

チューブ法の理論・応用とその周辺

栗木 哲†

(受付 2018 年 12 月 28 日; 改訂 2019 年 3 月 20 日; 採択 3 月 26 日)

要 旨

チューブ法は, 正規確率場の最大値の上側裾確率を精度良く近似する積分幾何学的手法である. 同じ目的のための手法として, オイラー標数法がある. 本稿では, チューブ法とオイラー標数法の考え方を概観し, 統計学への応用, ならびに最近の発展について解説する.

キーワード: オイラー標数法, 同時信頼区間, どこでも効果, 射影追跡, 特異モデル, VBM データ解析.

1. はじめに

表題のチューブ法は, 正規確率場の最大値の上側裾確率を精度良く近似する積分幾何学的手法である. 同じ目的のための手法として, オイラー標数法がある. 本稿では, チューブ法とオイラー標数法の考え方を概観し, 統計学への応用, ならびに最近の発展について解説する.

チューブ法では, 集合の管状近傍(チューブ)の体積評価を通して確率分布を評価する. 統計学において, チューブの体積評価の概念が初めて現れたのは, Hotelling (1939) の非線形回帰モデルである:

$$y_i = \beta f(x_i, \theta) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2), \quad i = 1, \dots, n.$$

ここで x_i, y_i はそれぞれ説明変数, 目的変数, $\beta \in \mathbb{R}$ と $\theta \in \Theta \subset \mathbb{R}$ は未知パラメータ, f は θ を含む基底関数である. ここで帰無仮説 $H: \beta = 0$ の尤度比検定を考えると, $y = (y_1, \dots, y_n)^\top$, $f(\theta) = (f(x_1, \theta), \dots, f(x_n, \theta))^\top$ とおくと

$$(1.1) \quad \text{LRT} = \sup_{\theta \in \Theta} \left| \left\langle \frac{y}{\|y\|}, \frac{f(\theta)}{\|f(\theta)\|} \right\rangle \right|$$

が大きな値をとるとき棄却する検定となる. 帰無仮説 $H: \beta = 0$ のもとでは, $y \sim N(0, \sigma^2 I)$, $y/\|y\| \sim \text{Unif}(\mathbb{S}^{n-1})$ (n 次元ユークリッド空間の単位球面 \mathbb{S}^{n-1} 上の一様分布)となる. Hotelling (1939) は, 尤度比検定統計量 LRT の帰無分布を求めることと, \mathbb{S}^{n-1} において, 基準化回帰基底ベクトルの θ を動かした時の軌跡 $M = \{\pm f(\theta)/\|f(\theta)\| \in \mathbb{S}^{n-1} \mid \theta \in \Theta\}$ のまわりのチューブ(図 1)

$$\text{Tube}(M, \theta) = \{u \in \mathbb{S}^{n-1} \mid \text{dist}(u, M) \leq \theta\}, \quad \text{dist}(u, M) = \min_{v \in M} \cos \langle v, u \rangle \quad (\text{大円距離})$$

の体積を求めることは同値であることを, 以下の関係式のように指摘した:

$$\Pr(\text{LRT} \geq a) = \frac{\text{Vol}_{n-1}(\text{Tube}(M, \cos^{-1} a))}{\text{Vol}_{n-1}(\mathbb{S}^{n-1})}.$$

ここで Vol_{n-1} は $n-1$ 次元体積である.

† 統計数理研究所: 〒190-8562 東京都立川市緑町 10-3

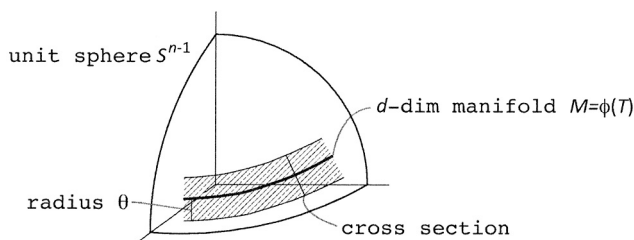


図 1. 球面チューブ.

Hotelling (1939) はさらに論文の中で M が 1 次元(曲線)の場合のチューブ体積公式を微分幾何の言葉で書き下している. この結果はプリンストン高等研究所の同僚であった H. Weyl によって M が一般次元閉多様体の場合に一般化され (Weyl, 1939), 現在では Hotelling-Weyl の定理とよばれている (丹野, 1976; 小林, 1997).

ところで, Hotelling の非線形回帰問題の尤度比検定統計量(1.1)は, 添字集合 θ の上で定義された確率場 $X(\theta) = \langle y/\|y\|, f(\theta)/\|f(\theta)\| \rangle$ の最大値 $\sup_{\theta \in \Theta} X(\theta)$ である. その視点より 1980 年代末から 90 年代にかけて, スタンフォード大で, チューブの体積と確率場の最大値分布に関するいくつかの研究が始まった (Knowles and Siegmund, 1989; Johansen and Johnstone, 1990; Sun, 1991, 1993). それがチューブ法 (tube method; volume-of-tube method) である.

なお近年 selective inference の名で, モデル選択後の条件付推論の方法論が, これもスタンフォード大の研究者を中心に発展している. そこではチューブ法におけるチューブ座標に沿って分解されたピボット統計量が利用される (Taylor et al., 2016 の Kac-Rice formula).

2. チューブ法の概要

$X(t), t \in T \subset \mathbb{R}^d$, を平均 0, 分散 1 のガウス確率場で, そのサンプルパスが滑らかであるものとする. チューブ法は, その最大値の上側裾確率

$$\Pr \left(\sup_{t \in T} X(t) \geq a \right) \quad (a \text{ が大きいとき})$$

を近似する方法である.

