

令和 4 年 6 月 15 日現在

機関番号：62603

研究種目：基盤研究(B) (一般)

研究期間：2016～2020

課題番号：16H02792

研究課題名(和文) 期待オイラー標数法の深化と実用化，および関連する数理の展開

研究課題名(英文) Developments of the theory and applications of the expected Euler characteristic method and related mathematics

研究代表者

栗木 哲 (Kuriki, Satoshi)

統計数理研究所・数理・推論研究系・教授

研究者番号：90195545

交付決定額(研究期間全体)：(直接経費) 9,600,000円

研究成果の概要(和文)：期待オイラー標数法は、確率場の最大値の近似分布を与える幾何学的方法論である。信号検出や変化点問題などの多重比較の多重性調整p値の計算のために用いられることが多く、例えば脳画像データ解析では標準ツールとして用いられている。しかしながら本方法には近似誤差評価が十分には解明されていないなど方法論として未完成な部分がある。また積分計算方法の改良等により更なる実用化を図ることができる余地もある。さらにランダム行列理論、代数統計など関連数理分野との境界領域の開拓も可能である。本研究ではこれらの各視点から総合的に期待オイラー標数法を研究し、その理論的深化と実用化、応用範囲の拡大を図る。

研究成果の学術的意義や社会的意義

データに基づく発見、すなわち統計的発見においては、常にデータのばらつきに起因する偽陽性の可能性を念頭におく必要がある。ここで偽陽性とは再現性のない発見ということができる。本研究課題である期待オイラー標数法は、偽陽性の確率を見積もるために用いられている典型的な方法であり、その適用範囲の拡大や誤差評価法の確立が望まれていた。

研究成果の概要(英文)：The expected Euler characteristic method is a geometric method to approximate the distribution of the maximum of a random field. It is available for the adjustment of the multiplicity p-value in multiple comparisons including signal detection and change point analysis. For example, it is used as a standard tool in brain image data analysis. However, this method has some immature parts as a methodology; e.g., the evaluation of approximation errors has not been fully elucidated. There is also room for further practical improvements such as efficient numerical calculations. Furthermore, it is possible to explore boundaries with related mathematical fields such as the random matrix theory and the theory of algebraic statistics. In this study, we comprehensively study the expected Euler characteristic method from these viewpoints.

研究分野：統計科学

キーワード：確率場 多重比較 積分幾何学 同時信頼領域 応用ホモロジー ランダム行列 ミンコフスキー汎関数

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様式 C-19、F-19-1、Z-19 (共通)

1. 研究開始当初の背景

添字 t を持った確率変数 $X(t)$ は確率過程とよばれるが、その添字 t がとくにベクトルの場合、そのことを強調する意味で確率場 (random field) とよばれることが多い。研究課題名にある「期待オイラー標数法」(オイラー標数法) とは、確率場 $X(t)$ の最大値 $\max_{t \in M} X(t)$ の分布を近似するための方法である。この方法を扱った最初の成書 R. J. Adler (1981), *The Geometry of Random Fields*, Wiley の表題にあるように、本方法では確率場を幾何学的対象ととらえて幾何学の言葉で最大値の確率分布を記述する。具体的には、確率場 $X(t)$ がある閾値 c よりも大きくなるような添字集合 $M_c = \{t \in M \mid X(t) \geq c\}$ をエクスカージョン集合 M_c とよび、その集合のオイラー標数 $\chi(M_c)$ の期待値 $\mathbb{E}[\chi(M_c)]$ を最大値の上側確率 $\Pr(\max_{t \in M} X(t) \geq c)$ の近似公式として用いるものである。その具体的な計算には、モースの定理 (Kac-Rice 公式) 等の微分・積分幾何が用いられる。

