in M} \cos^{-1}(u'v)$, is denoted by v_M. For a sufficiently small $\theta > 0$, each $v \in M_\theta$ has the unique projection v_M. The supremum θ_c of such θ is called *critical radius* of M. θ_c can be proved to be positive under assumptions of compactness and local convexity of M.

The $(n-1)$-dimensional spherical volume of M_θ is denoted by $\mathrm{Vol}(M_\theta)$. Let

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

be the volume of the unit sphere S^{n-1}. Since $z/\|z\|$ is distributed uniformly on the unit sphere, it holds by definition that

$$\begin{aligned} \mathrm{Vol}(M_\theta)/\Omega_n &= P\Big(\min_{u \in M} \cos^{-1}(u'z/\|z\|) \leq \theta\Big) \\ &= P\Big(\max_{u \in M} u'z/\|z\| \geq \cos\theta\Big) = P\Big(\max_{u \in M} Y(u) \geq \cos\theta\Big). \end{aligned}$$

The tube formula expresses $\mathrm{Vol}(M_\theta)$ in terms the *geometric invariants* $w_1, \ldots, w_{d+1}$ of the manifold M. The definition of the geometric invariants is given in (A.1) in the Appendix A.1. Let $\bar{B}_{m,n}(\cdot)$ denote the upper probability of beta distribution with parameter (m, n). We state the tube formula as follows.

Theorem 2.1 *Assume that M has a positive critical radius $\theta_c > 0$. Then, for $\theta \le \theta_c$, the volume of tube is expressed as*

$$\mathrm{Vol}(M_\theta) = \Omega_n \Big\{ w_{d+1} \bar{B}_{\frac{d+1}{2},\frac{n-d-1}{2}}(\cos^2\theta) + w_d \bar{B}_{\frac{d}{2},\frac{n-d}{2}}(\cos^2\theta) + \cdots + w_1 \bar{B}_{\frac{1}{2},\frac{n-1}{2}}(\cos^2\theta) \Big\}. \quad (2.2)$$

Historically, this tube formula was originally proved by Hotelling [10] when M is 1-dimensional. His result was generalized to the multi-dimensional case by Weyl [32] when M is a closed manifold without boundary. In the case where M has boundaries, the formula was given in Hotelling [10] (1-dimensional case), Knowles and Siegmund [12] (2-dimensional case), and Naiman [22] (multi-dimensional case). Takemura and Kuriki [28] gives the essentially same formula as Naiman [22] in terms of the mixed volumes (the Minkowski functionals) for the case of geodesically convex M.

From the tube volume formula (2.2), it follows immediately that

$$P(U \ge x) = w_{d+1} \bar{B}_{\frac{d+1}{2},\frac{n-d-1}{2}}(x^2) + w_d \bar{B}_{\frac{d}{2},\frac{n-d}{2}}(x^2) + \cdots + w_1 \bar{B}_{\frac{1}{2},\frac{n-1}{2}}(x^2)$$

for $x \ge \cos^{-1}\theta_c$.

Noting the independence of $z/\|z\|$ and $\|z\|$, we have

$$P\Big(\max_{u\in M} Z(u) \ge x\Big) = P\Big(\max_{u\in M} u'z/\|z\| \ge x/\|z\|\Big) = E\Big[\mathrm{Vol}(M_{\cos^{-1}(x/\|z\|)})\Big]\Big/\Omega_n, \quad (2.3)$$

where $\mathrm{Vol}(M_{\cos^{-1}(x/\|z\|)}) = 0$ for $x/\|z\| > 1$.

If the expression of the volume formula $\mathrm{Vol}(M_\theta)$ in (2.2) were valid for all θ, the distribution of $\max_{u\in M} u'z$ could be obtained by substituting $\cos^2\theta := x^2/\|z\|^2$ into (2.2) and taking an expectation with respect to $\|z\|^2 \sim \chi^2(n)$ using the relation

$$E\Big[\bar{B}_{\frac{1}{2}j,\frac{1}{2}(n-j)}(x^2/\|z\|^2)\Big] = \bar{G}_j(x^2).$$

Here $\bar{G}_j(\cdot)$ denotes the upper probability of χ^2 distribution with j degrees of freedom. This formal manipulation does not yield the exact answer because the formula (2.2) is valid only for small θ. However if x is large, then $\cos^{-1}(x/\|z\|)$ in the right hand side of (2.3) is small, and this formal method is expected to give a correct answer in some sense. In fact Sun [27] showed the following.

Theorem 2.2 *As $x \to \infty$,*

$$P(T \ge x) = w_{d+1}\bar{G}_{d+1}(x^2) + w_d \bar{G}_d(x^2) + \cdots + w_1 \bar{G}_1(x^2) + o(e^{-x^2/2}). \quad (2.4)$$

Kuriki and Takemura [18] examined the reminder term $o(e^{-x^2/2})$ in (2.4) precisely. They proved that the reminder term is actually of the order of $O(\bar{G}_n(x^2(1+\tan^2\theta_c)))$. They have also provided a lower bound Q_L and an upper bound Q_U

$$Q_L(x) \le P(T \ge x) \le Q_U(x), \qquad x > 0.$$

These bounds Q_L, Q_U depend on the critical radius θ_c and the geometric invariants $w_1, \ldots, w_{d+1}$.

