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

本プロジェクト研究では、Granger 因果性などを含む高度な因果指標を極めて一般的な乖離度から導入して、データ科学における今まで捉えられなかった潜在要因の統一的指標を提案することを目論んだ。この流れで、まず、多次元局所定常時系列に対して時変スペクトル行列に対する乖離度による多次元時系列の、ある成分ベクトルから他の成分ベクトルへの因果性のある、なしの検定手法の提案をする。局所定常過程の事変スペクトル行列に細谷因果性指標をいれこれに基づき検定統計量を導入し因果性検定を行った。また、金融データは2次モーメントを持たないと実証認識があるので、そのような現象の数理モデルである安定過程に対する Granger 因果性の研究を経験尤度解析と絡めて展開中である。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応用を脳波データや金融データに適用し、因果性因子の有用性を見ている。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求する。また保険数理では Levy 過程からの離散観測に基づく推測論の構築と surplus 過程へのプライシングも行った。

2. 主な研究成果

(2-1) 局所指標に基づく Granger 因果性と位相データ解析とその応用

この課題は上述下内容で遂行された。

(2-2) 安定過程に対する Granger 因果性

金融データは2次モーメントを持たないという実証認識があるので「多変量安定過程の応答行列に対して Granger 因果性を導入し、因果性検定を経験尤度により展開した。この流れは金融分野に新地平を築くもので、今後、この研究を強力に進める予定である。またポートマンントウ検定のカラクリを2段階 Whittle 尤度比から明らかにし、ポートマンントウ検定が漸近カイ2乗分布に従うための十分条件を与えた。

(2-3) 時系列 ANOVA モデルでランダム効果がある場合の漸近推測論の構築

独立標本に対して、分散分析は、古典的な話題であるが ANOVA モデルで攪乱項が多次元確率過程で群間平均が等しいという仮説のもとランダム効果のある分散分析モデルで、ランダム効果、有り無し
の検定統計量の漸近理論は、通常のそれと大きく異なることが判明し、興味ある理論が進行中である。また時系列攪乱項のあるモデルへの基礎理論構築も LAN 性が成り立たない場合も出てきて興味ある解析が進展中である。

(2-4) 計数過程の強度関数の推測

本研究では、計数過程の強度関数の推測と、その応用を展開した。犯罪発生率やコロナ感染率はこのモデルで記述される。基礎理論では、この強度関数の母数推定や非母数的推定論、変化点推測など議論した。特にコロナ感染率については強度関数の共変量に対して長期記憶構造を導入して強度の推測やコロナ感染者の予測など議論した。この場合、信頼区間も求めた。今後、実社会の問題として極めて重要であると思われる。

(2-5)

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、シミュ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。

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・准教授
劉 言 早稲田大学・数理科学研究所・講師
後藤 祐一 早稲田大学・理工学術院講師（任期付き）
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
Xue Yujie 早稲田大学・理工総研・次席研究員 基盤 (S) (谷口) による雇用

4. 研究業績

4.1 学術論文

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 Diseases*, 102, 327-331, Doi.org/10.1016/j.ijid.2020.10.050

Akashi, F., Taniguchi, M. and Tanida, Y. (2021). Estimation of linear functional of large spectral density matrix and application to Whittle's approach, To appear in Japanese Journal of Statistics and Data Science.

Nakajima, S. and Shimizu, Y. (2022). Asymptotic normality of least squares estimator to stochastic differential equations driven by fractional Brownian motions, arXiv:2112.12333; to appear in *Statistics and Probability Letters*.

Nakamura, S. Shimizu, Y. and Nalajima, S. (2022) Least squares estimators for discretely observed stochastic processes driven by small fractional noise, arXiv:2201.08462.

Nakajima, S. and Shimizu, Y. (2021). Parameter estimation of stochastic differential equation driven by small fractional noise, arXiv:2201.00372, 2021.

Kobayashi, M. and Shimizu, Y. (2021). Least squares estimators based on the Adams method for stochastic differential equations with small Lévy noise, arXiv:2201.06787.

Hayamizu, M., A Structure Theorem for Rooted Binary Phylogenetic Networks and Its Implications for Tree-Based Networks. *SIAM Journal on Discrete Mathematics* 35(4) 2490-2516, 2021年1月. <https://doi.org/10.1137/19M1297403>

Liu Y., Tanida Y. and Taniguchi, M. (2021). Shrinkage estimation for multivariate time series. *Statistical Inference for Stochastic Processes*. 24, 733-751.

Liu Y. and Taniguchi, M. (2021). Minimax estimation for time series models. *Metron*. 79, 353-359.

Xu, X., Li, Z., and Taniguchi, M. (2022). Comparison between the exact likelihood and Whittle likelihood for moving average processes, To appear in *STATISTICA*.

