

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 20 日現在

機関番号：12601

研究種目：挑戦的萌芽研究

研究期間：2014～2016

課題番号：26540013

研究課題名(和文) シェイプ空間上の確率分布族が持つ特異構造に対する情報幾何学的アプローチ

研究課題名(英文) Information geometric approach to singular structure of the family of probability distributions on shape spaces

研究代表者

清 智也 (SEI, TOMONARI)

東京大学・大学院情報理工学系研究科・准教授

研究者番号：20401242

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：シェイプ(形状)に関する統計的推測の研究として、主に3つの成果を得た。シェイプとは、ユークリッド空間の点列を平行移動・回転・拡大縮小によって同一視して得られるデータのことである。成果の一つ目は、シェイプ空間に類似した概念としてテキスタイル集合という集合を定義し、その幾何学的性質を明らかにしたことである。この集合はテキスタイルプロットと呼ばれる可視化技法に関係している。二つ目の成果として、球面上の分布族の正規化定数が満たす微分方程式を導出した。三つ目の成果は、行列のスケーリングの応用として客観的な総合指数の決定法を提案したことである。これらの研究成果は国際学会および雑誌論文において発表した。

研究成果の概要(英文)：We obtained three main results about statistical inference on shapes, where a shape is a sequence of spatial points under identification with respect to parallel displacement, rotation and scaling. First, we defined a set called the textile set related to the shape space, and elucidated its geometric properties. Secondly, we derived differential equations satisfied by the normalizing constant of spherical distributions on the sphere. Finally, we proposed a method of determining an objective general index. The results are presented through international conferences and journal papers.

研究分野：統計学

キーワード：テキスタイル集合 情報幾何 微分幾何 総合指数 textile plot シェイプ空間

1. 研究開始当初の背景

本研究開始当初の背景には、シェイプ空間の幾何学的特異性を統計学的にどう扱うか、という問題意識があった。

シェイプ空間とはユークリッド空間の点列に対し、平行移動・回転・拡大縮小の違いを同一視して得られる集合のことである。この空間は元となるユークリッド空間が2次元の場合は特異性を持たないが、3次元以上の場合には特異性を持つことが知られている。この特異性がシェイプ空間上で定義される確率分布族や統計的推測に与える影響について、これまで十分に研究されているとは言えなかった。

2. 研究の目的

本研究の目的は、シェイプ空間上で定義される確率分布族や統計的推測の構造を明らかにすることである。この目的を達成するために以下の2つの具体的な目標を設定した。

(1) テキスタイル集合の幾何学的性質の解明

シェイプ空間に関する事前調査を進めたところ、テキスタイルプロットと呼ばれる可視化技法の理解が重要であるという認識に到った。テキスタイルプロットとはKumasaka and Shibata (2008) で提案されたデータ可視化手法である。多変量データに対し、各個体を1本の折れ線グラフに対応させるという並行座標プロットの一つであり、データ全体として可能な限り水平に近づくようにする、という基準でスケールが決定される。テキスタイルプロットの出力全体をテキスタイル集合と定義する。この集合はユークリッド空間における点列を一定の規則で同一視する、という意味においてシェイプ空間と密接に関係している。そこで、テキスタイル集合の幾何学的性質を解明することを本研究の目的の一つとした。

(2) 球面上の確率分布族の計算手法の確立

シェイプ空間上の確率分布族は、方向統計学(方向データを対象とする統計学)で用いられる球面上の分布族と密接に関連している。なぜならば、球面とはユークリッド空間において拡大縮小の違いを同一視した集合であるからである。そこで、球面上の分布族の持つ性質や計算法を明らかにすることも重要となる。本研究では特に確率分布族の計算手法を確立することを目的とした。特に、種々の確率分布族に対し、近年注目を浴びているホロノミック勾配法の適用可能性を検討することが具体的な目標である。