一般に平均 0 のガウス確率場 $X(t)$ は, 正則条件の下で以下のような Karhunen-Loève 展開 (KL 展開) を持つ.

$$X(t) = \sum_{i=1}^{\infty} \phi_i(t) \xi_i, \quad \xi_i \sim N(0, 1) \text{ i.i.d.}$$

ここで $\text{Var}(X(t)) = \sum \phi_i^2(t) = 1$ である. この表現は一意ではないが, 標準的な表現として, $\phi_i(\cdot)$ を正值カーネル $R(s, t) = \text{Cov}(X(s), X(t))$ が定義する再生核ヒルベルト空間の正規直交基底ととることができる. $R(s, t)$ の $s = t$ における微分可能性が高次であるほど, サンプルパスは滑らかとなる. チューブ法は, KL 展開が有限項 (n 項) で打ち切られている場合を対象とする. $\xi = (\xi_1, \dots, \xi_n)^\top \sim N_n(0, I_n)$, $\phi(t) = (\phi_1(t), \dots, \phi_n(t))^\top$, $M = \{\phi(t) \mid t \in T\}$ とおけば

$$X(t) = \langle \phi(t), \xi \rangle_{\mathbb{R}^n}, \quad \Pr \left(\sup_{t \in T} X(t) \geq a \right) = \Pr \left(\sup_{p \in M} \langle p, \xi \rangle_{\mathbb{R}^n} \geq a \right)$$

である. $\|\phi_i(t)\| = 1$ なので, $M \subset \mathbb{S}^{n-1}$ である. 以降確率場 $\{\langle p, \xi \rangle_{\mathbb{R}^n}\}_{p \in M}$ を扱う. このガウス確率場のモーメント構造は, $E[\langle p, \xi \rangle] = 0$, $\text{Var}[\langle p, \xi \rangle] = 1$, $\text{Cov}(\langle p, \xi \rangle, \langle q, \xi \rangle) = \langle p, q \rangle$ である.

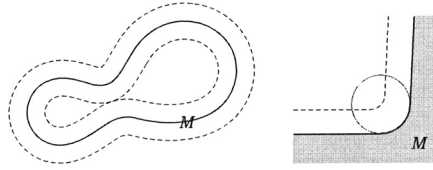


図 2. チューブの自己交差(左：大域的自己交差, 右：局所的自己交差).

M が \mathbb{S}^{n-1} の d 次元部分多様体であるとき, M が境界を持つ(閉多様体でない)場合を含めて, M のまわりのチューブの体積 $\text{Vol}_{n-1}(\text{Tube}(M, \theta))$ は, $0 \leq \theta \leq \exists \theta_c$ の範囲でベータ分布の上側確率の線形和で表すことができる (Takemura and Kuriki, 2002):

$$(2.1) \quad \frac{\text{Vol}_{n-1}(\text{Tube}(M, \theta))}{\text{Vol}_{n-1}(\mathbb{S}^{n-1})} = w_{d+1} \bar{B}_{\frac{d+1}{2}, \frac{n-d-1}{2}}(\cos^2 \theta) + w_d \bar{B}_{\frac{d}{2}, \frac{n-d}{2}}(\cos^2 \theta) + \cdots + w_1 \bar{B}_{\frac{1}{2}, \frac{n-1}{2}}(\cos^2 \theta).$$

ただし w_i は M のみに依存する定数で, 特に

$$w_{d+1} = \frac{\text{Vol}_d(M)}{\text{Vol}_d(\mathbb{S}^d)}, \quad w_d = \frac{\text{Vol}_{d-1}(\partial M)}{2\text{Vol}_d(\mathbb{S}^{d-1})},$$

$\bar{B}_{a,b}$ はパラメータ (a, b) のベータ分布の上側確率である. 定数 θ_c は M のみに依存する正定数で, 本節の最後で定義する.

ガウス確率ベクトル $\xi \sim N(0, I_n)$ の「長さ」 $\|\xi\|$ と「方位」 $\xi/\|\xi\|$ は独立に, それぞれ χ_n 分布(自由度 n のカイ 2 乗分布の平方根)と球面上一様分布 $\text{Unif}(\mathbb{S}^{n-1})$ に従うことから,

$$(2.2) \quad \Pr\left(\sup_{p \in M} \langle p, \xi \rangle_{\mathbb{R}^n} \geq a\right) = E\left[\Pr\left(\sup_{p \in M} \left\langle p, \frac{\xi}{\|\xi\|} \right\rangle_{\mathbb{R}^n} \geq \frac{a}{\|\xi\|} \mid \|\xi\|\right)\right] \\ = E\left[\frac{\text{Vol}_{n-1}\left(M, \cos^{-1}\left(\frac{a}{\|\xi\|}\right)\right)}{\text{Vol}_{n-1}(\mathbb{S}^{n-1})}\right] \\ \approx w_{d+1} \bar{G}_{d+1}(a^2) + w_d \bar{G}_d(a^2) + \cdots + w_1 \bar{G}_1(a^2).$$

ここで \bar{G}_ν は自由度 ν のカイ 2 乗分布の上側確率である. 最後の式は恒等式

$$E\left[\bar{B}_{\frac{\nu}{2}, \frac{n-\nu}{2}}\left(\frac{a}{\|\xi\|}\right)\right] = \bar{G}_\nu(a^2)$$

とチューブ体積公式 (2.1) の組合せで得られるが, (2.1) が θ が小さい ($\theta \leq \theta_c$) ときのみなりたっていたことに対応して, (2.2) は a が大きいときになりたつ近似式である. その誤差は $a \rightarrow \infty$ のとき $O(a^{n-2} e^{-\frac{1}{2}(1+\tan^2 \theta_c) a^2})$ である (Kuriki and Takemura, 2001).

