

科学研究費助成事業 研究成果報告書

平成 28 年 4 月 7 日現在

機関番号：32629

研究種目：若手研究(B)

研究期間：2012～2015

課題番号：24700274

研究課題名(和文) 統計的因果推論のための計算代数的手法に関する研究

研究課題名(英文) Computational Algebraic Methods for Statistical Causal Inference

研究代表者

田中 研太郎 (Tanaka, Kentaro)

成蹊大学・経済学部・准教授

研究者番号：00376948

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：本研究では、計算代数の方法を統計的因果推論に応用することで、新しい統計的推測方法につながる手法を構築した。例えば、2つの薬を同時に服用した時にのみ副作用が起こる、という現象をグラフを用いて表現することを考えると、実はそのような表現は不可能(または不自然)であるということが分かる。本研究では、そのような問題点を克服するために、代数的な手法を採用した。そして、その手法を用いて、統計的因果推論において条件付き独立性の関係を自動的に導出する新しいアルゴリズムを構築した。

研究成果の概要(英文)：Conditional independence is a fundamental concept in statistics and applied in a variety of fields. In order to analyze the properties of conditional independence, we deal with the implication problem of conditional independence statements, that is, testing whether a conditional independence statement is derived from a set of other conditional independence statements. In this research, we propose a new algebraic method for the implication problem among conditional independence statements, which is inspired by the factorization characterization of conditional independence.

研究分野：統計科学

キーワード：統計的因果推論 計算代数統計 条件付独立性

1. 研究開始当初の背景

近年、グレブナー基底を用いた計算代数の方法を統計学に応用した「計算代数統計」という分野が急速に発展している。一方、与えられたデータに基づいて事象の因果関係を推論する統計的手法のことを統計的因果推論と呼ぶ。統計的因果推論では、確率モデルの表現手法としてグラフを用いたグラフィカルモデルと呼ばれる手法が広く用いられている。これらの表現は、その表現に対応する確率モデルにおいて成立する条件付き独立性の構造を表している。ところが、実は、グラフでは表すことのできない条件付き独立性の構造が多数存在することが知られている。つまり、グラフによる表現が不十分であることが知られている。これらのグラフによる表現手法における欠点を克服するための手法として計算代数統計を用いた手法が注目されている。

2. 研究の目的

本研究は、計算代数統計と統計的因果推論の融合を目指している。データからその背後にある因果関係を推測するためには、条件付き独立性の代数的表現が非常に重要であることが近年の研究で明らかになりつつある。本研究では、今までは直観や発見的手法に頼っていた因果関係の構築方法について、計算代数を用いた手法に置き換えることを可能にする理論の開発を行う。その結果、直観や発見的手法では見過ごされがちな構造を取りこぼすことが無くなり、より客観的な確率モデルの推測方法の構築につながると期待される。さらに、条件付き独立性の推論ルールの効率的な列挙方法の構築にも取り組む。

3. 研究の方法

まず、変数の数が小さい単純な確率モデルの場合から理論的な解析を始めた。そして、確率モデルに対応しない条件付き独立性の代数的表現を列挙し整理していくことで、その性質の解明に取り組んだ。本研究は、計算代数統計と統計的因果推論の融合を目指して条件付き独立性の代数的表現について研究するという非常に新しい話題を扱っており、そのため、共同研究者が得られにくいといった問題点があった。しかし、幸いにも竹村彰通教授、宮川雅巳教授、清智也准教授などの研究協力者に恵まれた。さらに、この分野の先駆者である Studeny 教授との共同研究を行うことで、条件付き独立性の推論ルールを判定する新たなアルゴリズムの開発に取り組むことができた。国内外の様々な学会に参加して、研究についての情報の収集と発信を行うことで、実際の場面で求められている点についてのご指導を頂いた。それらに基づいて応用面を視野に入れた理論の展開を目指して研究を行った。

4. 研究成果

平成 24 年度においては、条件付き独立性の代数的表現(imset)についての理論的な解析に取り組んだ。そして、条件付き独立性に対して成り立つ関係式(推論ルール)を自動的に導出する新しいアルゴリズムを構築することに成功した。しかし、そのアルゴリズムは、確率分布が正の確率のみで成り立っている場合においてのみ適用可能な状況であった。平成 25 年度においては、この点を改善すべく、今までのアルゴリズムに他の新しいアルゴリズムを付け加えることを考えた。その結果、確率が 0 をとりうる場合でも条件付き独立性の推論ルールが適切に導出されるようなアルゴリズムを構築することに成功した。さらに、imset を拡張した概念を用いることによって、この新しく開発されたアルゴリズムをより効率的に動かすことができるということも示すことができた。平成 26 年度においては、前年度までに得られていた imset による推論ルールの導出方法を、多項式環の言葉で表現しなおすことができた。これにより、トーリックイデアルにおける種々の手法が条件付き独立性の推論ルールの問題に適用できるようになった。imset から作られる錘やその双対錘である優モジュラ関数から作られる錘の構造は、確率分布に対応する部分が非常に複雑な構造をしていることが分かってきている。その構造の解明のためにも、異なる数学的表現への翻訳はとても重要である。最終年度の平成 27 年度においては、前年度までに得られていたアルゴリズムを数理計画法の手法を使って R 言語において実装し、それらの性能を試すための計算機実験を行った。これらの研究成果については、学会発表や雑誌論文などで報告している。また、研究で得られたアルゴリズムを実装したプログラムについては、<https://github.com/tanaken-basis/>において公開している。

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 4 件)

(1)

Kentaro Tanaka, Milan Studeny, Akimichi Takemura and Tomonari Sei.