3. 研究の方法

本研究の2つの目的を達成するため、それぞれ以下の方法をとった。

(1) テキスタイル集合に関する幾何学的量の計算

テキスタイル集合は、与えられたサイズのデータ行列に対し、水平性基準と呼ばれる基準によってスケールして得られる行列の全体からなる集合である。この集合は陰関数の形で定義されるため、その具体的な形をイメージするのは必ずしも容易でない。そこでまず低次元の場合にどのような形になっているかを考察した。さらにそこで得た知見を手がかりに、高次元の場合の曲率や測地線の計算を行なった。また、テキスタイル集合は、データ行列にテキスタイルプロットを定義する変換を施すことにより定式化された行列全体の集合と見ることでもでき、これが部分多様体になるための十分条件を考察した。加えて、テキスタイル集合を解析幾何学の観点からも考察した。

(2) Fisher-Bingham 分布族の正規化定数に関するホロノミック勾配法の適用

Fisher-Bingham 分布族と呼ばれる球面上の確率分布族について、その正規化定数が満たすべき偏微分方程式(ホロノミック系)を調べた。特に、パラメータが退化した場合についての方程式の導出を試みた。なお、退化していない場合の偏微分方程式は既存研究(Nakayama et al. (2011) 他)によって得られていたが、退化する部分の完全な記述は得られていなかった。パラメータが退化するケースは、数学的に興味深いだけでなく、ケント分布など応用上も重要な部分クラスを含んでいる。

4. 研究成果

上記の2つの方法による成果に加え、当初は予期していなかった成果も1つ得られた。これら3つの研究成果を以下に記す。

(1) テキスタイル集合の微分幾何学的性質の解明

低次元のテキスタイル集合がどのような集合であるかを特徴付けることができた。例として、3変量3個体の場合の断面図を図1に示す。また高次元のテキスタイル集合についても、連結性、微分可能性について詳しく調べたほか、測地線や曲率の計算結果も得ることができた。さらに、微分幾何学における部分多様体論に関する「正則条件」をテキスタイル集合に対して与え、正則条件の下、実際これが部分多様体になることを示した。また、テキスタイル集合に対して解析幾何学を展開し、これを含む2次曲面を与えた。本結

果の一部は査読付き国際学会（GSI2015）において発表した。

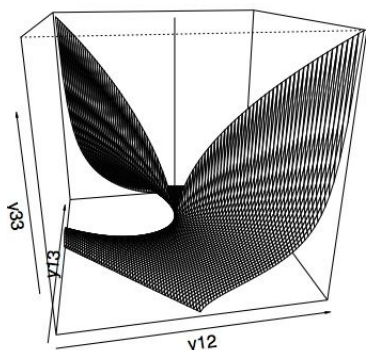


図 1. テキスタイル集合の断面（3変数3個体の場合）。

(2) Fisher-Bingham 分布族の正規化定数に関するホリノミック勾配法

Fisher-Bingham 分布族の正規化定数が満たす偏微分方程式を陽な形で導出した。特に、パラメータが退化した場合の方程式も導出することができた。具体的には、正規化定数が非心カイ二乗分布のモーメント母関数を用いて書けることに注目した。これは一種のラプラス変換と見ることができ、変換後の計算は部分分数分解を通じて見通しよく進めることができる。また、パラメータが退化したケースについては、変換後に現れる有理関数の極が重要な役割を果たすことも分かった。以上の結果により、重要なサブモデルであるケント分布などの最尤推定も扱えるようになった。本結果は共著論文として執筆し、投稿した。

当初は予期していなかった成果：

(3) 客観的総合指数の提案とその実データへの適用

シェイプ空間は、行列のスケーリングとも関連している。スケーリングとは、ある特定の性質を満たすように行列の各列（あるいは各行）を定数倍する手法の総称であり、様々な応用を持つことが知られている。本研究では、従来とは異なる応用例として、総合指数の決定法を提案した。何らかのデータからランキングを作成したいとき、一般に複数の指標を一つの指標にまとめる手順が必要となる。このとき、各指標の偏差値を求めてから単純平均を取るのが標準的な手法である。しかしこの方法は、指標の間に相関が見られる場合は、必ずしも公平でない可能性がある（図2）。提案手法では、スケーリングした結

果が、最終的に得られる総合指数と共通の共分散を持つように設計されている（図3）。このような指標が一意に定まることを示し、実データを用いてその有効性を確かめた。さらに、データがカテゴリカルな場合や、非線形変換を許した場合についても考察した。本結果は雑誌論文に投稿し、採択された（Sei 2016）。

