2025 年度科学研究費シンポジウム報告書

複雑・高次元データの統計科学の新展開：深化と融合

日時：2025 年 9 月 21 日（日）午後 - 23 日（火）正午
場所：〒 980-8579 仙台市青葉区荒巻字青葉 6-3-09
東北大学青葉山キャンパス情報科学研究科棟 2 階大講義室
開催責任者：荒木由布子（東北大学）
開催協力者：Xin Guan（東北大学）
共催：東北大学大学院情報科学研究科

内容・目的

近年，観測技術の進展および情報基盤の高度化に伴い，時間的・空間的構造を持つデータ，高次元・関数型データ，非独立，非線形性を含むデータなど，データの複雑多様化が進んでいます．こうした環境において，統計科学およびデータ科学に基づく研究は，理論と応用の両側面で進展しつつあり，さらに多様な分野との融合を通じて新たな知的枠組みが形成されつつあります．本シンポジウムでは，複雑・高次元データの解析に関する統計的理論や手法の新展開，さらにはさまざまな科学分野への応用事例を含む研究成果について，広く講演を募集します．理論と応用の各視点からの学際的な議論を通じて，統計科学における新たな研究の展開と，分野間連携の促進を図ることを目的としています．複雑データを対象とする統計的アプローチの将来像を展望し，分野を超えた知見の交流を通じて，新たな知的基盤の形成に資する場となることを期待しています．

科学研究費補助金

- 基盤研究 (A) 25H01107 「大規模複雑データの理論と方法論の深化と展開」
研究代表者：青嶋誠（筑波大学）
- 基盤研究 (B) 23K28042 「新たな複雑系データのための関数データ解析手法の開発と適用」
研究代表者：荒木由布子（東北大学）

プログラム

9 月 21 日（日）

12:50 開場

13:20–13:30 オープニング

13:30–14:00 佐藤圭悟（東北大学） p.1-2
再生可能エネルギー発電予測に向けた深層学習による気象モデルの構築

14:00–14:30 鈴木俊太郎（大阪大学） p.3-4
深層学習を用いた関数型自己回帰モデルの推定

14:30–15:00 屋良淳朝（大阪大学） p.5-6
深層学習を用いた共変量付きポアソン過程に対するノンパラメトリック回帰

15:00–15:20 休憩

15:20–15:50 市澤孝弥（岩手大学） p.7-8
Preferential attachment ランダムグラフの次数分布漸近評価の
数値解析について

15:50–16:20 永井勇（中京大学） p.9-10
GMANOVA モデルにおける直接的罰則付最尤推定量と最適化等

16:20–17:00 今野良彦（大阪公立大学） p.11-12
Shrinkage estimators in multivariate normal distributions
with block compound symmetry covariance structures

9 月 22 日（月）

10:00–10:30 竹内努（名古屋大学） p.13-14
Survival Analysis in Astrophysics: from Univariate to Multivariate
（宇宙物理学における生存時間解析: 1 変数から多変数へ）

10:30–11:00 島谷健一郎（統計数理研究所） p.15-16
クローナル植物に関する地上データと地下情報の紐づけ統合と統計モデル

11:00–11:30 藤原直哉（東北大学） p.17-18
Mobility analyses of geospatial agents

11:30–13:00 昼休み

13:00–13:30 河村和徳（拓殖大学） p.19-20
政治学研究における統計分析の応用動向

13:30–14:00 岡田陽介（拓殖大学） p.21-22
選挙研究における非言語情報の測定と分析：政治家の声の高低と印象形成

14:00–14:30 吐合大祐（高知県立大学） p.23-24
テキスト分析から見える自民党県連の組織構造

14:30–15:00 Xinhe Li（小樽商科大学） p.25-26
感情分析を用いた議員態度の可視化：北海道議会におけるオリンピック
議論の事例

15:00–15:20　休憩

15:20–15:50　Lingyu Jiang（東北大学）　p.27-28
Harnessing Kolmogorov–Arnold Networks for Accurate and Efficient Time Series Forecasting

