

位相データ解析における因果性とその系統学への応用

研究代表者 谷口 正信
(基幹理工学部 応用数理学科 教授)

1. 研究課題

因果性解析は、経済学者 Granger が、2本の経済時系列について、一方の系列の過去情報のみを使った予測誤差より、双方のそれを使った予測誤差が小さくなる時、関与の時系列に対して他系列からの因果性があるという概念を提唱したことから始められた。その後、この統計的議論は、経済、金融だけでなく、医学、遺伝子、グラフ、ネットワークにまで、広げられてきた。研究代表者谷口はすでにオーストラリアの感染症の州ごとの推移に関する Granger 因果性についての基礎的研究を遂行し今まで指摘されてない因果性を見出している。一方点群データに対してその位相的な情報を探る解析は位相データ解析(TDA)と呼ばれている。従来の統計解析では、突然の大変化を事前につかむのが難しいが、TDAには、それができる可能性が指摘されている。実際、金融データに対して位相データ解析の手法から得られた指標が金融危機の兆候を事前に検出することが報告されている。本研究では、TDAを用いた因果性解析の基礎理論構築とその最適統計推測論を展開し、金融データの突然の大変化の事前予測を因果性因子を使って行う。また、このメカニズムをつかむ統計モデルの提案をするとともに、生物系統学的なデータ解析への応用可能性を探索する。

2. 主な研究成果

統計変量はシグマ加法族で記述できるので、2つのシグマ加法族に対して一般的な離反度を提案し、これから一般化因果性を定義する。まず、この量を観測から推測する最適統計推測理論を構築し、極めて一般的な観測；経済・金融データ、医学・遺伝子データ、グラフ・ネットワークでの因果性、隠れ因子の推測への応用を行うが、初期解析としては、局所定常時系列の局所時間の尤度から求められる局所時間推定量を導き、そのバーコードからパーシステント・ランドスケープを求め、そこから位相情報を抽出し、まずは、シュミレーションで基礎モデルの有効性の検証をおこなった。また結果を脳波データや金融データに適用して新しい因果性を見た。さらに生物系統学や感染症医学の研究者とも協働して一般化因果性の生物学的な意義や有用性を専門的に考察し、因果性解析・TDA・系統解析の融合可能性について多角的に検討する。生物が辿った進化の道筋や進化的な類縁関係を解明する系統解析においては、系統樹の拡張版である「系統ネットワーク」というグラフを活用したデータ解析手法が強く求められているが、系統ネットワークは巨大なグラフのクラスであるため系統樹とは異なり数学的・計算機科学的に扱いづらく、有意義な問題を効率的に解くアルゴリズムはほとんど知られていないのが現状であった。しかし早水がこれまでの研究で証明した。「系統ネットワークの構造定理」は、系統ネットワークからノイズを除去してできる

全域系統樹という系統樹に関する様々な問題に対して線形時間（あるいは線形時間遅延）アルゴリズムを導出するものであった．今年度のプロジェクト研究では系統ネットワークの構造定理の応用問題のさらなる開拓を行い，既存のアルゴリズムでは $O(n^{3/2})$ 時間を要する。「逸脱度評価問題」（リーフを何個足せば系統ネットワークの中に全域系統樹が存在するようにできるかを求める問題）に対して初めて $O(n)$ 時間のアルゴリズムを与え，この問題の時間計算量が $\Theta(n)$ であることを解明した．さらに，全域系統樹の数え上げに関する新たな複数の未解決問題も提起し，それら#P 完全であるという予想を発表した (arXiv:1811.05849v4)．

3. 共同研究者

清水 泰隆 早稲田大学・理工学術院・教授

早水 桃子 早稲田大学・理工学術院・専任講師

劉 言 早稲田大学・数理科学研究所・講師

Xu Xiaofei 早稲田大学・理工総研・次席研究員 基盤 (S) (谷口) による雇用

池 祐一 早稲田大学・理工総研・客員次席研究員 基盤 (S) (谷口) による

後藤 祐一 早稲田大学・基幹理工学研究科・学振 DC2

4. 研究業績

4.1 学術論文

4-1. Liu, Y., Xue, Y. and Taniguchi, M. (2020). Robust linear interpolation and extrapolation of stationary time series in L_p . J. Time Ser. Anal. 41, 229-248.