In general it is not easy to evaluate the geometric invariants $w_1, \ldots, w_{d+1}$. However from our experiences, it is often possible to determine the coefficients $w_1, \ldots, w_{d+1}$ when M is a well-known manifold or its variation (e.g., sphere, Stiefel manifold, Grassmann manifold, spherical polyhedron, *etc*). Numerical studies show that approximate tail probabilities by the tube method are sufficiently close to true values and hence one can conclude that the tube method is practical for the purpose of determining relevant significance levels (e.g., 10% or smaller). See Figures 3 and 4.

In the following we discuss four testing problems. The critical radius θ_c is positive for the first three problems. However $\theta_c = 0$ for the problem in Section 6. In this case the the validity of the tube formula approximation becomes more complicated. The problem of zero critical radius is investigated in Takemura and Kuriki [29], [31].

3 Tests for interaction in multiway layouts.

As a model for two-way layout without replication, Johnson and Graybill [11] proposed a model with interaction terms of a bilinear form:

$$x_{ij} = \alpha_i + \beta_j + \phi u_i v_j + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma^2), \quad i = 1, \ldots, I,\ j = 1, \ldots, J,$$

and constructed a likelihood ratio statistic for testing the null hypothesis $H_0 : \phi = 0$. The likelihood ratio test statistic is reduced to the largest singular value of a residual matrix under the null hypothesis.

An extension of this model was proposed by Boik and Marasinghe [4]. They modeled three-way interaction terms as a trilinear form:

$$\begin{gathered} x_{ijk} = (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \phi u_i v_j w_k + \varepsilon_{ijk}, \quad \varepsilon_{ijk} \sim N(0, \sigma^2), \\ i = 1, \ldots, I,\ j = 1, \ldots, J,\ k = 1, \ldots, K. \end{gathered}$$

This model of interaction is a particular case of the PARAFAC model which is extensively studied in psychometrics and chemometrics (Leurgans and Ross [19]). In this model testing the three way interaction is reduced to testing the null hypothesis $H_0 : \phi = 0$. Making suitable changes of variables, the null distribution of the likelihood ratio test statistic for testing H_0 is reduced to U^2 (if σ^2 is unknown) or T^2 (if $\sigma^2 = 1$ is known) with U or T in (2.1), where z is an $(I-1) \times (J-1) \times (K-1)$-dimensional random vector consisting of independent standard random variables $N(0, 1)$, and

$$u = h_1 \otimes h_2 \otimes h_3, \qquad h_1 \in S^{I-2}, \quad h_2 \in S^{J-2}, \quad h_3 \in S^{K-2}.$$

Here $\otimes$ denotes the tensor (Kronecker) product. This statistic is regarded as an extension of the singular value into a three-way layout.

The index set is a tensor product space of the unit spheres

$$M = S^{I-2} \otimes S^{J-2} \otimes S^{K-2},$$

and geometric quantities of M are derived from those of the unit spheres.

Theorem 3.1 *Let $d_1 = I-2$, $d_2 = J-2$, $d_3 = K-2$, and $d = \dim M = \sum_{i=1}^3 d_i$. The nonzero geometric invariants of M are given by*

$$w_{d+1-2m} = \frac{(-1/2)^m \, \pi \, \Gamma(\frac{1}{2}(d+1)-m)}{\prod_{i=1}^3 \Gamma(\frac{1}{2}(d_i+1))} \sum_{\substack{l_1+l_2+l_3=m \\ l_i \ge \max(m-d_i,0), \ \forall i}} \prod_{i=1}^3 \frac{d_i!}{l_i! \, (l_i-m+d_i)!},$$

$m = 0, 1, \ldots, [d/2]$. *The critical radius of M is $\theta_c = \cos^{-1}(2/\sqrt{7}) \doteq 0.227\pi$.*

This theorem is a corollary to Theorem 3.2 of Kuriki and Takemura [18]. A proof is given in the Appendix A.2.

For example when $I = J = K = 3$, the upper probability of U, which corresponds to a significance level of the likelihood ratio test with σ^2 unknown, is given by

$$P(U \ge x) = \pi \bar{B}_{2,2}(x^2) - \frac{3\pi}{2} \bar{B}_{1,3}(x^2)$$

for $x \ge 2/\sqrt{7}$. In Figure 3 we depict the upper tail probability of U by the tube formula and Monte Carlo simulations with 10,000 replications. One can confirm that the tube method gives the exact upper probability for $x \ge 2/\sqrt{7} \doteq 0.756$ at least.

`--- Figure 3 around here ---`

Although we only treated the case of three-way layout here, an extension to higher multiway layout is not difficult.

4 Tests for multivariate normality.

Assume that i.i.d. sample vectors $x_1, \ldots, x_n \in R^q$ are observed. For a unit vector $u \in S^{q-1}$ let $\hat{K}_k(u)$ denote the k-th sample cumulant of $u'x_1, \ldots, u'x_n$, which are components of x_i's with respect to the direction u. $\hat{K}_k(u)$ is a symmetric k-linear combination of elements of multiway layout consisting of sample multivariate cumulants of $x_1, \ldots, x_n$. Malkovich and Afifi [20] defined measures of departure from multivariate normality (e.g., multivariate skewness or kurtosis) by the maxima of standardized cumulants as

$$\hat{B}_k = \max_{u \in S^{q-1}} |\hat{Z}_k(u)|$$

with $\hat{Z}_k(u) = \sqrt{n}\, \hat{K}_k(u)/\hat{K}_2(u)^{k/2}$, and proposed tests for multivariate normality based on the statistics $\hat{B}_k$.