4.2 総説・著書

Akashi, F., Taniguchi, M., Monti, A.C. and Amano, T. (2021). *Diagnostic Methods in Time Series* Springer Briefs in Statistics, Springer-Verlag, 108 pages
Doi.org/10.1007/978-981-16-2264-9

Shimizu, Y. (2022). *Asymptotic Statistics in Insurance Risk Theory*, SpringerBriefs in Statistics, Springer, Singapore.

4.3 招待講演

Taniguchi, M. The long and curved road, Waseda International Symposium Topological Data Science, Causality, Analysis of Variance, & Time Series, 早稲田大学, 2022年3月8日

Taniguchi, M. The long and curved road II, Otsu Seminar 2022年3月11日

Shimizu, Y. (2021). M-estimation based on quasi-processes from discrete samples of Levy processes / Workshop on Statistical modeling for stochastic processes and related fields (online), September 27-30, 2021

Shimizu, Y. (2021). A quite new approach to cohort-wise mortality prediction under survival energy hypothesis / ARLES seminar@online, Ritsumeikan University, October 21th, 2021.

Shimizu, Y. (2022). A quite new approach to cohort-wise mortality prediction under survival energy hypothesis / Waseda International Symposium: Topological Data Science, Causality, Analysis of Variance, & Time Series, Waseda University (Nishiwaseda campus), March 8, 2022.

Momoko Hayamizu, The structure of rooted binary phylogenetic networks and its applications to various computational problems in evolutionary analysis, RIMS Workshop: Toward an integration of fluids, ecology, and evolution in Biofluids 2021, 2021年7月6日

早水桃子, シングルセル RNA-Seq データを用いた細胞分化の軌跡推定を支える数理手法とソフトウェアの紹介, 2021年日本バイオインフォマティクス学会年会・第10回生命医薬情報学連合大会 (IIBMP2021), 企画セッション JST さきがけ「数理構造活用」, 2021年9月27日

早水桃子, 進化の系統樹と系統ネットワークに関する組合せ論, 第17回 MACS コロキウム, 京都大学理学部, 2021年11月19日

早水桃子, 1細胞の遺伝子発現データを用いた細胞分化の軌跡推定, AI・データ利活用研究会 第25回, 大阪大学, 2021年12月10日

早水桃子, 進化の系統樹と系統ネットワークにまつわる組合せ論, ネットワーク科学研究会 2021, 2021年12月12日

Momoko Hayamizu, Treefit: measuring the tree-likeness of point clouds and application to the analysis of single-cell gene expression data, Waseda International Symposium Topological Data Science, Causality, Analysis of Variance, & Time Series, 早稲田大学, 2022年3月9日

Liu, Y., Tamura, Y. and Taniguchi, M. Sphericity test for high-dimensional time series. Otsu Seminar. March 11, 2022.

Xu, X., Liu, Y. and Taniguchi, M. Higher order asymptotics of minimax estimators for time series. Otsu Seminar. March 11, 2022.

Liu, Y., Taniguchi, M. and Ombao, H. Statistical and Topological Inference for Local Granger Causality. Waseda International Symposium. March 9, 2022.

Xu, X., Chen, Y., Liu, Y., Goto, Y. and Taniguchi, M. Long-memory log-linear zero-inflated generalized Poisson autoregression for COVID-19 pandemic modeling. Waseda International Symposium. March 9, 2022.

Fujimori, K., Goto, Y., Liu, Y. and Taniguchi, M. Sparse principal component analysis for high-dimensional stationary time series. Waseda International Symposium. March 7, 2022.

Goto, Y., Arakaki, K., Liu, Y. and Taniguchi, M. Homogeneity tests for one-way models with dependent errors under correlated groups. Waseda International Symposium. March 7, 2022.

Liu, Y. 予測誤差に基づく時系列の統計的推測とその応用. 統計関連学会連合大会. September 6, 2021.

Goto, Y. 2021 Nov. 18 “Hypothesis tests for one- and two-way models with dependent errors” , (with Arakaki, K., Suzuki, K, Liu, Y., Xu, X., and M. Taniguchi), Data Science Workshop, Tohoku University

Goto, Y. 2022 Mar. 10—12 “Tests for the existence of group effects and interactions for two-way models with dependent errors”, (with Suzuki. K., Xu, X., and M. Taniguchi), Otsu Seminar “Recent Developments in Time Series and Related Topics” In honor of Professor Masanobu Taniguchi on the occasion of his retirement, Biwako Hotel.

Goto, Y. 2022 Mar. 7–9 “Homogeneity tests for one-way models with dependent errors under correlated groups” , (with K. Arakaki, Y. Liu, and M. Taniguchi), Waseda International Symposium “Topo-logical Data Science, Causality, Analysis of Variance & Time Series” dedicated to Professor Taniguchi's retirement, Waseda University.

Xu, X. “Long-memory Log-linear Zero-inflated Generalized Poisson Autoregression for COVID-19 Pandemic Modeling”, (with Chen. Y., Liu, Y., Goto, Y., and Taniguchi, M.), Waseda International Symposium “Topo-logical Data Science, Causality, Analysis of Variance & Time Series” dedicated to Professor Taniguchi's retirement, Waseda University. 2022 Mar. 7—9.