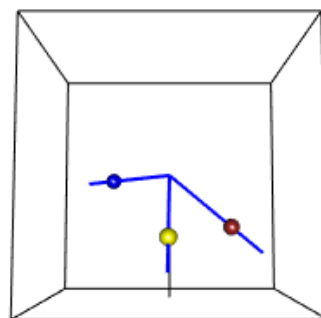


図 2. 不公平な指標。「やじろべえ」の支点から錘までの距離が各指標の重みを表す。同じ重みでは偏ってしまう。

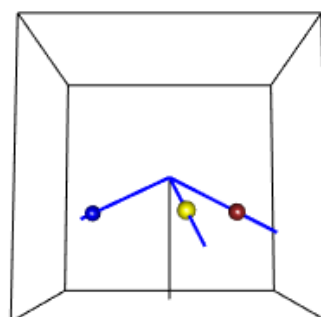


図 3. 公平な指標（提案手法）。距離を調整することで「やじろべえ」が水平につりあっている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 1 件)

1. Sei, T. (2016). An objective general index for multivariate ordered data, *Journal of Multivariate Analysis*, Vol. 147, pp. 247—264, 査読有. (doi:10.1016/j.jmva.2016.02.005)

〔学会発表〕(計 19 件)

1. 清 智也 (2017). 座標ごとの変換によって得られる Stein 型の等式とその性質, 日本数学会 2017 年度年会, 2017 年 3 月 26 日, 首都大学東京 (東京都八王子市).
2. 田中 潮 (2017). Gromov's problem: Bound the expansion coefficient from below in terms of the observable diameter of a metric measure space, and its diameter bounds, 測地線および関連する諸問題, 2017 年 1 月 7 日, 熊本大学 (熊本県熊本市).
3. 清 智也 (2016). 順序付き多変量データのための客観的総合指数 - TextilePlot との関係 -, 統計関連学会連合大会, 2016 年 9 月 7 日, 金沢大学 (石川県金沢市).
4. 田中 潮, 清 智也 (2016). 微分幾何の観点からの TextilePlot の理解, 統計関連学会連合大会, 2016 年 9 月 7 日, 金沢大学 (石川県金沢市).
5. Tanaka, U. and Sei, T. (2016). On submanifolds of textile set, The 4th Institute of Mathematical Statistics, Asia Pacific Rim Meeting, 2016 年 6 月 28 日, 香港 (中国).
6. Sei, T. and Tanaka, U. (2015). Differential geometric properties of textile plot, 2nd Conference on Geometric Science of Information (GSI2015), 2015 年 10 月 28 日, Palaiseau (フランス).
7. Kume, A. and Sei, T. (2015). Holonomic gradient methods for the Fisher-Bingham normalising constants, RSS 2015 Annual Conference, 2015 年 9 月 8 日, Exeter (イギリス).
8. 田中 潮, 清 智也 (2015). Differential geometric properties of textile plot, 統計関連学会連合大会, 2015 年 9 月 8 日, 岡山大学 (岡山県岡山市).
9. Nakamoto, K. and Sei, T. (2015). Computation of alternative

multivariate t-distributions by the holonomic gradient method, SIAM Conference on Applied Algebraic Geometry, 2015 年 8 月 7 日, Daejeon (韓国).

10. Koyama, T., Nakayama, H., Ohara, K., Sei, T. and Takayama, N. (2014). Software packages for holonomic gradient method, ICMS2014, 2014 年 8 月 6 日, ソウル (韓国).
11. Sei, T. and Kume, A. (2014). Holonomic gradient descent for the Bingham distributions, IMS-APRM2014, 2014 年 7 月 1 日, 台北 (台湾).

〔図書〕(計 1 件)

竹村 彰通, 日比 孝之, 原 尚幸, 東谷 章弘, 清 智也, 及び『グレブナー道場』著者一同, 「グレブナー教室 -- 計算代数統計への招待 --」, 共立出版, 2015, pp. 191--202.

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

清 智也 (SEI, Tomonari)
東京大学・大学院情報理工学系研究科・准教授
研究者番号: 20401242

(2) 研究分担者

田中 潮 (TANAKA, Ushio)
大阪府立大学・理学(系)研究科(研究院)・助教
研究者番号: 60516897

(3) 連携研究者

()

研究者番号:

(4) 研究協力者

()