Under the null hypothesis of multivariate normality, $\hat{B}_k$ has a relatively simple limiting null distribution as follows (Theorem 2.1 of Kuriki and Takemura [18]).

Theorem 4.1 *Let z be a q^k-dimensional random vector consisting of independent standard normal random variables. Let*

$$B_k = \max_{h \in S^{q-1}} |(\underbrace{h \otimes \cdots \otimes h}_{k})' z|.$$

Then $\hat{B}_k$ converges in distribution to B_k as $n \to \infty$.

B_k has a canonical form of T in (2.1) with

$$M = \{\epsilon \underbrace{h \otimes \cdots \otimes h}_{k} \mid h \in S^{q-1},\ \epsilon \in \{1, -1\}\}, \qquad \dim M = q - 1.$$

The geometric invariants and critical radius of the index set M are given in Theorem 3.3 of Kuriki and Takemura [18].

Theorem 4.2 *The nonzero geometric invariants of M are given by*

$$w_{q-2m} = k^{\frac{1}{2}(q-1)} \Big(-\frac{k-1}{k}\Big)^m \frac{\Gamma(\frac{1}{2}(q+1))}{\Gamma(\frac{1}{2}(q+1)-m)\, m!},$$

$m = 0, 1, \ldots, [(q-1)/2]$. *The critical radius is given by $\theta_c = \cos^{-1} \sqrt{(2k-2)/(3k-2)}$.*

For example consider the case $q = 2$. Then the approximate tail probability is given by

$$P(B_k \geq x) \sim \sqrt{k}\, \bar{G}_2(x^2) = \sqrt{k}\, e^{-x^2/2}.$$

The upper tail probabilities of B_3 by Monte Carlo simulation and its approximation by the tube method are depicted in Figure 4. In addition, the speed of convergence of $\hat{B}_3$ to B_3 as n goes to infinity is examined by Monte Carlo simulation. Here the Monte Carlo simulations are based on 100,000 replications. One can see that the tube formula is accurate enough for approximating tail probabilities.

`--- Figure 4 around here ---`

As we have seen, the test statistic by Malkovich and Afifi [20] is constructed by searching a direction of nonnormality. A test by Anderson and Stephens [2] for testing uniformity of directional data is based on a similar idea. They constructed a test statistic by searching a dense or sparse direction among observations. The tube formula is also available for deriving the limiting null distribution of the Anderson-Stephens statistic (Kuriki and Takemura [17]).

5 Order restricted correspondence analysis.

Let (n_{ij}) be an $a \times b$ contingency table with ordinal row and column categories. As statistical models for such ordered categorical data, the following models of cell probabilities have been proposed (Nishisato and Arri [25], Goodman [6]):

$$p_{ij} = p_{i\cdot}p_{\cdot j}(1 + \phi\mu_i\nu_j) \qquad \text{(correspondence analysis)} \tag{5.1}$$

or

$$p_{ij} = \exp\{\alpha_i + \beta_j + \phi\mu_i\nu_j\} \qquad \text{(RC association model)}, \tag{5.2}$$

where μ_i's and ν_j's are order restricted scores

$$\mu_1 \le \cdots \le \mu_a, \qquad \nu_1 \le \cdots \le \nu_b$$

with side conditions

$$\sum_i p_{i\cdot}\mu_i = \sum_j p_{\cdot j}\nu_j = 0, \quad \sum_i p_{i\cdot}\mu_i^2 = \sum_j p_{\cdot j}\nu_j^2 = 1.$$

A natural estimator of ϕ in the order restricted correspondence analysis is given by

$$\hat{\phi} = \max\{\textstyle\sum_{ij}\hat{p}_{ij}\mu_i\nu_j \;\mid\; \sum_i \hat{p}_{i\cdot}\mu_i = \sum_j \hat{p}_{\cdot j}\nu_j = 0,\ \sum_i \hat{p}_{i\cdot}\mu_i^2 = \sum_j \hat{p}_{\cdot j}\nu_j^2 = 1,$$
$$\mu_1 \le \cdots \le \mu_a,\ \nu_1 \le \cdots \le \nu_b\},$$

where $\hat{p}_{ij} = n_{ij}/n$. Hirotsu [8] suggested a test for independence

$$H_0 \ : \ p_{ij} = p_{i\cdot}p_{\cdot j} \quad \text{(or equivalently } \phi = 0) \tag{5.3}$$

based on the statistic $\hat{\phi}$.