15:50–16:20　新川裕也（佐賀大学）　p.29-30
クロスオーバー試験における持続血糖データの階層ガウス混合モデリング

16:20–16:50　岡野遼（一橋大学）　p.31-32
関数合成コントロール法とその拡大

16:50–17:00　休憩

17:00–17:40　Jeng-Min Chiou（National Taiwan University）　p.33-34
Modeling Structural Changes in Distributional Data

18:30–20:30　懇談会

9 月 23 日（火）

10:00–10:30　Sai Yao（東北大学）　p.35-36
Nonlinear Trivariate Modal Interval Regression and Its Application to Spatial Precipitation Data

10:30–11:00　Atina Husnaqilati（Universitas Gadjah Mada, Indonesia）　p.37-38
Understanding Excess Return –An Investigation into the Applicability of CAPM, Fama-French, PCA, and Random Matrix Theory in Indonesia

11:00–11:40　Kengo Kamatani（統計数理研究所）　p.39-40
Practical Monte Carlo Methods Using Piecewise-Deterministic Markov Processes

11:40–11:50　クロージング

再生可能エネルギー発電予測に向けた深層学習による気象モデルの構築

東北大学大学院経済学研究科　佐藤 圭悟　松田 安昌　李 銀星

はじめに

電力系統の安定供給には、需要と供給の一致が不可欠である。近年は脱炭素化の流れを背景に再生可能エネルギーの導入が進み、その割合は増加している。しかし、再生可能エネルギーは季節や天候により発電量が大きく変動するため、それのみで安定的にエネルギーを賄うことは困難である。したがって、火力発電など出力調整が可能な電源が依然として必要となるが、その調整には追加的なコストが伴う。この課題を解決するためには、再生可能エネルギー発電量を高精度に予測することが不可欠であり、その基盤として正確な気象予測が求められる。近年、気象予測は従来の数値予報モデルに加え、深層学習が注目されており、すでに数値予報モデルを上回る精度を示す成果も報告されている。本研究では、深層学習を用いたデータドリブン型気象予測モデルの構築を目標とする。

提案手法

本研究では、日本周辺を対象とした風速予測において、再解析データと数値予報モデルの予測値を入力とする sequence-to-sequence を用いた Convolutional LSTM（ConvLSTM）による予測手法を提案する。図 1 にモデルの概要を示す。ConvLSTM は Shi et al. (2015) により提案された手法であり、LSTM の行列演算を畳み込み演算に置き換えることで、画像データの位置情報を保持したまま時系列処理を行うことが可能である。本研究では、直近 24 時間の再解析データを Encoder に入力し、予測対象である未来 18 時間の欧州中期予報センターの数値予報モデル IFS に基づく ERA5 短期予報を Decoder に入力する。この構成により、再解析データを活用して数値予報モデルの予測を補完することができる。また、本モデルは差分予測を採用しており、IFS の予測値に本モデルが推定した差分を加算することで最終的な風速予測を得る。

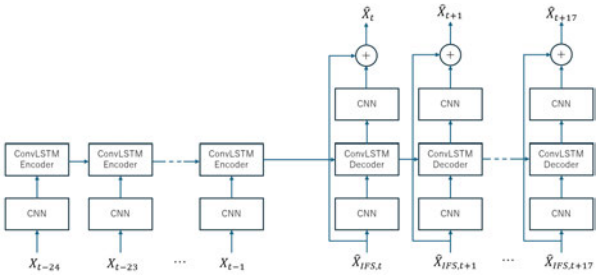


図 1: 提案モデルの概要

実験方法

実験には欧州中期予報センターが提供する再解析データ ERA5 を用いた。再解析データとは、観測データと数値予報モデルをデータ同化手法により統合し、過去の大気の状態を再現したデータである。ERA5 は全球を対象とし、1 時間間隔・0.25 度解像度の格子データを提供する。本研究では、ERA5 に含まれる地上 10 m の東西風（u10）および南北風（v10）を変数として用い、日本を中心とする 64×64 グリッドを対象領域とした。学習期間は 2022 年 7 月から 2024 年 6 月、検証期間は 2024 年 7 月から 8 月、テスト期間は 2024 年 9 月から 12 月と設定した。また、比較対象としては、IFS に基づく ERA5 の短期予報を用いた。