誤差を定めている定数 θ_c はチューブの臨界半径とよばれ, 以下で定義される.

$$\theta_c = \sup\{\theta \mid \text{Tube}(M, \theta) \text{ は自己交差しない}\}.$$

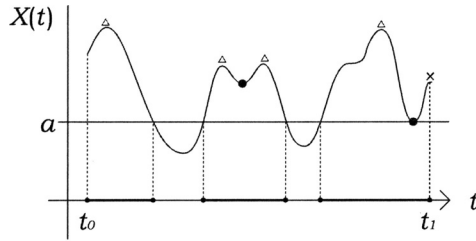
チューブの自己交差には, 大域的な自己交差と局所的な自己交差が考えられる (図 2).

M が与えられたとき, 臨界半径 θ_c を解析的ないしは数値的に評価することができる (栗木・竹村, 2008). 具体的な計算例は Kuriki and Takemura (2001) などにみられる.

3. オイラー標数法の概要

3.1 一次元オイラー標数法

チューブ法と同じ目的の手法に, オイラー標数法がある. ここではその方法を, 添字集合が



$$\chi(T_a) = 4(\Delta) - 2(\bullet) + 1(\times) = 3$$

図 3. エクスカージョン集合とそのオイラー標数.

1次元の場合を中心に概観する.

$X(t)$ を, 添字集合 $T = [t_0, t_1]$ の上で定義された, 滑らかなサンプルパスを持つ実数値確率過程とする. ここで添字の一意性 " $X(s) = X(t)$ a.s. $\Rightarrow s = t$ " を仮定する.

与えられた閾値 a に対して, $X(t)$ が a 以上となるような添字の全体

$$T_a = \{t \in T \mid X(t) \geq a\}$$

をエクスカージョン集合(excursion set)という. エクスカージョン集合 T_a のオイラー標数を $\chi(T_a)$ とおく. T_a は一次元集合なので, $\chi(T_a)$ は T_a の連結成分の個数である. $X(t)$ とエクスカージョン集合の例を図 3 に与える. この例では $\chi(T_a) = 3$ である.

エクスカージョン集合の定義から, $X(t)$ が a 以上となることは T_a が空集合とならないことと同値で, さらにそれは T_a の連結成分の個数が 0 でないことと同値である. すなわち

$$\sup_{t \in T} X(t) \geq a \Leftrightarrow T_a \neq \emptyset \Leftrightarrow \chi(T_a) \geq 1$$

である. ここで a の値が大きいとき, $X(t)$ が 2 つの離れた時点で a を超える事象は稀であることが予想される. すなわち事象 $\chi(T_a) = k$ ($k \geq 2$) が起こる確率は小さいことが期待できる. この近似のもとでは,

$$(3.1) \quad \mathbb{1}\left\{\sup_{t \in T} X(t) \geq a\right\} - \chi(T_a) = -\sum_{k=2}^{\infty} (k-1)\mathbb{1}(\chi(T_a) = k)$$

が無視できる量となり, 両辺の期待値をとることにより

$$(3.2) \quad \Pr\left(\sup_{t \in T} X(t) \geq a\right) \approx E[\chi(T_a)] \quad (a \text{ が大きいとき})$$

がなりたつことが期待される. この近似法を, オイラー標数法(The Euler characteristic method) という (Adler and Hasofer, 1976; Worsley, 1995; Adler and Taylor, 2007). この方法は, 信号処理分野のライス公式 (Rice, 1944) の一般化に相当する. なお (3.1) は常に非正なので, 不等式

$$\Pr\left(\sup_{t \in T} X(t) \geq a\right) \leq E[\chi(T_a)] \quad (\text{すべての } a \text{ について})$$

がなりたつ (Naiman, 1986). この不等式は多重検定や同時信頼区間構成において, 近似が保守側であることを保証する.

次に (3.2) の右辺を具体的に評価する.

定義 1. 区間 $T = [t_0, t_1]$ で定義された滑らかな関数 $X(t)$ の、(拡張された)臨界点 (augmented critical point) とは

$$\dot{X}(t) = 0 \quad (t \in (t_0, t_1)), \quad < 0 \quad (t = t_0), \quad > 0 \quad (t = t_1)$$

をみたす点 $t = t^*$ である。(拡張された)臨界点集合を T^* とおく。

仮定 1. $a > 0$ を所与とする。 t^* を $X(t)$ の臨界点で $X(t^*) \geq a$ であるような点とする。確率 1 で $\dot{X}(t^*) \neq 0$ であること、また t^* は孤立点であることを仮定する。

チューブ法で扱う確率場では、本仮定がなりたつことは容易に確認できる。一般には Worsley (1995) の正則条件 C1-C3 の下でなりたつ。このとき、 $\mathbb{1}_{\{\cdot\}}$ を事象の指示関数とすると、図を描けば分かるように (図 3)

$$(3.3) \quad \chi(T_a) = \sum_{t \in T^* \cap \text{int}(T)} \mathbb{1}_{\{X(t) \geq a\}} \text{sgn}(-\ddot{X}(t)) + \sum_{t \in T^* \cap \partial T} \mathbb{1}_{\{X(t) \geq a\}}$$

がなりたつ。これが一次元版のモースの定理である。

以下では、式(3.3)の第 1 項、第 2 項の期待値を別々にとる。最初に第 1 項を考える。 $t^* \in \text{int}(T)$ とする。まず

$$\lim_{\varepsilon \rightarrow 0} \int_{\text{small region } (\ni t^*)} f(t) \frac{\mathbb{1}_{\{\dot{X}(t) \in (-\varepsilon, \varepsilon)\}}}{2\varepsilon} |\ddot{X}(t)| dt = f(t^*)$$