4-2. Goto, Y. and Taniguchi, M. (2020) Discriminant analysis based on binary time series. Metrika 83, 569-595 DOI.org/10.1007/s00184-019-00746-1

4-3. Xue, Y. and Taniguchi, M. (2020) Local Whittle likelihood approach for generalized divergence. Scand. J. Statist. 4, No. 1, 182-195. DOI:10.1111/sjos.12418

4-4. Akashi, F., Taniguchi, M. and Monti, A.C. (2020) Robust causality test of infinite variance processes. J. Econometrics 216 No. 1, 235-245..

4-5. Xue, Y. and Taniguchi, M. (2020) Modified LASSO estimators for time series regression models with dependent disturbances. Statistical Methods and Applications
DOI 10.1007/s10260-020-00506-w

4-6. Taniguchi, M., Kato, S., Ogata, H. and Pewsey, A. (2020) Models for circular data from time series spectra. J. Time Series Anal. 41: 808 – 829.
DOI:10.1111/jtsa. 12549

4-7. Chen, C.W.S., Lee, S., Dong, M.C. and Taniguchi, M. (2021). What factors

drive the satisfaction of citizens on governments' responses to COVID-19 ?
International J. of Infectious Disease 102, 327 – 331.
DOI:org/10.1016/j.ijid2020.10.050

4-8 Shimizu, Y., Minami, Y. and Ito, R.(2020). Why does a human die? A structural approach to cohort-wise mortality prediction under Survival Energy Hypothesis, ASTIN Bulletin. 2020, 51, (2) 1-29; (Open Access).

4-9 Shimizu, Y., (2020). Asymptotic distributions for estimated expected functionals of general random elements, arXive. 2020; (arXive.2005.02532.).

4-10 Shimizu, Y. and Zhang, Z.(2020). Asymptotically normal estimators of the ruin probability for Lévy Insurance surplus from discrete samples, Risks MDPI. 2019, 7 (2), 1-22; (Article: DOI: Open Access).

4-11. Martin Royer, Frédéric Chazal, Clément Levrard, Yuhei Umeda, and Yuichi Ike; ATOL: Measure Vectorization for Automatic Topologically-Oriented Learning Proceedings of 24rd International Conference on Artificial Intelligence and Statistics (AISTATS 2021)

4-12. Yasuaki Hiraoka, Yuichi Ike, and Michio Yoshiwaki; Algebraic stability theorem for derived categories of zigzag persistence modules preprint (2020), arXiv:2006.06924

4-23 Momoko Hayamizu, Katharina T. Huber, Vincent Moulton, Yukihiro Murakami. Recognizing and realizing cactus metrics, Information Processing Letters 157(105916) 2020. DOI:10.1016/j.ipl.2020.105916 (査読付き国際誌)

4.2 総説・著書

2020 年度はなし。

4.3 招待講演

(4 – 1) 講演題目: Time Series Analysis under Non-Standard Settings

By Masanobu TANIGUCHI, Waseda University

“Statistical Topological Data Analysis Workshop, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, 6 January”

(4 – 2) Liu, Y., Taniguchi, M. and Ombao, H. Statistical Inference for Persistence Landscapes of the Granger Causality. Waseda International Symposium. February 27, 2020.

(4-3) Liu, Y., Taniguchi, M. and Ombao, H. Statistical Inference for Persistence Landscapes of the Granger Causality. Waseda Cherry Blossom Workshop. March 21, 2020.

(4-4) Liu, Y., Taniguchi, M. and Ombao, H. Statistical Inference for Persistence Landscapes of the Granger Causality. Waseda International Symposium. February 25, 2021.

(4-5) Liu, Y., Taniguchi, M. and Ombao, H. Statistical Inference for Persistence Landscapes of the Granger Causality. Waseda Cherry Blossom Workshop. March 19, 2021.

(4-6) Goto, Y., T. Kaneko, S. Kojima, M. Taniguchi "Likelihood Ratio Processes under Non-Standard Settings", Waseda International Symposium
"Topological Data Science, Causality & Time Series Analysis", Waseda University.

(4-7) Ike, Y. Topological data analysis from a practical and mathematical perspective, iTHEMS Math Seminar, RIKEN, online, July 2020.

(4-8) Ike, Y. Convergence Result of Stochastic Subgradient Descent for Persistence-based Functions, Waseda Cherry Blossom Workshop on Topological Data Science, online, March 2021.