ところで確率場の最大値の分布は、多重比較や変化点解析における検定の p 値の多重性調整のために用いられることが多い。脳画像 (fMRI) データ解析においては、多くの信号を同時に観測する必要があるため、偽陽性の排除が本質的である。その目的で K. Worsley はこの分野にオイラー標数法を導入し、現在では標準的なツールとして確立している。また一般にはあまり知られていないが、実験物理でも同種の多重性が問題となる。例えば、新素粒子の発見のための閾値設定には、オイラー標数法の考え方が用いられている。(D. A. van Dyk (2014), *The Role of Statistics in the Discovery of a Higgs Boson*, *Annu. Rev. Stat. Appl.*, 1, 41-59.) 一方で申請者は連携研究者の竹村彰通氏と 2000 年頃から、確率ベクトルの錐へ射影や最大値型尤度比検定統計量の研究を端緒として、オイラー標数法についての一連の研究を行っている。とくに 2003 年の論文では確率場が平均 0、分散 1、共分散関数が有限展開を持つガウス確率場という応用上重要な場合において、オイラー標数法の近似誤差評価を具体的に与えた。また平行して、従来法では導出が不可能であった問題 (多層配列の最大特異値の分布、順序制約付き最大特異値の分布、モーメント射影追跡指標の最大値分布、等) においてオイラー標数法が分布の近似値を具体的に与えることができることを示してきた。

2. 研究の目的

このようにオイラー標数法は、特に統計学への応用の側面では一定の成功を収めてきたが、方法自体に理論的に未完成的な側面がある。また関連数理分野との境界領域の開拓の可能性もある。本研究の目的は、オイラー標数法を総合的に研究することであり、その範囲は

- ・期待オイラー標数法の理論的深化と実用化
- ・関連する数理との接点の研究
- ・応用分野の開拓
- ・研究集会の開催

の全項目をにらんだものとする。これらの研究を通して、方法論としてのオイラー標数法の基礎付けと適用範囲の拡大、ならびに周辺数理領域との接点の開拓を行うことを目的とした。

3. 研究の方法

目的に掲げた各項目のそれぞれについてより具体的な研究課題を設定した。

項目 I. 期待オイラー標数法の理論的深化と実用化。

本項目では、近似法としての誤差評価の研究、とくに確率場が非ガウスの場合を目標とした。またホロノミック勾配法を用いた数値計算の確立も目標とした。

項目 II. 関連する数理との接点の研究.

本項目では、オイラー標数法に基づく新しい最適実験計画法、ランダム行列理論との接点、トポロジカルデータ解析 (TDA)、特異統計モデルとの関連との関連について研究を行う。

項目 III. 応用分野の開拓.

特に宇宙データ解析への応用を念頭におくこととする。

項目 IV. 研究集会の開催.

応用数理全般を対象とする研究集会の開催する。

全体を通して国内外の研究協力者との連携をとり、いくつかのものは共同研究の形で研究を行うこととした。

4. 研究成果

I. 期待オイラー標数法の理論的深化と実用化

(1) ガウス確率場に対するチューブ法と期待オイラー標数法は、相対近似誤差が指数的小量となるという意味で精度の良い近似を与える。 t 分布から生成される t 確率場に対するチューブ法の漸近誤差を解析したところ、自由度が有限である限りボンフェロニ法を含むチューブ法の相対近似誤差は微量とはならないことが明らかになった。チューブ法近似はモデルによっては常に保守側、あるいは常に反保守側の近似を与えるが、その不等号の向きはガウスや t 分布といった分布の違いにはよらないことも結論づけられた。本結果は Electronic Journal of Statistics に刊行された H. Wynn 教授 (LSE, UK) との共著論文の一部に収録された。

(2) 非心実ウィシャート行列の最大固有値の分布は、多変量分散分析の検定の検出力の解析や MIMO における SN 比評価で必要となる。しかしながら、その分布関数の記述にはゾーナル多項式を拡張したある種の不変多項式が必要であり、数値的評価の方法は知られていない。本研究では、最大固有値分布を期待オイラー標数法で近似するという方針のもと、その数値計算の方法を検討した。特に HGM 法 (Holonomic Gradient Method) の適用を試みた (高山信毅, Yi Zhang との共同研究, Journal of Multivariate Analysis に刊行)。

II. 関連する数理との接点の研究

(3) 期待オイラー標数法の考え方から導かれる新しい最適実験計画を提案した。Naiman (1986) は、非線形回帰モデルの回帰曲線の同時信頼区間を期待オイラー標数法によって構成できることを示した。同時信頼区間のバンド幅は回帰モデルのパラメータの推定量の分散と説明変数ベクトルのとりうる集合の体積で定まる。それらはデザイン行列の関数であり、逆にバンド幅を小さくすることを目的にデザイン行列を設計するという実験計画が提案できる。本基準をフーリエ回帰モデル、あるいはそれと数学的に同等な重み付き多項式回帰モデルに適用すると、多項式回帰モデルとフーリエ回帰モデルの最適化問題にメビウス群が目的関数を保存する群として作用するという著しい特徴が見いだされ、最適化の次元を減らすことができることが明らかになった。これらは H. Wynn 教授 (LSE, UK) との共同研究であり、成果は同氏との共著として Electronic Journal of Statistics に刊行された。