である。なぜならば、 $\dot{X}(t) = y$ とおき、 $\ddot{X}(t^*) > 0$ の場合は左辺は $\lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} f((\dot{X})^{-1}(y)) dy = f(t^*)$ 、 $\ddot{X}(t^*) < 0$ の場合は左辺は $\lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_{\varepsilon}^{-\varepsilon} f((\dot{X})^{-1}(y)) d(-y) = f(t^*)$ 。これより

$$\begin{aligned} E \left[\sum_{t \in T^* \cap \text{int}(T)} \mathbb{1}_{\{X(t) \geq a\}} \text{sgn}(-\ddot{X}(t)) \right] &= E \left[\lim_{\varepsilon \rightarrow 0} \int_{t_0}^{t_1} \mathbb{1}_{\{X(t) \geq a\}} \text{sgn}(-\ddot{X}(t)) \frac{\mathbb{1}_{\{\dot{X}(t) \in (-\varepsilon, \varepsilon)\}}}{2\varepsilon} |\ddot{X}(t)| dt \right] \\ &= \lim_{\varepsilon \rightarrow 0} \int_{t_0}^{t_1} E \left[\mathbb{1}_{\{X(t) \geq a\}} (-\ddot{X}(t)) \frac{\mathbb{1}_{\{\dot{X}(t) \in (-\varepsilon, \varepsilon)\}}}{2\varepsilon} \right] dt \\ &= \lim_{\varepsilon \rightarrow 0} \int_{t_0}^{t_1} \frac{E[\mathbb{1}_{\{X(t) \geq a\}} (-\ddot{X}(t)) \mathbb{1}_{\{\dot{X}(t) \in (-\varepsilon, \varepsilon)\}}]}{E[\mathbb{1}_{\{\dot{X}(t) \in (-\varepsilon, \varepsilon)\}}]} \\ &\quad \times \frac{E[\mathbb{1}_{\{\dot{X}(t) \in (-\varepsilon, \varepsilon)\}}]}{2\varepsilon} dt \\ &= \int_{t_0}^{t_1} E[\mathbb{1}_{\{X(t) \geq a\}} (-\ddot{X}(t)) | \dot{X}(t) = 0] \theta_{\dot{X}(t)}(0) dt, \end{aligned}$$

ただし

$$\theta_{\dot{X}(t)}(0) = \lim_{\varepsilon \rightarrow 0} \frac{E[\mathbb{1}_{\{\dot{X}(t) \in (-\varepsilon, \varepsilon)\}}]}{2\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{\Pr(\dot{X}(t) \in (-\varepsilon, \varepsilon))}{2\varepsilon}$$

は $\dot{X}(t)$ の密度関数の 0 での値である。(3.3)の第 2 項の期待値評価は容易であり、

$$E[\chi(T_a)] = \int_{t_0}^{t_1} \rho(t) dt + \Pr(X(t_0) \geq a, \dot{X}(t_0) < 0) + \Pr(X(t_1) \geq a, \dot{X}(t_1) > 0),$$

$$\rho(t) = E[\mathbb{1}_{\{X(t) \geq a\}} (-\ddot{X}(t)) | \dot{X}(t) = 0] \theta_{\dot{X}(t)}(0)$$

を得る。これは Worsley (1995) の公式の一次元版である。

3.2 平均 0, 分散 1 のガウス場の場合

以下ではチューブ法と同じ設定 $X(t) = \phi(t)^\top \xi$, $\xi \sim N_n(0, I_n)$ で $\rho(t)$ の具体形をもとめる。 $\phi(t)$ を $n \times 1$ 単位ベクトルで t について C^3 級とする。また $\phi: T \rightarrow \mathbb{S}^{n-1}$ は単射とする。

$X(t)$ の 1 階, 2 階微分は $\dot{X}(t) = \dot{\phi}(t)^\top \xi$, $\ddot{X}(t) = \ddot{\phi}(t)^\top \xi$. これより t を固定して考えると $(X(t), \dot{X}(t), \ddot{X}(t))$ は 3 変量正規分布に従う。正規分布の標準的な計算から $g(t) = \|\dot{\phi}(t)\|^2$ とおくと $\theta_{\dot{X}(t)}(0) = 1/\sqrt{2\pi g(t)}$, $\rho(t) = g(t)^{\frac{1}{2}} e^{-\frac{1}{2}a^2}/(2\pi)$ が得られ,

$$E[\chi(T_a)] = \frac{1}{2\pi} \int_{t_0}^{t_1} g(t)^{\frac{1}{2}} dt \times \Pr(\chi_2^2 \geq a^2) + \frac{1}{2} \Pr(\chi_1^2 \geq a^2)$$

を得る。ここで $\int_{t_0}^{t_1} g(t)^{\frac{1}{2}} dt$ は $M = \{\phi(t) \mid t \in [t_0, t_1]\}$ の一次元体積であるので, (2.2) の $d = 1$ の場合に一致する。

添字集合が多次元 (d 次元) の場合もオイラー標数法は (3.2) で定義される。その場合の期待値 $E[\chi(T_a)]$ の評価は, 拡張された臨界点とモースの定理の一般次元への一般化が必要となる。

定理 1. (Takemura and Kuriki, 2002) $X(t)$, $t \in T \subset \mathbb{R}^d$ を平均 0, 分散 1 の滑らかなサンプルパスを持つガウス確率場とする。エクスカージョン集合 $T_a = \{t \in T \mid X(t) \geq a\}$ のオイラー標数の期待値は

$$(3.4) \quad E[\chi(T_a)] = \frac{\text{Vol}_d(T)}{\text{Vol}(\mathbb{S}^d)} \Pr(\chi_{d+1}^2 \geq a^2) + \frac{\text{Vol}_{d-1}(\partial T)}{2\text{Vol}(\mathbb{S}^{d-1})} \Pr(\chi_d^2 \geq a^2) + \dots$$

ここで \mathbb{S}^d は d 次元単位球面, 添字集合 T は計量

$$g_{ij} = \text{Cov} \left(\frac{\partial X(t)}{\partial t_i}, \frac{\partial X(t)}{\partial t_j} \right) = \left\langle \frac{\partial \phi(t)}{\partial t_i}, \frac{\partial \phi(t)}{\partial t_j} \right\rangle_{\mathbb{R}^n}$$