結果および考察

テストデータに対する予測を平均二乗誤差（MSE）で評価した結果を表 1 および図 2 に示す。短時間（1–3 時間先）では IFS の誤差が本手法よりも小さく、数値予報モデルの予測精度が高いことが確認された。一方、4 時間先以降では本手法の誤差が IFS を下回り、18 時間先まで持続的に低い値を示した。したがって、本手法は中期的な時間スケールにおける予測性能向上に寄与することが明らかとなった。

この改善は、本手法が直近の再解析データを活用することで IFS の予測を効果的に補完していることに起因すると考えられる。一方、短時間（1–3 時間先）では数値予報モデルの予測が優位であるため、提案手法の補完効果が顕著に現れるまでには一定のリードタイムが必要であることが示唆される。

表 1: MSE

モデル	1h	2h	3h	4h	5h	6h
IFS	0.0307	0.0611	0.0906	0.9129	0.9781	1.0415
IFS+ConvLSTM	0.0322	0.0630	0.0988	0.7402	0.8157	0.8824

モデル	7h	8h	9h	10h	11h	12h
IFS	1.0902	1.1383	1.1834	1.2188	1.2439	1.2687
IFS+ConvLSTM	0.9355	0.9833	1.0254	1.0606	1.0899	1.1188

モデル	13h	14h	15h	16h	17h	18h
IFS	1.2906	1.3139	1.3358	1.9586	1.9792	2.0035
IFS+ConvLSTM	1.1481	1.1860	1.2304	1.6506	1.6490	1.6731

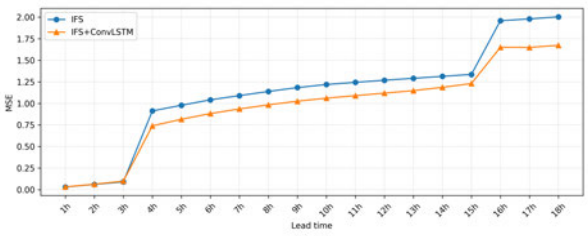


図 2: MSE 推移

参考文献

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c. (2015). Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28.

深層学習を用いた関数型自己回帰モデルの推定

大阪大学大学院基礎工学研究科 鈴木譲研究室 博士後期課程 1 年 鈴木 俊太郎

標本が関数であるデータを解析対象とする関数データ解析と呼ばれる手法が近年盛んに研究されている．その中でも個々のデータが関数として記述される時系列データを関数型時系列と呼ぶ．関数型時系列はエネルギー・人口統計・金融・環境などで様々な現象のモデリングに応用される．例を挙げると Hörmann, Kokoszka, and Nisol [6], Hyndman and Ullah [8], Hyndman and Shang [7] 等があげられる．本研究では AR モデルの拡張である以下の関数型自己回帰モデル (FAR) を考察対象とする:

$$X_{t+1}(\boldsymbol{u}) = \Psi_0\left(X_t(\cdot)\right)(\boldsymbol{u}) + \xi_t(\boldsymbol{u}), \quad \boldsymbol{u} \in [0,1]^d, \quad t \in \mathbb{Z}_{\geq 0}, \tag{0.1}$$

ただし Ψ_0 は $\mathcal{H} := L^2([0,1]^d : \mathbb{R})$ から $\mathcal{H}$ への非線形作用素であるとする．$\{\xi_t\}_{t\in\mathbb{N}}$ は独立同分布な $\mathcal{H}$ 上の確率要素で平均関数は $\mathbf{0}_{\mathcal{H}}$ で与えられる．また各 $t \in \mathbb{N}$ に対して ξ_t は $X_{t-1}, ..., X_0$ と独立である．