(4-9) Momoko Hayamizu. A structure theorem for tree-based phylogenetic networks: from theory to algorithms. Waseda Cherry Blossom Workshop on Topological Data Science, 19 March 2021. (招待講演)

4.4 受賞・表彰

- (i) Taniguchi M. "Distinguished Author Award 2020", Journal of Time Series Analysis 誌 (英国)
- (ii) 後藤祐一 早稲田大学 小野梓学術賞受賞 (2021年3月)

4.5 学会および社会的活動

谷口 正信 : (i) 国民生活の基礎となる年金について、国民の巨大な年金積立金(150兆円)をポートフォリオで運用している GPIF の調査数理室とは運用の基礎数理に関する共同研究を推進している。

(ii) 自動車産業とは“燃料電池に係る共同研究”で基礎数理面から貢献している。
学術誌編集 :

(iii) Editorial Board, Statistical Inference for Stochastic Processes (仏誌), 2007 -

現在。

(iv) Editorial Board, *Mathematica Japonicae* (国内誌), 1998- 現在。

学術団体役職：

(v) Institute of Mathematical Statistics(USA), 1987-

米国数理統計学会特別会員(Fellow)

(vi) International Statistical Society (国際統計協会), 1987-

正会員(Elected Member)

清水泰隆：日本アクチュアリー会理事会客員、日本保険・年金リスク学会(JARIP) 理事

早水桃子：特記事項 (インタビュー取材)

「新しい数学の理論を創出し、生命科学の進歩に貢献する」, 朝日新聞出版 早稲田理工
by AERA 2021, REVIEW 明日をつくる6人の研究者, 2021年2月26日.

5. 研究活動の課題と展望

2つのシグマ加法族の離反性から因果性を導入すれば、その適用範囲はグラフ、遺伝子、トポロジカルな対象まで飛躍的に広がる。またこの量への最適推測論の構築も極めてチャレンジングである。さらに複数の観測を点群データに変換し、パーシステント・ランドスケープという位相グラフ化し、これらの一般化離反度から、因果性、隠れ因子を推定すれば、従来の金融・統計手法では捉えられなかった突然の大変化の事前因子、隠れ因子が、一部見出せてきつつある。これは、従前の統計解析ではできなかった革新的手法になりうると思われる。またこのカラクリを説明する統計数理理論の構築は、極めて新奇性のあるものである。特に位相データ解析手法が、どのような信号を抽出するか？どのようなノイズに対してロバストか？この主張の理論構築は基礎理論として極めて独創性を持つもので、諸方面に大きな影響を及ぼすと思われる。また、この解析を系統解析に適用すれば、いままでなされなかった因果性・TDA・系統学と3分野の協業による革新的なパラダイムが出現し、進化生物学や感染症医学を含む生命科学分野への波及効果も生み出せるであろう。

谷口は、因果性研究を関数時系列、位相データ解析やカテゴリーカル・データ解析につなげる。また長年文理融合の研究を進めてきた。特に、計量経済、計量ファイナンスのトップジャーナルレベルで研究発展をはかっている。これらは、年金積立金運用独立行政法人や自動車産業との協業にも展開する。

また、谷口は基盤(S)研究による、欧州ワークショップ(イタリアの Rome, Bologna, Bergamo, ベルギーの Gent の4か所)を開催し、時系列解析、金融解析、球面分布解析等の早稲田研究をグローバル、かつ企業や官庁研究まで展開する。

清水は、外部資金としては科研費基盤研究(A)等に応募予定。また、国際研究集会については、ドイツとの金融保険関連との集会や、2021年5月-6月頃にはスイス(ETH)の保険数理教授を招聘し、国際共同研究と日本アクチュアリー会と共同でソルベンシーに関する産学共同研究集会を行う予定。保険・金融数理をグローバルかつ産学共同研究の形で展開する。

劉は、国立研究開発法人科学技術振興機構(JST)のさががけ「信頼されるAIの基盤技術」、同じく創発的研究支援事業に応募予定で、加えて、科研費ではステップアップを図り、基盤研究(B)を申請する予定。また、経済・社会的諸課題を横断的に解決すべく、数理的手法を

切り口として、計数時系列を含めた現代的統計推測論、データサイエンスを発展させる。