Let $Z = (z_{ij}) \in R^{a\times b}$ be an $a \times b$ random matrix consisting of independent standard random variables $N(0,1)$. Let

$$\begin{aligned} P &= \{(v_1,\ldots,v_a)' \in S^{a-1} \mid \textstyle\sum_i \sqrt{p_{i\cdot}}v_i = 0,\ v_1/\sqrt{p_{1\cdot}} \le \cdots \le v_a/\sqrt{p_{a\cdot}}\} \subset S^{a-1}, \\ Q &= \{(w_1,\ldots,w_b)' \in S^{b-1} \mid \textstyle\sum_j \sqrt{p_{\cdot j}}w_j = 0,\ w_1/\sqrt{p_{\cdot 1}} \le \cdots \le w_b/\sqrt{p_{\cdot b}}\} \subset S^{b-1} \end{aligned}$$

be convex spherical polyhedra with $\dim P = a-2$, $\dim Q = b-2$. Let

$$T = \max_{v\in P,\, w\in Q} v'Zw = \max_{v\in P,\, w\in Q} \operatorname{tr}((vw')'Z). \tag{5.4}$$

Then by the continuous mapping theorem, we have the following result (Theorem 2.1 of Kuriki [14]).

Theorem 5.1 *Under the independence model $\phi = 0$, $\sqrt{n}\,\hat{\phi}$ converges in distribution to T in (5.4).*

Consider the likelihood ratio test for testing H_0 in (5.3). The limiting null distribution of the likelihood ratio criterion does not converge to the χ^2 distribution, because the null parameter vector is located at the boundary of the whole parameter space. Applying the general theory for this type of nonregular case by Chernoff [5] (see also Self and Liang [26]), we obtain the following result (Theorem 2.2 of Kuriki [14]).

Theorem 5.2 *The likelihood ratio test criterion for testing the independence model against the order restricted models (5.1) or (5.2) converges in distribution to T^2 with T given in (5.4).*

Since T in (5.4) is of the form (2.1) with

$$M = P \otimes Q = \{vw' \in R^{a\times b} \mid v \in P,\ w \in Q\}, \qquad \dim M = a + b - 4,$$

upper tail probabilities of the limiting null distribution T can be obtained by the tube method. The coefficients w_i's in the tube formula are given in Theorem 2.6 of Kuriki [14] as follows.

Theorem 5.3 *Let $w_e(P)$, $1 \le e \le a-1$, be geometric invariants of P such that the volume of the tube P_θ around P in S^{a-1} is expressed as*

$$\mathrm{Vol}(P_\theta) = \Omega_a \sum_{e=1}^{a-1} w_e(P) \bar{B}_{\frac{1}{2}e,\frac{1}{2}(a-e)}(\cos^2\theta). \tag{5.5}$$

Let $w_f(Q)$, $1 \le f \le b-1$, be defined similarly for Q. Then the geometric invariants w_i's of M are given by

$$w_i = \sum_{e+f-1-2k=i} w_e(P)\, w_f(Q)\, c_{e,f,k}, \qquad i = 1, 2, \ldots, a-b-3,$$

where

$$c_{e,f,k} = (-1)^k\, 2^{e+f-1-k}\, \frac{\Gamma(\frac{1}{2}(e+1))\,\Gamma(\frac{1}{2}(f+1))\,\Gamma(\frac{1}{2}(e+f-1)-k)}{\sqrt{\pi}\,\Gamma(e-k)\,\Gamma(f-k)\,k!},$$

$0 \le k \le \min(e,f) - 1$. *The critical radius of M is $\theta_c = \pi/4$.*

The coefficients $w_e(P)$ and $w_f(Q)$ are known as the level probabilities of Bartholomew's test (Barlow, *et al.* [3]). Consider a one-way ANOVA model $x_i \sim N(\theta_i, \sigma^2/n_i)$, $i = 1, \ldots, k$. Denote by $\hat{\theta}_i$ the maximum likelihood estimator of θ_i under the simple order restriction $\theta_1 \le \cdots \le \theta_k$. The level probability $P(l, k; n_1, \ldots, n_k)$, $1 \le l \le k$, is defined to be the probability under $\theta_1 = \cdots = \theta_k$ that the estimators $\hat{\theta}_i$, $i = 1, \ldots, k$, take exactly l distinct values. Then $w_e(P)$ in (5.5) is equal to

$$w_e(P) = P(e+1, a; p_{1\cdot}, \ldots, p_{a\cdot}), \qquad e = 1, 2, \ldots, a-1.$$

6 A change-point problem in a two-way layout.

Consider a two-way layout where both of row and column factors are ordinal variables. Observations x_{ij}, $i = 1, \ldots, I$, $j = 1, \ldots, J$, are assumed to be independent random variables $N(\mu_{ij}, 1)$. Hirotsu [9] proposed the following model with a change-point at (τ_1, τ_2),

$$\mu_{ij} = \begin{cases} \theta + \alpha_i + \beta_j + \gamma & (i \le \tau_1 \ \text{ and } \ j \le \tau_2), \\ \theta + \alpha_i + \beta_j & (\text{otherwise}), \end{cases}$$

where θ, α_i, β_j, γ and (τ_1, τ_2) are unknown parameters. Then the critical region of the likelihood ratio test for testing $H_0 : \gamma = 0$ against $H_1 : \gamma > 0$ is written by

$$\max_{1 \le k \le I-1,\ 1 \le l \le J-1} \omega'_{kl} x > c, \tag{6.1}$$

where $x = (x_{11}, x_{12}, \ldots, x_{IJ})$ is a lexicographically rearranged vector, and

$$\omega_{kl} = \sqrt{\frac{kl(I-k)(J-l)}{IJ}} \Bigl(\underbrace{-\frac{1}{k}, \ldots, -\frac{1}{k}}_{k}, \underbrace{\frac{1}{I-k}, \ldots, \frac{1}{I-k}}_{I-k}\Bigr)' \otimes \Bigl(\underbrace{-\frac{1}{l}, \ldots, -\frac{1}{l}}_{l}, \underbrace{\frac{1}{J-l}, \ldots, \frac{1}{J-l}}_{J-l}\Bigr)'.$$

Note that $\|\omega_{kl}\| = 1$ and each $\omega'_{kl} x$ has the distribution $N(0, 1)$ under the null hypothesis.