(0.1) において Ψ_0 が線形作用素である場合は解釈性が高いため様々な現象のモデリングに応用される．例を挙げると Shah, Khan, Javed, et al. [12], Chen, Kokoszka, et al. [3], が挙げられる．Ψ_0 の推定手法としては Bosq [1] が提案した Yule - Walker 法や Cardot, Ferraty, and Sarda [2] が提案した B-Spline による近似方法が提案されている．これらの研究では固有値の減衰や有限次元空間への射影の打ち切り誤差に対して適切な正則条件を加えたもとで提案推定量の収束レートを導出している．

このように FAR モデルは推定手法が盛んに研究されており実データ解析への応用例も豊富なモデルである．一方で時間方向についてのデータの変動については線形な表現力しか持たない．そこで本研究では Ψ_0 を線形作用素から非線形な作用素へ拡張した非線形なモデルについて考える．Ψ_0 が非線形な場合について扱った論文としては Zhu and Politis [13] Masry [10], Kurisu [9] があげられる．

これらの文献は推定量の漸近挙動を導出するために関数型時系列 $\{X_t\}_{t\in\mathbb{Z}_{\geq 0}}$ に指数混合性と呼ばれる性質を満たすことを課している．指数混合性は (0.1) のように時系列データの構造をノンパラメトリックに推定する際に推定量の漸近挙動を導出する上で重要な性質である．そのため時系列データが指数混合性をもつ条件については盛んに研究されている．例を挙げると Chen and Chen [4], Meyn and Tweedie [11], Hairer and Mattingly [5] などの研究が代表的である．一方で無限次元空間の測度は特異になりやすく $\{X_t\}_{t\in\mathbb{N}}$ のような確率過程が指数混合性をもつ条件は自明ではない．しかし Zhu and Politis [13] Masry [10], Kurisu [9] では関数型時系列 $\{X_t\}_{t\in\mathbb{N}}$ が指数混合条件を満たすための具体的な条件について言及されていない．そこで本研究では (0.1) にしたがう関数型時系列が指数混合性をもつための十分条件を与えた．また (0.1) が指数混合性を持つもとで Ψ_0 の推定量全般に対して分散評価を与え, 推定量の一例として深層学習を用いた際の汎化誤差を与えた．

References

[1] Denis Bosq. *Linear Processes in Function Spaces: Theory and Applications*. New York: Springer, 2000. ISBN: 9780387987071. URL: https://d-nb.info/959841423/04.

[2] Hervé Cardot, Frédéric Ferraty, and Pascal Sarda. "Spline estimators for the functional linear model". In: *Statistica Sinica* 13.3 (2003), pp. 571–591.

[3] Dazhi Chen, Piotr Kokoszka, et al. "A review study of functional autoregressive models with application to energy forecasting". In: *Energy Systems* 11.2 (2020), pp. 209–241. DOI: 10.1007/s12667-019-00355-y. URL: https://www.researchgate.net/publication/343286956_A_review_study_of_functional_autoregressive_models_with_application_to_energy_forecasting.

[4] M Chen and G Chen. "Geometric ergodicity of nonlinear autoregressive models with changing conditional variances". In: *Canadian Journal of Statistics* 28.3 (2000), pp. 605–614. DOI: 10.2307/3315968.

[5] Martin Hairer and Jonathan C. Mattingly. "Yet another look at Harris' ergodic theorem for Markov chains". In: *Seminar on Stochastic Analysis, Random Fields and Applications VI*. Vol. 63. Progress in Probability. Birkhäuser / Springer Basel AG, 2011, pp. 109–117. DOI: 10.1007/978-3-0348-0021-1_7.

[6] Siegfried Hörmann, Piotr Kokoszka, and Rainer Nisol. "Functional dynamic factor models with application to yield curve forecasting". In: *Journal of Econometrics* 157.2 (2010), pp. 294–303. DOI: 10.1016/j.jeconom.2010.03.012.