(4) ランダム行列理論の最大固有値・特異値分布は、例えば対称行列の場合、その 2 次形式として定義される確率場の最大値分布としてとらえることができ、分布の裾確率評価のために期待オイラー標数法を適用することができる。実対称、複素エルミートガウスランダム行列の場合 (GOE, GUE)、期待オイラー標数法による裾確率近似分布はエルミート関数で陽に書き下すことができ、その行列サイズが発散するときの極限分布は、真の極限分布である Tracy-Widom 分布の上側裾確率を正しく近似することを見いだした。同様な結果は統計学で

より重要なウィシャート行列, 多変量ベータ行列についても成立する. 本結果は ims APRM 2018 (Singapore) の招待講演で発表され, 現時点で論文執筆中である.

(5) 確率場の分布形に関する直交多項式に基づく適合度検定の研究を行った. 直交多項式から分布の局外パラメータの影響を除去することが一つの問題である. Morris (1982) で提案された指数型分布族に付随する直交多項式の場合は, 局外パラメータに最尤推定量をプラグインしても, 2次以上の多項式については漸近的挙動が変わらないことを示すことができた.

(6) 特異点を持つ統計モデルは特異モデルとよばれ, そのモデルの解析にはオイラー標数法がしばしば用いられる. ここではサンプル数が少ないことに起因する特異性をもつ統計モデルの解析を行った. ここで扱ったモデルは分散共分散がクロネッカー積構造をもつガウスモデルである. サンプルサイズが2の場合には, マトリックスペンシルのクロネッカー標準形によってモデルが正準系に変換され, それを出発点として, 最尤推定量の存在, 一意的存在, 非存在が決定されることを明らかにした. また計算の過程では, グレブナー基底計算が有用であることも分った (Mathias Drton, Peter Hoff との共同研究, Annals of Statistics に掲載).

(7) 分割表モデル, グラフィカルモデルの一般形である A 超幾何分布の指数型分布族としての解析, とくに基準化定数の漸近評価を行った. このモデルもサンプル数が少ない場合には特異性を持つ (高山信毅, 竹村彰通と共同, Advances in Applied Mathematics に発表).

(8) Bernstein コピュラを一般化する形で B-Spline コピュラを定義した. 半無限区間に等間隔のノットを持つ B-Spline 関数のモーメントを陽の形であたえ, それを用いて B-Spline コピュラが取りうる相関の範囲を確定した (Xiaoling Dou, Gwo Dong Lin, Donald Richards との共同研究, Sankhya に掲載).

(9) オイラー標数よりもさらに情報量をもったベッチ数を用いた新しいデータ解析手法である位相データ解析 (TDA) のロバスト化の試みの研究を行い, 国際学会 NeurIPS で発表した.

III. 応用分野の開拓

(10) 一般次元の弱い非ガウス性を持った等方的確率場について, そのレベルセットの期待ミンコフスキー汎関数の期待値を, 3, 4次の高次スペクトル (すなわち 3, 4点相関関数のフーリエ変換) を用いた摂動計算の形で与えた. 期待ミンコフスキー汎関数は形態解析の分野で用いられることの多い幾何量であるが, とくに天体物理の分野では宇宙マイクロ波背景放射 (CMB) や宇宙の大規模構造 (Large-scale structure of the Universe) のデータに適用され宇宙理論の妥当性の検証に用いられてきた経緯がある. 研究結果の一部を, 天体物理系の国際誌 Physical Review D に掲載した, 引き続き宇宙シミュレータで生成されたシミュレーションデータと理論結果の比較数値実験を行った. さらに上記の計算はいわゆる形式的な計算であるため, それにどのように数学的な正当性を与えることができるか検討を行った (松原隆彦 (KEK) との共同研究). なお本項目の研究の一部は, 科学研究費基盤研究 (B) 21H03403 として遂行された.