The test statistic in (6.1) is considered as the maximum of a Gaussian random field with a discrete index set

$$M = \{\omega_{kl} \mid 1 \le k \le I-1,\ 1 \le l \le J-1\} \subset S^{I \times J - 1}.$$

If a volume formula for M_θ is available, we can calculate exact significance levels in (6.1). Noting that $\omega'_{kl} x$ and $\omega'_{mn} x$ are highly correlated when $k \doteq m$ and $l \doteq n$, Ninomiya [23], [24] considers a 2-dimensional piecewise linear submanifold of $S^{I \times J - 1}$ defined as

$$\begin{aligned} \tilde{M} \ &= \ \left\{ \frac{s\omega_{k+1,l} + t\omega_{k,l+1} + u\omega_{k,l}}{\|s\omega_{k+1,l} + t\omega_{k,l+1} + u\omega_{k,l}\|} \mid {}_{1 \le l \le J-2}^{1 \le k \le I-2}, {}_{s,t,u \ge 0}^{s+t+u=1} \right\} \\ &\quad \cup \left\{ \frac{s\omega_{k+1,l} + t\omega_{k,l+1} + u\omega_{k+1,l+1}}{\|s\omega_{k+1,l} + t\omega_{k,l+1} + u\omega_{k+1,l+1}\|} \mid {}_{1 \le l \le J-2}^{1 \le k \le I-2}, {}_{s,t,u \ge 0}^{s+t+u=1} \right\}. \end{aligned}$$

Since $\tilde{M} \supset M$, it holds that $\mathrm{Vol}(\tilde{M}_\theta) \ge \mathrm{Vol}(M_\theta)$. Therefore evaluation of $\mathrm{Vol}(\tilde{M}_\theta)$ yields a conservative testing procedure.

$\tilde{M}$ consists of linear submanifolds pasted together and it contains non-smooth edges and vertices. In particular the critical radius of $\tilde{M}$ is zero. Therefore justification of the tube formula for $\mathrm{Vol}(\tilde{M}_\theta)$ becomes complicated. For 1-dimensional (non-smooth) M, Naiman's inequality (Naiman [21]) gives a simple upper bound for $\mathrm{Vol}(\tilde{M}_\theta)$. However generalization of Naiman's inequality to 2-dimensional non-smooth manifolds is very difficult. Ninomiya [23], [24] succeeded in deriving an upper bound for $\mathrm{Vol}(\tilde{M}_\theta)$ and proposed several conservative rejection regions for testing H_0.

Appendix.

A.1 Geometric invariants.

Let M be a d-dimensional C^2-submanifold of S^{n-1} with piecewise smooth boundaries. Assume that the critical radius θ_c of M is positive. Let

$$K = K(M) = \bigcup_{c\geq 0} cM \subset R^n$$

be the smallest cone containing M. For each $u \in K$ define the normal cone at u by the dual cone

$$N_u(K) = S_u(K)^* = \{v \in R^n \mid v'\tilde{u} \leq 0,\ \forall \tilde{u} \in S_u(K)\},$$

where $S_u(K)$ is the support cone (tangent cone) of K at u. Let $H(u,v)$ be the second fundamental form of M at u with respect to the direction $v \in N_u(K)$ (see, e.g., Section 2.3 of Takemura and Kuriki [28]).

Let $\partial M_{\tilde{d}}$ be the $\tilde{d}$-dimensional boundary of M. Let $u \in \partial M_{\tilde{d}}$ and let $v \in N_u(K) \cap S^{n-1}$. The l-th symmetric function of the principal curvatures of M, i.e., the eigenvalues of the second fundamental form $H(u,v)$, is denoted by $\mathrm{tr}_l H(u,v)$. The geometric invariants (curvature invariants) of M are given as follows (Proposition 2.1 of Takemura and Kuriki [30]).

Theorem A.1 *For $e = 0, \ldots, d$,*

$$w_{d+1-e} = \frac{1}{\Omega_{d+1-e}\Omega_{n-d-1+e}} \sum_{\tilde{d}=d-e}^{d} \int_{\partial M_{\tilde{d}}} \mathrm{d}u \int_{N_u(K)\cap S^{n-1}} \mathrm{d}v\, \mathrm{tr}_{\tilde{d}-d+e} H(u,v), \tag{A.1}$$

where for each $0 \leq \tilde{d} \leq d$, $\mathrm{d}u$ and $\mathrm{d}v$ are the volume elements of $\partial M_{\tilde{d}}$ and $N_u(K) \cap S^{n-1}$, respectively.

In (A.1) the geometric invariants w_i's are expressed in terms of the second fundamental form $H(u,v)$, which depends on the way of embedding $M \subset S^{n-1}$. When M is a closed manifold without boundary, Weyl [32] showed that w_i's can be expressed in terms of the curvature tensor, which is intrinsic, independent of the embedding. Weyl's formula can be rewritten in a more sophisticated manner using the double forms of the curvature. See Gray [7] for the case of the tube formula in the Euclidean space.