[7] Rob J. Hyndman and Han Lin Shang. "Grouped functional time series forecasting: An application to age-specific mortality rates". In: *Journal of Computational and Graphical Statistics* 25.4 (2016), pp. 1189–1208. DOI: 10.1080/10618600.2015.1027774.

[8] Rob J. Hyndman and Md Shahid Ullah. "Robust forecasting of mortality and fertility rates: A functional data approach". In: *Computational Statistics & Data Analysis* 51.10 (2007), pp. 4942–4956. DOI: 10.1016/j.csda.2006.07.028.

[9] Daisuke Kurisu. "Nonparametric regression for locally stationary random fields under stochastic sampling design". In: *Electronic Journal of Statistics* 14.1 (2020), pp. 2032–2061. DOI: 10.1214/20-EJS1719. URL: https://doi.org/10.1214/20-EJS1719.

[10] Elias Masry. "Nonparametric regression estimation for dependent functional data: asymptotic normality". In: *Stochastic Processes and their Applications* 115.1 (2005), pp. 155–177.

[11] Sean P. Meyn and Richard L. Tweedie. *Markov Chains and Stochastic Stability*. 2nd. Cambridge: Cambridge University Press, 2009. DOI: 10.1017/CBO9780511626630.

[12] Muhammad Bilal Shah, Muhammad Nauman Khan, Muhammad Javed, et al. "A functional autoregressive approach for modeling and forecasting short-term air temperature". In: *Frontiers in Environmental Science* 12 (2024), p. 1411237. DOI: 10.3389/fenvs.2024.1411237. URL: https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2024.1411237.

[13] Tingyi Zhu and Dimitris Politis. "Kernel estimates of nonparametric functional autoregression models and their bootstrap approximation". In: *Electronic Journal of Statistics* 11 (2017), pp. 2876–2906.

深層学習を用いた共変量付きポアソン過程に対するノンパラメトリック回帰

大阪大学 大学院基礎工学研究科 屋良 淳朝
大阪大学 大学院基礎工学研究科，理研 AIP 寺田 吉壱

概 要

本研究では，共変量を伴うポアソン過程の強度関数に対するノンパラメトリック推定問題を考える．ノンパラメトリック推定においてはカルバック・ライブラー（KL）ダイバージェンスが容易に発散するため，ノンパラメトリック最尤推定量（NPMLE）の理論解析は困難である．そこで本研究では，NPMLE の解析に対し単純かつ統一的な枠組みを導入し，真の強度関数と NPMLE とのヘリンジャー距離を評価するオラクル不等式を導出する．さらに，その応用として深層ニューラルネットワークを用いた NPMLE の収束レートを明らかにする．我々の解析によって，真の強度関数が合成関数の構造を有する場合に，深層学習による NPMLE が次元の呪いを克服し得ること，また低次元リーマン多様体構造に自動的に適応可能であることが示された．加えて，真の強度関数が合成関数の構造を有する場合には，導出した収束レートがほぼミニマックス最適な速度を達成することを証明した．

はじめに

ポアソン（点）過程は，時間と空間におけるランダムな事象の基本的な確率モデルであり，疫学，生態学，地質学，地震学，都市情報学などの分野で幅広く応用されている．ポアソン過程（とその一般化）は例えば，地震の発生，環境暴露の関数としての疾病患者の空間分布，ある都市におけるタクシーの乗車要求の頻度などをモデル化するのに利用されてきた．ポアソン過程は，時間または空間の任意の時点における事象の平均発生率を表す強度関数によって特徴付けられる．イベントデータからこの強度関数を推定することは，背後のイベント発生メカニズムについての洞察を与え，将来の事象を予測することを可能にするため，重要な問題である．

