

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 12 日現在

機関番号：22604

研究種目：研究活動スタート支援

研究期間：2015～2016

課題番号：15H06531

研究課題名(和文)代数構造を起点とする統計的推測および実験計画の研究

研究課題名(英文)Studies on Statistical Inference and Experimental Designs Based on Algebraic Structures

研究代表者

小川 光紀(Ogawa, Mitsunori)

首都大学東京・社会科学部研究科・助教

研究者番号：50758290

交付決定額(研究期間全体)：(直接経費) 1,700,000円

研究成果の概要(和文)：統計的推測の手法では、複雑な積分や大規模な和の計算が障害となって、実施が困難になることがある。このような問題を念頭にホロノミック勾配法とよばれる微分方程式を利用した数値計算の手法が提案され、盛んに研究が行われている。本研究では、Veronese 配置に付随するGKZ-超幾何系を利用したホロノミック勾配法の計算手順を与えた。また、複数の分解能を利用した実験計画の指標に関連して、これまで2水準の場合に得られていた結果を多水準の場合に拡張した。

研究成果の概要(英文)：In statistical inference, the computation of complicated integrals or large-size summation sometimes makes the problem intractable. To overcome such problems, the numerical computation method, called the holonomic gradient method that utilize the differential equations was proposed and applied to many statistical problems. In this project, we gave the concrete calculation procedure for the holonomic gradient method using GKZ-hypergeometric system associated with Veronese configurations. We also studied the problem of experimental design. Some results on designs of variable resolutions were extended to the case of multi-level designs.

研究分野：計算代数統計学

キーワード：統計計算 代数統計

1. 研究開始当初の背景

統計的推測の手法を実行する際、規格化定数の計算が困難であることが障害となることがある。たとえば、規格化定数が陽な表示をもたない多重積分の形で与えられる場合や、直接的な計算が困難な規模の和の形である場合がこれに相当する。このような問題に対して、規格化定数の効率的な計算手法を構築するアプローチと、規格化定数の計算そのものを回避するアプローチの二通りの方針が考えられる。

実験計画の分野では、効率的な一部実施計画の指標を定義し、具体的な計画を構成する手法が重要であり、多くの研究がなされてきた。レギュラーな一部実施計画のクラスは技術的に取り扱いが容易であり、このクラスに対しては古くから分解能や最小 aberration 基準のようなよい一部実施計画の指標が提案され、研究が進められてきた。一方、近年では、非レギュラーな一部実施計画の研究が盛んであり、そこでは指示関数の概念が重要な役割を果たす。

近年、計算代数統計とよばれる研究分野が急速に発展している。計算代数統計における数学的道具の基礎は、多項式環のグレブナー基底の理論である。たとえば、計算代数統計の初期の結果として、Diaconis と Sturmfels によって離散指数型分布族の条件付き分布からのサンプリング手法が考案された。この手法によれば、条件付き分布に基づく仮説検定を実施する際に、規格化定数の計算を回避することができる。また、前述の指示関数の理論も、実験計画に関する問題の代数的取り扱いと深く関係している。

2. 研究の目的

本研究の目的は、計算代数的手法によって、統計的推測や実験計画に関する具体的問題を解決することである。具体的には、次に述べる研究課題に取り組んだ。

(1) 統計的推測

分割表やランダムグラフの統計モデルに対する統計的推測では、条件付き分布を用いることが有用である場合がある。しかし、条件付き分布の規格化定数の計算が困難であることが原因で、実施することが困難な場合が多い。そのようなものの一つとして、Hardy—Weinberg モデルを一般化したモデルを帰無仮説とし、さらに交互作用を表すパラメータを追加したモデルを考える。このとき、興味の対象は新たに追加したパラメータであり、その他のパラメータは“興味のないパラメータ”になる。“興味のあるパラメータ”を精度よく推定するためには条件付き尤度に基づく最尤推定が有用であるが、規格化定数が複雑かつ大規模な和の形になってし

まうため、直接的な計算方法は現実的ではない。本研究では、この規格化定数が GKZ-超幾何系とよばれる微分方程式の非常に特殊な場合の解であることに着目し、ホロノミック勾配法とよばれる数値計算手法による解決を目指した。