Xu, X. “Higher order asymptotics of minimax estimators for time series”, Otsu Seminar “Recent Developments in Time Series and Related Topics” In honor of Professor Masanobu Taniguchi on

the occasion of his retirement, Biwako Hotel. 2022 Mar. 10—12.

Xue, Y. : A new estimator which improves MLE uniformly (with Taniguchi, M.), Otsu Seminar “Recent Developments in Time Series and Related Topics” In honor of Professor Masanobu Taniguchi on the occasion of his retirement, 琵琶湖ホテル, 2022年3月11日

Hellinger distance estimation for non-regular spectra (with Taniguchi, M.), Waseda International Symposium “Topo-logical Data Science, Causality, Analysis of Variance & Time Series” dedicated to Professor Taniguchi's retirement, 早稲田大学, 2022年3月9日

4.4 受賞・表彰

谷口正信 令和4年度 科学技術分野の文部科学大臣表彰科学技術賞 研究部門

Hayamizu, M. WASEDA e-Teaching Award, 2021年度春学期 離散数学入門, 2022年3月15日

劉言、第35回 小川研究奨励賞 日本統計学会

青山祐介, 各務文雄, 柳沼基, 塩見岳史, 劉言, 谷口正信. 感性工学活用による燃料電池研究の価値向上研究. 日本感性工学会大会. September 2, 2021. 優秀発表賞

4.5 学会および社会的活動

谷口正信：自動車産業との“数理統計学活用による共同研究”
研究代表者：谷口正信、

早稲田大学総合理工学研究所、名誉研究員

学術誌編集：

(i) Associate Editor, *Statistica Sinica* (国際誌) 2020 - 現在

(ii) Editorial Board, *Statistical Inference for Stochastic Processes* (仏誌), 2007 - 現在。

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

学術団体役職：

(i) Institute of Mathematical Statistics(USA), 1987—
米国数理統計学会特別会員(Fellow)

- (ii) International Statistical Society (国際統計協会) , 1987—
正会員(Elected Member)
- (iii) イタリア Bologna 大学の学位審査委員 2021--

清水泰隆： 日本アクチュアリー会 評議員
日本年金保険リスク学会誌 編集委員
Japanese Journal of Statistics and Data Science 編集委員
Journal of the Japanese Association of Risk, Insurance and Pensions Editor

早水桃子：

1. 研究教育アウトリーチ活動：YouTube チャンネル登録者 3 万人達成
<https://www.youtube.com/channel/UCZFe7-vmShA0oRQ55FomoUg>
2. 早稲田大学 大学総合研究センターWASEDA OpenCourseWare 広報用 PV (YouTube)
動画制作への協力
3. 2023 年度版 早稲田大学 学部入学案内パンフレット 基幹理工学部ページへの取材協力
4. JST/CRDS 俯瞰ワークショップ (WS) パネリスト・取材協力, 2021 年 10 月 28 日

後藤祐一：

- 2022 Mar. 28-31 “Tests for the existence of group effects and interactions for two-way models with dependent error” , (with Xu, X., Suzuki, K., and M. Taniguchi), Mathematical Society of Japan Spring Meeting 2022. Saitama Univ.
- 2021 Sep. 14-17 “Homogeneity tests for one-way models with dependent errors” , (with K. Arakaki, Y. Liu, and M. Taniguchi), Mathematical Society of Japan Autumn Meeting 2021. Online
- 2021 Sep. 5-9 “従属誤差を持つ一元配置モデルにおける均一性の検定” , (with K. Arakaki, Y. Liu, and M. Taniguchi), Japan Joint Statistical Meeting 2021. Online

Xu, X. : “Comparison between the exact likelihood and Whittle likelihood for moving average processes”, (with Li, Z. and M. Taniguchi), Mathematical Society of Japan Spring Meeting 2022. Saitama Univ. 2022 Mar. 28—31.

“Higher order asymptotics of minimax estimators for time series”, (with Liu, Y. and M. Taniguchi), Mathematical Society of Japan Autumn Meeting 2021. Online. 2021 Sep. 14-17.

“Higher order asymptotics of minimax estimators for time series”, (with Liu, Y. and M. Taniguchi), Japan Joint Statistical Meeting 2021. Online. 2021 Sep. 5—9.

5. 研究活動の課題と展望

本研究の設定は極めて一般的であり、複数のシグマ加法族の離反度から因果性を導入すれば、その適用範囲は通常の、金融、医学、環境データだけでなく、グラフ、遺伝子、トポロジカルな対象まで広汎に広がる。またこの離反度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未来の、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。またこのメカニズムを説明する統計数理理論の構築は、極めて独創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革新性を持つもので、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や保険数理に適用すれば、いままでなされなかった因果性・TDA・系統学、保険数理の4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や保険、年金数理への波及効果も生み出せるであろう。