本研究では，同一の条件下で得られた複数の独立な共変量によって駆動されるポアソン過程の実現を観測できる場合を考える．すなわち，共変量 Z と，それを条件付けた下で定まる条件付き強度関数 $S_0(\cdot \mid Z)$ で決定されるポアソン過程からの i.i.d. 観測 $(Z_1, \{x_{1,1}, \ldots, x_{1,N_1}\}), \cdots, (Z_n, \{x_{n,1}, \ldots, x_{n,N_n}\})$ から真の条件付き強度関数 $S_0(\cdot \mid \cdot)$ を推定する問題を考える．

強度関数のノンパラメトリック推定についての理論研究は，カーネル平滑化など推定量が陽な表現を持つ場合を中心に行われている．しかし, 実際の応用で強力な予測性能を発揮する深層学習を用いた強度関数の推定のように，推定量が陽な表現を持たない場合については理論解析はまだ限定的である．特に，ノンパラメトリック最尤推定においては，通常の汎化誤差解析では, 真のモデルと推定モデルのカルバック・ライブラー (KL) 情報量の収束を議論するが, これは成立しない場合がある．そのため, 真の強度関数に仮定を課さない限り理論解析ができない. 一方で, 実世界での応用を考えると，KL 情報量の意味での収束は必ずしも必要ではなく，適切な metric での一致性を示せば十分である．そこで，本研究では，van de Geer (2000) のテクニックを応用して，KL 情報量の代わりに強度関数間のヘリンジャー距離を評価することで，現実的な仮定の下で強度関数のノンパラメトリック最尤推定量の一致性を示すための, オラクル不等式を導出する．オラクル不等式の重要な応用として，(i) 真の強度関数が合成関数の構造を持つ場合，(ii) 真の強度関数がリーマン多様体上に台をもつ場合において，深層学習による強度関数の最尤推定量の収束レートを導出する．これによって，深層学習は，(i) の場合には次元の呪いを回避することができ，(ii) の場合には明示的にリーマン計量などの情報を与えなくても自動的にリーマン多様体の構造に適応できることが明らかとなった．また，(i) の場合において，深層学習がほとんどミニマックス最適な収束レートを達成することを示した．

主結果

まずは，強度関数のノンパラメトリック最尤推定を定式化する．コンパクト集合 $A \subset \mathbb{R}^d$ 上のポアソン過程を

考え，与えられた条件付き強度関数 S に対して，観測 $(X_i, \boldsymbol{z}_i)$ の負の対数尤度関数を

$$\ell(X_i, \boldsymbol{z}_i \mid S) \coloneqq -\int_A \log(S(\boldsymbol{x} \mid \boldsymbol{z}_i))\, X_i(d\boldsymbol{x} \mid \boldsymbol{z}_i)) + \int_A S(\boldsymbol{x} \mid \boldsymbol{z}_i)\, d\boldsymbol{x} = -\sum_{j=1}^{N_i} \log(S(\boldsymbol{x}_{ij} \mid \boldsymbol{z}_i)) + \int_A S(\boldsymbol{x} \mid \boldsymbol{z}_i)\, d\nu(\boldsymbol{x})$$

で定める．ここで，$\{\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{iN_i}\} \subset A$ は共変量 $\boldsymbol{z}_i$ を与えた下での X_i に対応する観測点であり，$N_i = X_i(A \mid \boldsymbol{z}_i)$ である．以上より，共変量 Z とそれを与えた下でのポアソン過程 $X(\cdot \mid Z)$ からの i.i.d. 観測 $(X_1, \boldsymbol{z}_1), \ldots, (X_n, \boldsymbol{z}_n)$ に対する負の対数尤度関数（の $1/n$ 倍）は，

$$\ell_n(X_1, \ldots, X_n, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_n \mid S) \coloneqq \frac{1}{n}\sum_{i=1}^{n} \ell(X_i, \boldsymbol{z}_i \mid S) = \frac{1}{n}\sum_{i=1}^{n} \left\{ -\sum_{j=1}^{N_i} \log(S(\boldsymbol{x}_{ij} \mid \boldsymbol{z}_i)) + \int_A S(\boldsymbol{x} \mid \boldsymbol{z}_i)\, d\nu(\boldsymbol{x}) \right\} \tag{1}$$