が付与されたリーマン多様体と考えている。

(3.4) は (2.2) と全く同じ式であり, オイラー標数法はチューブ法と同等であることが分かる。

4. 統計学への応用

4.1 同時信頼区間の構成

本節では, チューブ法の統計学への応用を紹介する。最初の例は Naiman (1986) による回帰モデルの同時信頼区間構成である。

観測値 $\{(x_i, y_i)\}$ に対して回帰モデル

$$y_i = b^\top f(x_i) + \varepsilon_i, \quad x_i \in \mathcal{X} \subset \mathbb{R}, \quad \varepsilon_i \sim N(0, 1) \text{ i.i.d.}$$

を想定する。ここで b は未知パラメータベクトル, $f(x)$ は所与の基底関数ベクトル f である。例えば $f(x) = (1, x, x^2, \dots)^\top$ とおくと多項式回帰モデルになる。簡単のため, 誤差項の分散は既知で 1 とする。

同時信頼区間とは, 真の回帰曲線 $\{(x, b^\top f(x)) \mid x \in \mathcal{X}\}$ を確率 $1 - \alpha$ 以上で含む集合である。その集合の取り方には任意性があるが, ここでは \hat{b} を b の最小二乗推定量 (OLS) とし,

$$\{(x, y) \mid y \in \hat{b}^\top f(x) \pm \text{s.d.}(\hat{b}^\top f(x)) c_\alpha, x \in \mathcal{X}\}$$

の形を考える。ここで $A \pm B$ は区間 $(A - B, A + B)$ であり, c_α は α に応じて定める定数である。情報行列の逆行列を $\Sigma = (\sum_i f(x_i) f(x_i)^\top)^{-1}$ とおくと, c_α は

$$1 - \alpha \leq \Pr(b^\top f(x) \in \widehat{b}^\top f(x) \pm \|\Sigma^{\frac{1}{2}} f(x)\| \cdot c_\alpha, \forall x \in \mathcal{X}) = \Pr\left(\max_{x \in \mathcal{X}} |\phi(x)^\top \xi| \leq c_\alpha\right),$$

ただし $\xi = \Sigma^{-\frac{1}{2}}(\widehat{b} - b)$, $\phi(x) = \Sigma^{\frac{1}{2}} f(x) / \|\Sigma^{\frac{1}{2}} f(x)\|$ をみたくし閾値として選ぶことができる。ここで $\xi \sim N_n(0, I_n)$, $\phi(x) \in \mathbb{S}^{n-1}$ なので、チューブ法が利用可能である。Naiman (1986) の不等式により、チューブ法で構成した同時信頼区間はわずかに保守側の近似となる。

また Lu and Kuriki (2017) は、複数の回帰曲線

$$\{(x, b_{(h)}^\top f(x)) \mid x \in \mathcal{X}\}, \quad h = 1, \dots, k$$

のコントラスト

$$\left\{ \sum_{h=1}^k c_{(h)} b_{(h)}^\top f(x) \mid x \in \mathcal{X}, (c_{(h)}) \in \mathcal{C} \right\}$$

の保守側同時信頼区間を与えている。ここで \mathcal{C} は $\sum_{h=1}^k c_{(h)} = 0$ をみたくし係数ベクトル $(c_{(h)}) = (c_{(1)}, \dots, c_{(k)})$ の全体である。この問題は複数の成長曲線の多重比較として現れる。

なお、回帰曲線の同時信頼区間構成法として、コーシー・シュワルツの不等式

$$\max_{x \in \mathcal{X}} |\phi(x)^\top \xi| \leq \|\phi(x)\| \cdot \|\xi\| = \|\xi\|$$

を用いる方法 (Scheffé 法) がある。本方法はラオ (1986) で紹介されていることからしばしば使われるが、全実軸を定義域とする単回帰モデル $(f(x) = (1, x)^\top, x \in \mathcal{X} = \mathbb{R})$ という特別の場合を除いてコーシー・シュワルツの不等式の上界はタイトにならないため、しばしば非常に保守的となり、使うべきではない。

4.2 Look-Elsewhere Effect (LEE, どこでも効果)

実験物理、特に新素粒子探索において、探索する領域を広くすると見せかけの新発見が起きることが知られている。この現象は、Look-Elsewhere Effect (LEE, どこでも効果) とよばれているが、統計学の言葉では検定の多重性に他ならない (Gross and Vitells, 2010)。

正規性のもとでは標準偏差の 2 倍が 95% に相当するため、通常、実験結果の図示には 2 シグマのエラーバーをつけることが慣習的に行われている。しかしながら素粒子探索においては 5 倍の偏差の検出をもって、新発見と認められる (村山, 2013)。この 5 シグマルールの正当化はオイラー標数法に基づいてなされている (van Dyk, 2014)。

4.3 ランダム行列の最大特異値

成分が独立に標準正規分布に従う $n \times p$ ランダム行列 $\Xi = (\xi_{ij}) \in \mathbb{R}^{n \times p}$, $\xi_{ij} \sim N(0, 1)$ i.i.d. を考える。その最大特異値 $\sigma_{\max}(\Xi)$ は $g \in \mathbb{R}^n, h \in \mathbb{R}^p$ を長さ 1 のベクトルとするときの双線形式 $g^\top \Xi h = \text{tr}((hg^\top)^\top \Xi)$ の最大値である。 hg^\top の値はノルムが 1 でランク 1 の $n \times p$ 行列全体を取りうるので

$$\sigma_{\max}(\Xi) = \max_{H \in M} \text{tr}(H^\top \Xi), \quad M = \{H \in \mathbb{R}^{n \times p} \mid \text{tr}(H^\top H) = 1, \text{rank}(H) = 1\}.$$