A.2 Proof of Theorem 3.1.

First we define a combinatorial quantity. Let d_1, d_2, d_3 be positive integers. Let

$$A_1 = \{1, 2, \ldots, d_1\},\ \ A_2 = \{d_1+1, d_1+2, \ldots, d_1+d_2\},\ \ A_3 = \{d_1+d_2+1, d_1+d_2+2, \ldots, d\}.$$

Then A_1, A_2, A_3 form a partition of $A = \{1, 2, \ldots, d\}$. The cardinality of A_i is $d_i = |A_i|$. Let a map $\tau : \{1, \ldots, d\} \to \{1, 2, 3\}$ be defined by $\tau(a) = i$ for $a \in A_i$.

Consider a set of m pairings

$$\{(a_1,a_2),\dots,(a_{2m-1},a_{2m}) \mid a_1<a_3<\cdots<a_{2m-1},\ a_1<a_2,\dots,a_{2m-1}<a_{2m}\} \qquad \text{(A.2)}$$

such that

(i) $2m$ indices $a_1,a_2,\dots,a_{2m}$ are distinct elements of A.

(ii) For each pairing in (A.2), say (a_{2l-1},a_{2l}), $\tau(a_{2l-1})\neq\tau(a_{2l})$, $l=1,\dots,m$.

Furthermore let $n(d_1,d_2,d_3;m)$ denote the total number of sets of m pairings satisfying (i) and (ii). Then the nonzero geometric invariants w_i's are expressed as

$$w_{d+1-2m}=\frac{(-1/2)^m\,\pi\,\Gamma(\frac{1}{2}(d+1)-m)}{\prod_{i=1}^3\Gamma(\frac{1}{2}(d_i+1))}\,n(d_1,d_2,d_3;m).$$

(Theorem 3.2 of Kuriki and Takemura [18]). So we have to evaluate $n(d_1,d_2,d_3)$.

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 such that $|B_{12}|=l_3,\ |B_{13}|=l_2,\ B_{12}\cap B_{13}=\emptyset$. Similarly choose $B_{21}\subset A_2$ and $B_{23}\subset A_2$ such that $|B_{21}|=l_3,\ |B_{23}|=l_1,\ B_{21}\cap B_{23}=\emptyset$; choose $B_{31}\subset A_3$ and $B_{32}\subset A_3$ such that $|B_{31}|=l_2,\ |B_{32}|=l_1,\ B_{31}\cap B_{32}=\emptyset$. 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 (A.2) satisfying (i) and (ii). Taking summation for feasible triplets $(l_1,\ l_2,\ l_3)$ proves the theorem.

The fact that $\cos\theta_c=2/\sqrt{7}$ is proved in Theorem 3.2 of Kuriki and Takemura [18] as well.

References

[1] Adler, R. J. On excursion sets, tube formulas and maxima of random fields. *Ann. Appl. Probab.*, 2000; **10**: 1–74.

[2] Anderson, T. W. and Stephens, M. A. Tests for randomness of directions against equatorial and bimodal alternatives. *Biometrika*, 1972; **59**: 613–621.

[3] Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. *Statistical Inference Under Order Restrictions*, Wiley: London, 1972.

[4] Boik, R. J. and Marasinghe, M. G. Analysis of nonadditive multiway classifications. *J. Amer. Statist. Assoc.*, 1989; **84**: 1059–1064.

[5] Chernoff, H. On the distribution of the likelihood ratio. *Ann. Math. Statist.*, 1954; **25**: 573–578.

[6] Goodman, L. A. The analysis of cross-classified data having ordered and/or unordered categories: association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *Ann. Statist.*, 1985; **13**: 10–69.

[7] Gray, A. *Tubes*. Addison-Wesley: Redwood City, 1990.

[8] Hirotsu, C. Use of cumulative efficient scores for testing ordered alternatives in discrete models. *Biometrika*, 1982; **69**: 567–577.

[9] Hirotsu, C. Two-way change-point model and its application. *Austral. J. Statist.*, 1997; **39**: 205–218.

[10] Hotelling, H. Tubes and spheres in n-spaces, and a class of statistical problems. *Amer. J. Math.*, 1939; **61**: 440–460.

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

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

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

[14] Kuriki, S. Asymptotic distribution of inequality restricted canonical correlation with application to tests for independence in ordered contingency tables. Research Memorandum No. 803, The Institute of Statistical Mathematics, 2001, submitted for publication.

[15] Kuriki, S. and Takemura, A. Distribution of the maximum of Gaussian random field: tube method and Euler characteristic method. *Proc. Inst. Statist. Math.*, 1999; **47**: 199–219 (in Japanese).

[16] Kuriki, S. and Takemura, A. Some geometry of the cone of nonnegative definite matrices and weights of associated $\bar{\chi}^2$ distribution. *Ann. Inst. Statist. Math.*, 2000; **52**: 1–14.

[17] Kuriki, S. and Takemura, A. Tail probabilities of the limiting null distributions of the Anderson-Stephens statistics. Discussion Paper CIRJE-F-77, Faculty of Economics, Univ. of Tokyo, 2000, submitted for publication.