当初は Hardy—Weinberg モデルの一般化という設定を主眼においていたが、研究の過程で数学的に同一な問題が、グラフィカルモデルのロバスト推定で有用な代替的 t 分布の計算に関連することが判明したため、研究の途中からは主目標を後者の問題に移して推進することとした。

(2) 実験計画

一部実施計画の理論では、少ない実施回数で効率よく識別可能性を保証することが重要である。よい一部実施計画の指標として基本的なものの一つに、分解能とよばれるものがある。この指標はある意味で汎用的なものであるが、現実に起き得る様々な状況に応じて、より適切な基準を定義し、その基準のもとでよい性質をもつ一部実施計画を採用した方がよい場合がある。その一つが、複数の分解能を用いた指標である。要因がいくつかのグループに分割されており、交互作用は各グループ内でのみあり得ることが事前にわかっている状況を考える。このような場合、全体としての分解能だけでなく、各グループにおける分解能も指標に組み込むことで、事前知識を反映したより適切な一部実施計画を考えることができる。複数の分解能を用いた議論は、先行研究では2水準の要因に限定して行われていたが、本研究ではその理論や関連する計画の構成手法を多水準の場合に一般化することを目指した。

3. 研究の方法

(1) 統計的推測の研究課題では、ホロノミック勾配法とよばれる数値計算手法の枠組みを用いた。ホロノミック勾配法は、2011年に Nakayama et al.によって提案された数値計算の手法であり、計算対象となる関数を解にもつ微分方程式を利用して数値計算を実行する。その数学的基礎となっているのは微分作用素のなす代数構造であり、この手法を適切に実施できれば、有理関数を成分とする行列計算のみで数値計算を実行することができる。

実際にこの手法を適用するためには、計算対象の関数を解にもつ微分方程式の導出、パフィアン系とよばれる連立の常微分方程式の導出、パフィアン系の基底に含まれる関数に対する初期値計算手法の構築などが必要である。本研究では、初年度に主にパフィアン系に関する理論的考察を行い、次年度にその整備と数値計算への実装の作業を行なった。

本研究で考える問題では、ホロノミック勾配法で必要になる微分方程式は、Veronese 配置に付随する GKZ-超幾何系として得られることが比較的容易に確認できる。しかし、対応するパフィアン系は複雑な形をしており、その全体を陽に書き下すことは困難である。そこで、パフィアン系の計算をいくつかの計算手順に分割した形で導出した。当初に得られた手順には、一部の計算方法が一意に定まらないことがあり、実装の際の妨げとなっていた。そこで、計算手順の再整備を行い、より実装に適した計算手順を導出した。

パフィアン系の計算手順の他に、特別な状況における初期値計算の方法を与え、それらを実装し、本研究で得られた手法の性能について調べた。

(2) 多水準の一部実施計画を扱うための数理的道具として、各要因の水準を 1 の n 乗根によるエンコードと指示関数の理論を用いた。指示関数の理論は、実験計画の分野におけるグレブナー基底の理論の最初の応用である Pistone と Wynn の結果に刺激を受けて、Fontana et al. が 2 水準の場合に発展させたものである。Fontana et al. の研究では繰り返しなしの 2 水準計画の場合に限定して議論が行われていたが、その後 2 水準の繰り返しありの場合や、いくつかのエンコードの仕方のもとでの多水準の場合へと理論が拡張されている。今回用いた 1 の n 乗根へのエンコードでは、交互作用項に対するある種の直交性が成り立つため、指示関数を用いた計算を素直に実行することができる。

以上の準備のもとで、2 水準の場合の複数の分解能を用いた指標の構成法を自然な形で多水準の場合にも拡張した。さらに、いくつかの基本的性質を有する計画から、複数の分解能を用いた指標についてよい性質をもつ計画を構成する手法についても検討した。

4. 研究成果

(1) 代替的 t 分布の計算や Hardy-Weinberg モデルを一般化したモデルにおける興味あるパラメータの推定に対して応用することのできる関数の数値計算手法を、ホロノミック勾配法に基づいて与えた。具体的には、対象とする関数が Veronese 配置の特別な場合に付随する GKZ-超幾何系を満たすことを確認し、対応するパフィアン系の計算を再現する計算手順を与えた。この計算手順を実行する際、パフィアン系に特異点が存在すると、計算が困難になる。今回の主目的である代替的 t 分布への応用では、計算すべき領域では特異点をもたないことから、導出した計算手順を実行することが可能である。以下、その概要を具体的に述べる。