M は $\mathbb{R}^{n \times p}$ の単位球面 $\mathbb{S}^{n \times p - 1} = \{H \in \mathbb{R}^{n \times p} \mid \text{tr}(H^\top H) = 1\}$ の部分多様体なので、その分布はチューブ法で近似できる (Kuriki and Takemura, 2001, 2008b) :

$$\Pr(\sigma_{\max}(\Xi)^2 \geq a) \approx \frac{2^{p+n-2} \Gamma(\frac{p+1}{2}) \Gamma(\frac{n+1}{2})}{\sqrt{\pi}} \sum_{j=0}^{p-1} \left(-\frac{1}{2}\right)^j \frac{\Gamma(\frac{n+p-1-2j}{2})}{\Gamma(p-j) \Gamma(n-j) j!} \Pr(\chi_{n+p-1-2j}^2 \geq a).$$

$\sigma_{\max}(\Xi)^2$ は自由度 n の $p \times p$ ウィシャート行列 $\Xi^\top \Xi$ の最大固有値である。

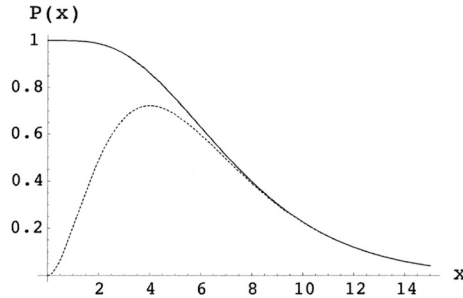


図 4. 2×2 自由度 5 Wishart 分布の最大固有値の上側確率
(実線：真値，点線：チューブ法).

図 4 に， $n = 5$ ， $p = 2$ の場合の σ_{\max}^2 の上側確率とそのチューブ法近似を図示する．なお最近の研究結果として，ランダム行列の手法を用いて，チューブ法近似の n と p が発散するときの極限を論じることができる (栗木, 2019)．

4.4 多次元正規性検定/探索的射影追跡の有意性

q 次元 i.i.d. ベクトル値観測系列 $x_t = (x_t^1, \dots, x_t^q) \in \mathbb{R}^q$ ($t = 1, \dots, n$) が得られているとする．このサンプルが q 次元正規分布に従っているかどうかを検定し，正規分布に従っていない場合にはどのような正規性からの乖離が見られるかを検出したいとする．

サンプルの方向ベクトル $h \in \mathbb{S}^{q-1}$ 方向の成分 $h^\top x_t \in \mathbb{R}$ ($t = 1, \dots, n$) の j 次サンプルキウムラントを $K_j(h)$ ($j = 2, 3, 4$) とする．これらから，射影データの歪度と尖度が

$$B_1(h) = \frac{K_3(h)}{K_2(h)^{\frac{3}{2}}}, \quad B_2(h) = \frac{K_4(h)}{K_2(h)^2}$$

と計算される．正規性の仮定のもとでは，全ての h に対して $K_3(h)$ と $K_4(h)$ の絶対値は小さくなるという性質に基づいて，Jones and Sibson (1987) は射影追跡のためのモーメント指標

$$I_{JS}(h) = \frac{1}{6} B_1(h)^2 + \frac{1}{24} B_2(h)^2$$

を提案した．その最大値 $\max_{h \in \mathbb{S}^{q-1}} I_{JS}(h)$ は多変量正規性の検定統計量として用いることができ，また最大値をとる h が，正規性の乖離がもっとも大きい方向と解釈することができる．

多変量正規性検定の有意水準は，帰無分布のもとでの上側裾確率から評価される：

$$\Pr\left(\max_{h \in \mathbb{S}^{q-1}} I_{JS}(h) \geq a \mid H_0\right), \quad H_0 : x_t \sim q \text{ 次元正規分布.}$$

この確率を，サンプル数 n が大きいときの漸近近似で考える． \mathbb{S}^{q-1} 上で定義される連続関数の空間 $C(\mathbb{S}^{q-1})$ の中心極限定理により，法収束

$$\sqrt{\frac{n}{6}} B_1(\cdot) \xrightarrow{d} Z_1(\cdot), \quad \sqrt{\frac{n}{24}} B_2(\cdot) \xrightarrow{d} Z_2(\cdot) \quad (\text{indep.})$$

が示される．ただし $Z_1(\cdot)$ と $Z_2(\cdot)$ は独立な \mathbb{S}^{q-1} 上正規確率場で，それらの平均は 0，分散 1，相関関数は $\text{corr}[Z_1(u), Z_1(v)] = (u^\top v)^3$ ， $\text{corr}[Z_2(u), Z_2(v)] = (u^\top v)^4$ である．チューブ法によって，極限分布

$$\max_{h \in \mathbb{S}^{q-1}} n I_{JS}(h) \xrightarrow{d} \max_{h \in \mathbb{S}^{q-1}} [Z_1(h)^2 + Z_2(h)^2]$$

の上側裾確率を近似できる．その際，左辺の最大値は確率場 $u_1 Z_1(h) + u_2 Z_2(h)$ ， $(h, u) \in \mathbb{S}^{q-1} \times \mathbb{S}^1$

の最大値であることを用いる (Kuriki and Takemura, 2008a).