“A linear-algebraic tool for conditional independence inference”, *Journal of Algebraic Statistics*, **6**(2), 150-167, 2015. (査読あり)

(2)

鈴木理史, 田中研太郎, 宮川雅巳.

“カンファレンス行列を用いた 3 水準スクリーニング実験における追加試行”, *品質*, **44**(4), 65-74, 2014. (査読あり)

(3)
高野祐一, 田中未来, 鮭川矩義, 竹山光将, 神里栄, 千代竜佑, 小林健, 田中研太郎, 中田和秀.

“ファジィクラスタワイズ回帰を用いた 共同購入型クーポンサイトの閲覧傾向分析”, *オペレーションズ・リサーチ*, **59**(2), 81-87, 2014. (査読あり)

(4)
Takuya Kashimura, Tomonari Sei, Akimichi Takemura and Kentaro Tanaka.

“Cones of Elementary Imsets and Supermodular Functions: A Review and Some New Results”, *Proceedings of the Second CREST-SBM International Conference "Harmony of Groebner Bases and the Modern Industrial Society"*, 117-152, 2012. (査読あり)

[学会発表] (計 16 件)

(1)
田中研太郎,
“実験計画法と Group Lasso”,
第 27 回 RAMP シンポジウム,
2015 年 10 月 15 日～10 月 16 日, 静岡大学.

(2)
Kentaro Tanaka.
“Linear algebraic methods and Groebner bases for experimental design and group lasso”,
Current Trends on Groebner Bases,
Osaka, Japan, 1-10 July, 2015.

(3)
田中研太郎, 宮川雅巳.
“実験計画法における Group Lasso の利用”,
科研費シンポジウム「統計的推測の理論的基礎とその応用」,
2014 年 12 月 1 日～12 月 2 日, 筑波大学.

(4)
田中研太郎.
“条件付き独立性に関する問題について”,
名古屋統計セミナー,
2014 年 9 月 11 日, 名古屋大学.

(5)
田中研太郎.
“条件付き独立性と線形計画法”,
2014 年度 統計関連学会連合大会,
2014 年 9 月 14 日～16 日, 東京大学.

(6)
Kentaro Tanaka, Milan Studeny,
Akimichi Takemura, Tomonari Sei.
“Linear programming methods for Conditional Independence Inference”,
Joint Workshop on Limit Theorems and Algebraic Statistics, Prague,
Czech Republic, 25-29 August, 2014.

(7)
Kentaro Tanaka, Masami Miyakawa.
“The Group Lasso Method for Finding an Optimal Design”,
The 3rd Institute of Mathematical Statistics Asian Pacific Rim Meeting,
台北, 中華民国, 29 June - 3 July, 2014.

(8)
田中研太郎, Milan Studeny, 竹村彰通,
清智也.
“Algebraic methods for conditional independence inference”,
日本数学会 2013 年度秋季総合分科会,
2013 年 9 月 24 日～27 日, 愛媛大学.

(9)
田中研太郎, 宮川雅巳.
“実験計画法, 数理計画法, Group Lasso”,
2013 年度 統計関連学会連合大会,
2013 年 9 月 8 日～11 日, 大阪大学.

(10)
Kentaro Tanaka, Milan Studeny,
Akimichi Takemura, Tomonari Sei.
“Machine Learning Methods for Conditional Independence Inference”,
Bernoulli Society Satellite Meeting to the ISI World Statistics Congress 2013, Asymptotic Statistics and Related Topics: Theories and Methodologies,
Tokyo, Japan, 2-4 September, 2013.

(11)
田中研太郎, Milan Studeny, 竹村彰通,
清智也.
“条件付き独立性と線形代数”,
統計数学セミナー,
2012 年 12 月 7 日, 東京大学.

(12)
田中研太郎, Milan Studeny, 竹村彰通,
清智也.
“条件付き独立性と代数統計”,
数理統計学の沃野,
2012 年 11 月 23 日～24 日, 慶應義塾大学.

(13)

田中研太郎, Milan Studeny, 竹村彰通,
清智也.

“計算機代数的手法の統計的因果推論への
応用”,
松江シンポジウム (統計科学における深化と
横断的展開),
2012 年 10 月 24 日~26 日, 島根県.

(14)

Kentaro Tanaka, Akimichi Takemura,
Tomonari Sei.

“A new method for Conditional
Independence Inference”,
*The 9th Workshops on Uncertainty
Processing*, Marianske Lazne,
Czech Republic, 12-15 September, 2012.

(15)

Kentaro Tanaka, Masami Miyakawa.

“Optimization problems for intervention
effects in structural equation models”,
*The 2nd Institute of Mathematical
Statistics Asia Pacific Rim Meeting*,
Tsukuba, Japan, 2-4 July, 2012.

(16)

田中研太郎.

“条件付き独立性における関係式の導出”,
研究集会「現代の産業社会とグレブナー基底
の調和」,
2012 年 6 月 1 日~2 日, 東京大学.

[図書] (計 1 件)

(1)

“日本統計学会公式認定 統計検定 1 級対応
統計学”,
二宮嘉行, 大西俊郎, 小林 景, 椎名 洋,
笛田 薫, 田中研太郎, 岡田謙介, 大屋幸輔,
廣瀬英雄, 折笠秀樹, 日本統計学会(編集),
竹村彰通(編集), 岩崎学(編集) (担当:分担
執筆, 範囲:第 6 章), 東京図書 2013.

[その他] (計 1 件)

(1)

研究成果の R のプログラムを載せた
Web ページ
<https://github.com/tanaken-basis/>

6. 研究組織

(1) 研究代表者

田中 研太郎 (TANAKA KENTARO)
成蹊大学・経済学部・准教授
研究者番号: 00376948

(2) 研究分担者

なし

(3) 連携研究者

なし