ホロノミック勾配法に基づく数値計算手法を確立するためには、

計算対象とする関数を解にもつ微分方程式を十分多く発見する。

得られた微分方程式をもとに、パフィアン系とよばれる連立の常微分方程式を導出する。

計算対象の関数およびそれを適切に微分した関数を有限個並べたパフィアン系の基底に対し、計算を開始する初期点における値を数値計算する方法を与える。という三つの要素が必要である。

本研究の場合、を満足する微分方程式として、Veronese 配置の特殊な場合に付随する GKZ-超幾何系という微分方程式を採用した。

次に、におけるパフィアン系の導出を行った。パフィアン系の形は全体としては非常に複雑であるが、より簡便な計算を系統的に積み重ねる形での計算手順を与えた。この計算手順の導出は、Veronese 配置に付随するトーリックイデアルのグレブナー基底を理論的に考察するとき有用な、ソーティング作用素の考え方に基づいている。研究前半に得られたパフィアン系の計算手順を再度見直すことで、パフィアン系計算の実装の観点からも見通しのよい計算手順が得られた。さらに、得られたパフィアン系の基底が最小サイズのものであることも理論的に確認した。このことは、今回採用している GKZ-超幾何系に基づくホロノミック勾配法において、本研究が与える計算手順が計算コストの観点からある意味で最小を達成していることを意味する。

の初期値計算に関しては、ある種の独立性が成り立つ特殊な場合に対する計算対象の関数の陽な表示を用いることとした。また、これに近い場合には元の多重積分を複数個の一次元積分の計算によって数値計算できることもわかった。

最終年度には、本研究で与えた手法の試験的な実装を行った。複数個の一次元積分を用いて精度よく計算可能な場合を用いて提案手法の性能を確認したところ、このような特殊な場合については適切な実装のもとでは精度よく計算できていることがわかった。また、計算時間について計測したところ、理論通りのオーダーで計算時間が増加している様子がうかがえた。一方、より一般の場合について適用し、モンテカルロ法などの他の近似計算手法と比較したところ、不安定な場合があることが判明した。代替的 t 分布における統計的推測への応用など、実践的な状況を考えれば、一般の場合についても数値的に安定して動作することが重要であり、本手法の実装面での改良が今後の課題として残った。

(2) 複数の分解能を用いたよい一部実施計画の評価基準について、従来は 2 水準の場合に限定して議論されていたものを多水準の場合に拡張した。指示関数の理論を多水準に拡張する際に用いられた一つのアプローチである 1 の n 乗根によるエンコードを採用し、

この設定のもとで aberration 基準などを定義する場合の流儀に従って分解能を用いた指標を定義した。この設定のもとで、2水準のときに J-特徴量とよばれる量に対して成り立つ有用な性質が、多水準の場合にも自然に拡張されることを示した。この性質を用いて、複数の分解能を用いた指標についてよい性質をもつ一部実施計画を、より小さな一部実施計画から構成する手法を与えた。具体的には、次の二つの構成法を与えた。

(a)各グループ内の要因数が小さい計画から、各グループがより多くの要因を含む計画を、分解能に関する性能を維持しつつ構成する方法。

(b)要因に関するグループの数が小さい計画から、より多くのグループの計画を、分解能に関する性能を維持しつつ構成する方法。

これらの結果は、2水準で知られていた結果を指示関数の理論を用いて多水準の場合に拡張できることを利用したものである。

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)

[学会発表](計2件)

小川光紀：

状態空間モデルによる独立構造を伴う動的ネットワークのモデリング。

2016年度統計関連学会連合大会，金沢大学，石川県，2016年9月4-7日。

Mitsunori Ogawa：

Some results on multi-level factorial designs in complex coding.

The 4th Institute of Mathematical Statistics Asia Pacific Rim Meeting, The Chinese University of Hong Kong, China, June 27-30, 2016.

6. 研究組織

(1)研究代表者

小川 光紀 (OGAWA, Mitsunori)

首都大学東京・社会科学部研究科・助教

研究者番号： 50758290

(2)研究分担者

なし

(3)連携研究者

なし

(4)研究協力者

なし