4.5 特異モデル：有限混合モデルのコンポーネント数検定

正則な統計モデルでは、モデルの各点で、スコア関数が有限次元の線形空間を張る。そうでない点を特異点といい、特異点を含む統計モデルを「特異モデル」という (福水 他, 2004)。

特異モデルでは、その特異点を帰無仮説にする尤度比は、漸近カイ 2 乗性がなりたたない。一つの典型例は有限混合モデルである。最も簡単な例として、2 つの 2 項分布 $\text{Bin}(k, p)$ と $\text{Bin}(k, p)$ の混合分布を考える：

$$x_1, \dots, x_n \sim \alpha \text{Bin}(k, p) + (1 - \alpha) \text{Bin}(k, 1/2) \text{ i.i.d.}$$

ここで k は定数、 α, p は推定対象のパラメータである。これは $p = 1/2$ のとき α は推定不能になり、 $\alpha = 0$ のとき p が推定不能となるという特異モデルである。

このモデルにおいて、コンポーネント数の検定

$$\begin{aligned} H_0 : \alpha = 0 \quad \text{or} \quad p = 1/2 \quad (1 \text{ コンポーネント}), \\ H_1 : \alpha \neq 0 \quad \text{and} \quad p \neq 1/2 \quad (2 \text{ コンポーネント}) \end{aligned}$$

を考える。 x_t の取り得る値は $\{0, \dots, k\}$ であり、 x_t をダミー変数表示したベクトルは多項分布

$$(\delta_{x_t,0}, \dots, \delta_{x_t,k}) \sim \text{Mult}_{k+1}(1; q_0, \dots, q_k), \quad t = 1, \dots, n, \text{ i.i.d.}$$

に埋め込むことができる。ここで

$$q_i = \frac{1}{2^k} \binom{k}{i} \{ \alpha(1 + \phi)^i (1 - \phi)^{k-i} + (1 - \alpha) \} \quad (\phi = 2p - 1).$$

特異モデルの一般論より、 H_0 vs. H_1 の尤度比検定統計量は H_0 の下で正規確率場の最大値

$$\max\{0, \sup_{|\phi| \leq 1} X(\phi)\}^2, \quad X(\phi) = \frac{2^{k/2} \sum_{i=0}^k \{(1 + \phi)^i (1 - \phi)^{k-i} - 1\} \xi_i}{\sqrt{\sum_{i=0}^k \binom{k}{i} \{(1 + \phi)^i (1 - \phi)^{k-i} - 1\}^2}}, \quad \xi_i \sim N(0, 1) \text{ i.i.d.}$$

に分布収束する。この極限分布に基づいて、チューブ法により p 値の近似評価式が得られる。

$$\Pr(\sup_{|\phi| \leq 1} X(\phi) \geq a) \approx \frac{\text{Vol}(M)}{2\pi} \Pr(\chi_2^2 \geq a^2) + \frac{1}{2} \Pr(\chi_1^2 \geq a^2),$$

$$\text{Vol}(M) = 2 \int_0^1 \frac{\sqrt{k \{(1 + \phi^2)^k - 1 - k\phi^2\} (1 + \phi^2)^{k-2}}}{(1 + \phi^2)^k - 1} d\phi.$$

4.6 特異モデル：順序分割表解析

$a \times b$ 2 元表データ $N = (n_{ij})_{a \times b}$ を、 (i, j) セルの確率が p_{ij} の多項分布の実現値としてモデル化する。 p_{ij} のモデリングとして、対応分析モデル (CA モデル; 数量化 3 類)

$$p_{ij} = p_i \cdot p_j (1 + \phi \mu_i \nu_j)$$

と RC 連関モデル

$$\log p_{ij} = \alpha_i + \beta_j + \phi \mu_i \nu_j$$

がよく知られている (Kateri, 2014)。ここで行と列が順序尺度の場合、不等式制約

$$\mu_1 \leq \dots \leq \mu_a, \quad \nu_1 \leq \dots \leq \nu_b$$

を仮定したモデルを考える．このモデルのもとでは局所対数オッズ比は非負

$$\log\left(\frac{p_{i+1,j+1}p_{i,j}}{p_{i,j+1}p_{i+1,j}}\right) = \phi(\mu_{i+1} - \mu_i)(\nu_{j+1} - \nu_j) \geq 0, \quad \forall i, j$$

となる．このモデルにおいて，独立性の仮説 $H: \phi = 0$ は， μ_i, ν_j が識別不能となるモデルの特異点である．この特異点を帰無仮説とする尤度比検定の極限帰無分布は，正規確率場の最大値の分布となり，その裾確率の近似にチューブ法を使うことができる (Kuriki, 2005)．

4.7 脳画像データ解析

K. Worsley と K. Friston は一連の研究で，fMRI データをガウス確率場とみなし，そのピークの値の統計的有意性をオイラー標数法によって評価する方法を提案した (Friston et al., 1994; Ashby, 2011)．彼らの方法は，VBM (Voxel based morphometry) データなどの脳画像解析の標準ソフト SPM に実装されている (The FIL Methods Group, 2018; 根本, 2014; 川口, 2017)．

5. おわりに

本項では滑らかなサンプルパスを持つガウス確率場の最大値の分布を求めるためのチューブ法とオイラー標数法について，その考え方と歴史を概観した．また統計学，データ解析への応用をいくつか紹介した．いままで述べたように，チューブ法は，(i) 多くの最大値型検定統計量，特に尤度比検定統計量 p 値の計算に用いることができる．(ii) 近似公式は a の有限項の漸近展開で与えられ計算が容易．(iii) 実用的な範囲 (例えば p 値 ≤ 0.3) で非常に精度がよい．(iv) $d = 1$ のとき，近似公式は保守側のバウンドを与える，といった実用的にも優れた特徴を持つ．

一方で理論的に残された問題として，添字集合の次元や体積が無限に発散する場合のチューブ法の性質は，まだほとんど解明されていない (栗木, 2019 を参照)．数学的な定式化を含めて，今後研究すべき研究分野である．

謝 辞

本稿執筆の機会を与えてくださった田村義保名誉教授，チューブ法研究の共著者である竹村彰通教授，Henry Wynn 教授，加藤直広博士，Xiaolei Lu 博士，4.7 節の内容についてご教示いただいた川口淳教授に感謝いたします．また詳細に論文を査読いただいたレフリーの方に感謝いたします．本稿は 2016 年度日本統計学受賞者記念講演での同名の講演に基づきます．関係者の方々に感謝いたします．