IV. 研究集会の開催

(11) 応用数理全般を対象としたワークショップ「確率・統計・行列ワークショップ」を毎年開催した (2017年11月9日(木), 10日(金) 信州大学理学部; 2018年10月24日(水), 25日(木) 滋賀大学データサイエンス学部; 2019年11月11日(月), 12日(火) 統計数理研究所; 2020年12月14日(月) 信州大学理学部; 2021年11月15日(月) 滋賀大学データサイエンス学部, 以上共同主催者 沼田泰英, 竹村彰通).

5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 7件/うち国際共著 4件/うちオープンアクセス 1件）

1. 著者名 Satoshi Kuriki, Henry P. Wynn	4. 巻 13
2. 論文標題 Optimal experimental design that minimizes the width of simultaneous confidence bands	5. 発行年 2019年
3. 雑誌名 Electronic Journal of Statistics	6. 最初と最後の頁 1099-1134
掲載論文のDOI（デジタルオブジェクト識別子） 10.1214/19-EJS1546	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Xiaoling Dou, Satoshi Kuriki, Gwo Dong Lin, Donald Richards	4. 巻 A
2. 論文標題 Dependence Properties of B-Spline Copulas	5. 発行年 2019年
3. 雑誌名 Sankhya	6. 最初と最後の頁 1, 29
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s13171-019-00179-y	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Nobuki Takayama; Satoshi Kuriki; Akimichi Takemura	4. 巻 99
2. 論文標題 A-hypergeometric distributions and Newton polytopes	5. 発行年 2018年
3. 雑誌名 Advances in Applied Mathematics	6. 最初と最後の頁 109-133
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.aam.2018.05.001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Nobuki Takayama, Lin Jiu, Satoshi Kuriki, Yi Zhang	4. 巻 179
2. 論文標題 Computation of the expected Euler characteristic for the largest eigenvalue of a real Wishart matrix	5. 発行年 2022年
3. 雑誌名 Journal of Multivariate Analysis	6. 最初と最後の頁 104642
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.jmva.2020.104642	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Mathias Drton, Satoshi Kuriki, Peter Hoff	4. 巻 49
2. 論文標題 Existence and uniqueness of the Kronecker covariance MLE	5. 発行年 2021年
3. 雑誌名 The Annals of Statistics	6. 最初と最後の頁 2721-2754
掲載論文のDOI (デジタルオブジェクト識別子) 10.1214/21-AOS2052	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Takahiko Matsubara, Satoshi Kuriki	4. 巻 104
2. 論文標題 Weakly non-Gaussian formula for the Minkowski functionals in general dimensions	5. 発行年 2021年
3. 雑誌名 Physical Review D	6. 最初と最後の頁 103522
掲載論文のDOI (デジタルオブジェクト識別子) 10.1103/PhysRevD.104.103522	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Takahiko Matsubara, Chiaki Hikage, Satoshi Kuriki	4. 巻 105
2. 論文標題 Minkowski functionals and the nonlinear perturbation theory in the large-scale structure: Second-order effects	5. 発行年 2022年
3. 雑誌名 Physical Review D	6. 最初と最後の頁 23527
掲載論文のDOI (デジタルオブジェクト識別子) 10.1103/PhysRevD.105.023527	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計2件 (うち招待講演 1件 / うち国際学会 1件)

1. 発表者名 Siddharth Vishwanath
2. 発表標題 Robust persistence diagrams using reproducing kernels
3. 学会等名 NeurIPS 2020
4. 発表年 2020年

1. 発表者名 Satoshi Kuriki
2. 発表標題 The Euler characteristic method for multivariate analysis and random matrices
3. 学会等名 ims-APRM 2018 (招待講演) (国際学会)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
連携研究者	竹村 彰通 (Takemura Akimichi) (10171670)	滋賀大学・データサイエンス学部・学部長・教授 (14201)	
連携研究者	高山 信毅 (Takayama Nobuki) (30188099)	神戸大学・理学研究科・教授 (14501)	
連携研究者	白井 朋之 (Shirai Tomoyuki) (70302932)	九州大学・マス・フォア・インダストリ研究所・教授 (17102)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
英国	London School of Economics			
台湾	Academia Sinica			
米国	University of Washington	Duke University	The Pennsylvania State University	

共同研究相手国	相手方研究機関			
Germany	Technical University of Munich	Ulm University		