[18] Kuriki, S. and Takemura, A. Tail probabilities of the maxima of multilinear forms and their applications. *Ann. Statist.*, 2001; **29**, to appear.

[19] Leurgans, S. and Ross, R. T. Multilinear models: Applications in spectroscopy (with discussion). *Statistical Science*, 1992; **7**: 289–319.

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

[21] Naiman, D. Q. Conservative confidence bands in curvilinear regression. *Ann. Statist.*, 1986; **14**: 896–906.

[22] Naiman, D. Q. Volumes of tubular neighborhoods of spherical polyhedra and statistical inference. *Ann. Statist.*, 1990; **18**: 685–716.

[23] Ninomiya, Y. Construction of conservative testing for detection of two-way change-point. *Japanese Journal of Applied Statistics*, 2000; **29**: 117–130 (in Japanese).

[24] Ninomiya, Y. Construction of conservative testing for change-point problems in two-dimensional random fields. Research Memorandum No. 746, The Institute of Statistical Mathematics, 2000, submitted for publication.

[25] Nishisato, S. and Arri, P. S. Nonlinear programming approach to optimal scaling of partially ordered categories. *Psychometrika*, 1975; **40**: 525–548.

[26] Self, S. G. and Liang, K-Y. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Amer. Statist. Assoc.*, 1987; **82**: 605–610.

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

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

[29] Takemura, A. and Kuriki, S. Tail probability via tube formula when critical radius is zero, 2000, submitted for publication.

[30] Takemura, A. and Kuriki, S. On the equivalence of the tube and Euler characteristic methods for the distribution of the maximum of Gaussian fields over piecewise smooth domains. *Ann. Appl. Probab.*, 2001, conditionally accepted.

[31] Takemura, A. and Kuriki, S. Maximum covariance difference test for equality of two covariance matrices. In *Algebraic Methods in Statistics and Probability*, Viana, M. and Richards, D. (eds). Contemporary Mathematics Vol. 287, American Mathematical Society, 2001; 283–302.

[32] Weyl, H. On the volume of tubes. *Amer. J. Math.*, 1939; **61**: 461–472.

[33] Worsley, K. J. Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *Ann. Statist.*, 1995; **23**: 640–669.

Address:

Satoshi Kuriki
Institute of Statistical Mathematics
4-6-7 Minami-azabu, Minato-ku
Tokyo 106-8569, Japan
Email: `kuriki@ism.ac.jp`

Akimichi Takemura
Graduate School of Information Science and Technology
University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Email: `takemura@stat.t.u-tokyo.ac.jp`

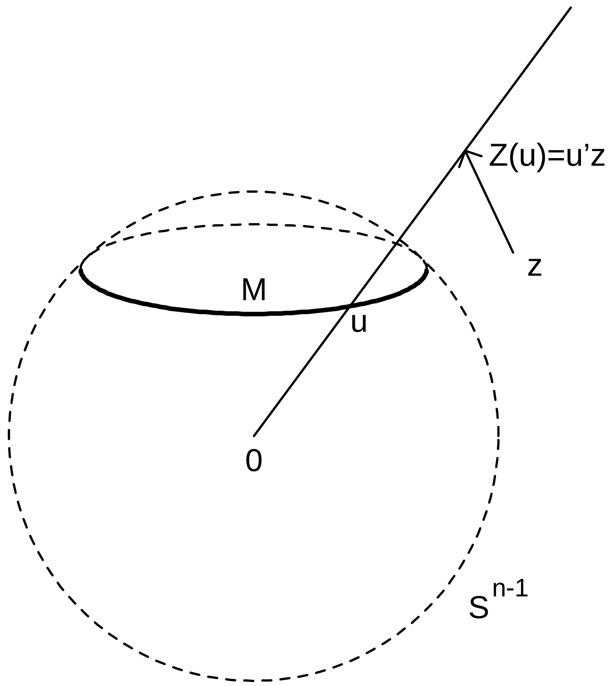

Figure 1. Gaussian random field $Z(u)$, $u \in M$.

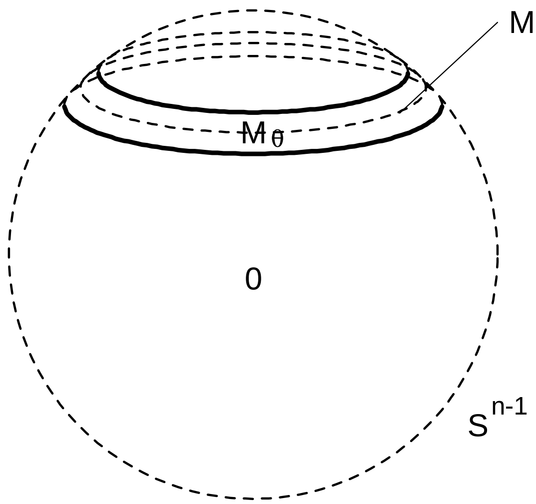

Figure 2. Tube M_θ around M.