参 考 文 献

- Adler, R. J. and Hasofer, A. M. (1976). Level crossings for random fields, *The Annals of Probability*, **4**(1), 1–12.
- Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*, Springer, New York.
- Ashby, F. G. (2011). *Statistical Analysis of fMRI Data*, The MIT Press, Cambridge, Massachusetts.
- Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C. and Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent, *Human Brain Mapping*, **1**, 214–220.
- 福水健次, 栗木哲, 竹内啓, 赤平昌文 (2004). 『特異モデルの統計学—未解決問題への新しい視点』, 統計科学のフロンティア 7, 岩波書店, 東京.
- Gross, E. and Vitells, O. (2010). Trial factors for the look elsewhere effect in high energy physics, *The European Physical Journal C*, **70**(1–2), 525–530.

- Hotelling, H. (1939). Tubes and Spheres in n -spaces, and a class of statistical problems, *American Journal of Mathematics*, **61**(2), 440–460.
- Johansen, S. and Johnstone, I. M. (1990). Hotelling’s theorem on the volume of tubes: Some illustrations in simultaneous inference and data analysis, *The Annals of Statistics*, **18**(2), 652–684.
- Jones, M. C. and Sibson, R. (1987). What is projection pursuit?, *Journal of the Royal Statistical Society, Series A*, **150**(1), 1–37.
- Kateri, M. (2014). *Contingency Table Analysis: Methods and Implementation Using R*, Springer.
- 川口淳 (2017). 脳画像解析における統計学的クラスター推測, 日本磁気共鳴医学会雑誌, **37**(2), 39–49.
- Knowles, M. and Siegmund, D. (1989). On Hotelling’s geometric approach to testing for a nonlinear parameter in regression, *International Statistical Review*, **57**(3), 205–220.
- 小林昭七 (1997). 1940 年代, 50 年代の日本の微分幾何, 数学, **49**(3), 225–234.
- Kuriki, S. (2005). Asymptotic distribution of inequality-restricted canonical correlation with application to tests for independence in ordered contingency tables, *Journal of Multivariate Analysis*, **94**(2), 420–449.
- 栗木哲 (2019). オイラー標数法によるウィシャート行列の最大固有値分布の近似, 統計数理, **67**(2), 366–367.
- Kuriki, S. and Takemura, A. (2001). Tail probabilities of the maxima of multilinear forms and their applications, *The Annals of Statistics*, **29**(2), 328–371.
- Kuriki, S. and Takemura, A. (2008a). The tube method for the moment index in projection pursuit, *Journal of Statistical Planning and Inference*, **138**(9), 2749–2762.
- Kuriki, S. and Takemura, A. (2008b). Euler characteristic heuristic for approximating the distribution of the largest eigenvalue of an orthogonally invariant random matrix, *Journal of Statistical Planning and Inference*, **138**(11), 3357–3378.
- 栗木哲, 竹村彰通 (2008). チューブの体積と正規確率場の最大値の分布, 数学, **60**(2), 134–155.
- Lu, X. and Kuriki, S. (2017). Simultaneous confidence bands for contrasts between several nonlinear regression curves, *Journal of Multivariate Analysis*, **155**, 83–104.
- 村山齊 (2013). インタビュー「ファビオラ・ジャノッティ博士に聞く」, *Kavli IPMU News*, **24**, 40–47.
- Naiman, D. Q. (1986). Conservative confidence bands in curvilinear regression, *The Annals of Statistics*, **14**(3), 896–906.
- 根本清貴 (2014). 『すぐできる VBM : 精神・神経疾患の脳画像解析』, 学研メディカル秀潤社, 東京.
- ラオ, C. R. (1986). 『統計的推測とその応用』 (奥野忠一 監訳), 東京図書, 東京.
- Rice, S. O. (1944). Mathematical analysis of random noise, *The Bell System Technical Journal*, **23**, 282–332.
- Sun, J. (1991). Significance levels in exploratory projection pursuit, *Biometrika*, **78**(4), 759–769.
- Sun, J. (1993). Tail probabilities of the maxima of Gaussian random fields, *The Annals of Probability*, **21**(1), 34–71.
- Takemura, A. and Kuriki, S. (2002). On the equivalence of the tube and Euler characteristic methods for the distribution of the maximum of Gaussian fields over piecewise smooth domains, *The Annals of Applied Probability*, **12**(2), 768–796.
- 丹野修吉 (1976). 『多様体の微分幾何学』, 実教出版, 東京.
- Taylor, J. E., Loftus, J. R. and Tibshirani, R. J. (2016). Inference in adaptive regression via the Kac-Rice formula, *The Annals of Statistics*, **44**(2), 743–770.
- The FIL Methods Group (2018). SPM12 Manual, <https://www.fil.ion.ucl.ac.uk/spm/>.
- van Dyk, D. A. (2014). The role of statistics in the discovery of a Higgs boson, *Annual Review of Statistics and Its Application*, **1**, 41–59.
- Weyl, H. (1939). On the volume of tubes, *American Journal of Mathematics*, **61**(2), 461–472.
- Worsley, K. J. (1995). Boundary corrections for the expected Euler characteristic of excursion sets of random fields, with an application to astrophysics, *Advances in Applied Probability*, **27**(4), 943–959.

The Tube Method: Theory and Applications

Satoshi Kuriki

The Institute of Statistical Mathematics

The tube method is an integral geometric method that is used to estimate the upper tail probabilities of the maxima of Gaussian random fields. The Euler characteristic method is used for the same purpose. In this paper, the concepts underlying the tube method and Euler characteristic method are investigated, and their associated applications in statistics and recent technical developments are discussed.