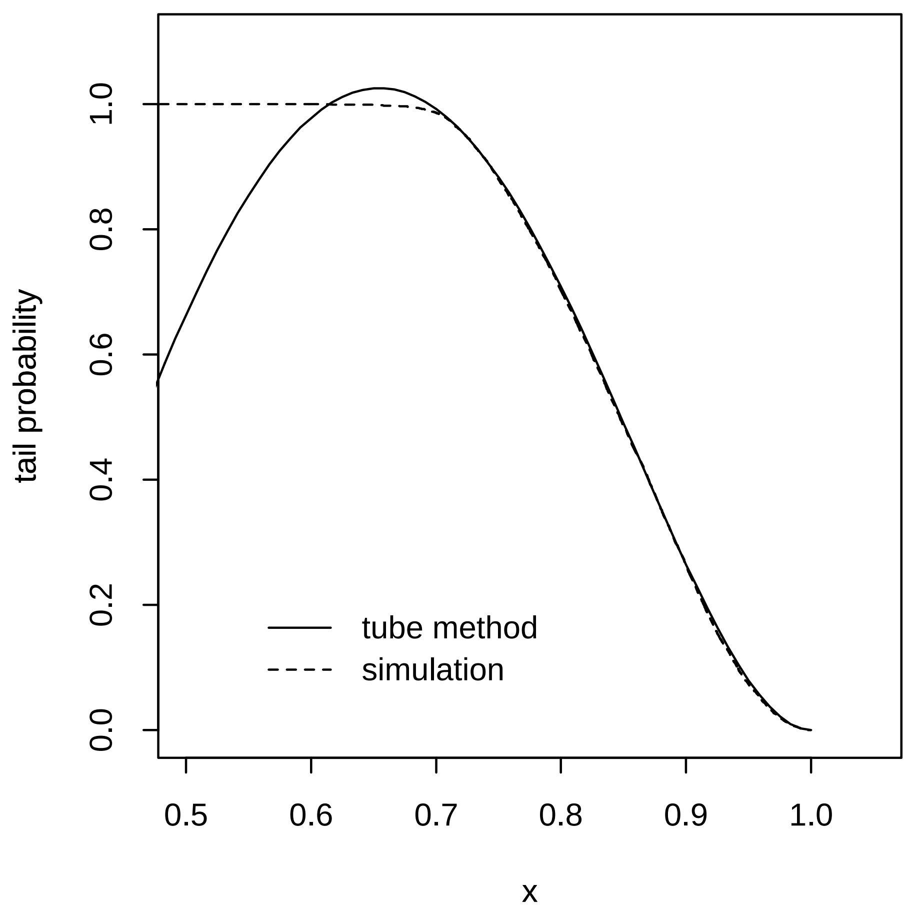

Figure 3. The upper tail probability $P(U \geq x)$.

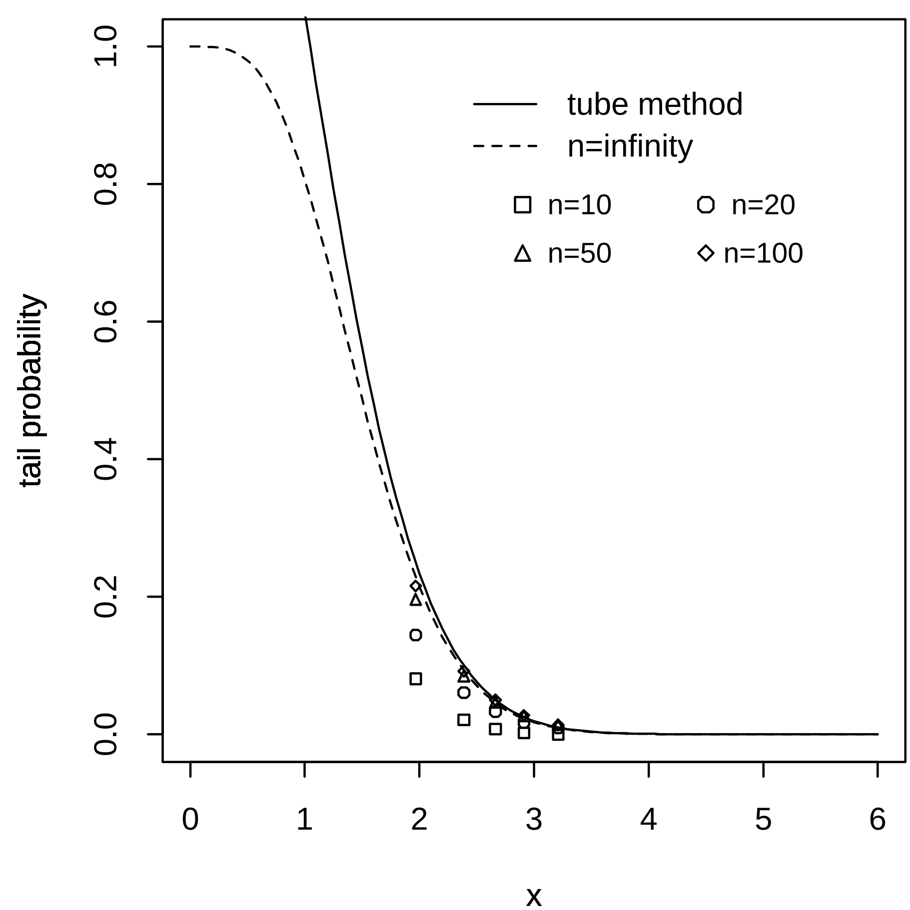

Figure 4. The upper tail probability $P(B_3 \geq x)$.