となる．今後，誤解がなければ観測データを省略して $\ell_n(S) \coloneqq \ell_n(X_1, \ldots, X_n, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_n \mid S)$ とも書く．ノンパラメトリック最尤推定では，必ずしも有限次元のパラメータでパラメトライズされない関数の空間（仮説空間）$\mathcal{F}_n$ 上で負の対数尤度 (1) を最小化する関数を推定量とする．得られた推定量は，ノンパラメトリック最尤推定量 (Nonparametric maximum likelihood estimator, NPMLE) と呼ばれる．すなわち，共変量 $\boldsymbol{z}$ を与えた下でのポアソン過程 $X(\cdot \mid \boldsymbol{z})$ の条件付き強度関数 $S_0(\boldsymbol{x} \mid \boldsymbol{z})$ の NPMLE $\hat{S}_n(\boldsymbol{x} \mid \boldsymbol{z})$ は，

$$\hat{S}_n \in \min_{S \in \mathcal{F}_n} \ell_n(S) \tag{2}$$

で定義される．

この設定で，以下のオラクル不等式は真の強度関数と NPMLE とのヘリンジャー距離 $\widetilde{H}$ に関する評価を与える．

定理 1 (オラクル不等式)**.** ポアソン過程 $X(\cdot \mid Z)$ の条件付き強度関数 $S_0(\cdot \mid Z)$ の推定を考える．適当な条件の下で，普遍定数 c と

$$\sqrt{n}\delta_n^2 \geq c\Psi(\delta_n) \tag{3}$$

なる任意の $\delta \geq \delta_n$ に対して，

$$\mathbb{P}(\widetilde{H}^2(\hat{S}_n, S_0) > 1224(c_0+1)^2(\delta^2 + \widetilde{H}^2(S_n^*, S_0))) \leq c\exp\left(-\frac{n\delta^2}{c^2}\right) \tag{4}$$

が成り立つ．

このオラクル不等式によって，様々な真の強度関数が属する関数クラスと推定モデルの設定のもとで NPMLE の収束レートを導出することができる．例えば，推定モデルとして適当な大きさの DNN を用いた場合，S_0 が合成関数の構造を持つ場合（Schmidt-Hieber (2020) を見よ）は

$$\mathbb{E}[\widetilde{H}^2(\hat{S}_n, S_0)] \leq C'\phi_n \log(n)^3,$$
$$\phi_n \coloneqq \max_{i=0,\ldots,q} n^{-\frac{(1+\alpha\wedge 1)\beta_i^*}{(1+\alpha\wedge 1)\beta_i^* + t_i}},\ \beta_i^* \coloneqq \beta_i \prod_{l=i+1}^{q} (\beta_l \wedge 1)$$

が成立する．この収束レートは入力の次元に依存しておらず，次元の呪いを回避していることがわかる．さらに，この収束レートはほとんどミニマックス最適であることもわかる．すなわち，深層学習はこの設定においてこれ以上改善できない最適な性能を発揮する．

紙面の都合上省略するが，当日は低次元多様体上に台をもつポアソン過程に対する収束レートについても紹介する．

参考文献

Johannes Schmidt-Hieber. Nonparametric regression using deep neural networks with ReLU activation function. *Annals of Statistics*, 48:1875 – 1897, 2020.

Sara van de Geer. *Empirical Processes in M-estimation*, volume 6. Cambridge university press, 2000.

Preferential attachment ランダムグラフの次数分布漸近評価の数値解析について

岩手大学総合科学研究科理工学専攻　市澤孝弥　川崎秀二

1　序論

複雑ネットワークをモデル化する構成法の一つに preferential attachment (PA) がある。PA では次の二種のステップにより、パラメータ $p \in [0,1]$ を持つグラフ $G(p)$ を構成する：

- Vertex-step: 新しいノード v を追加する。既存のノードから一つのノード u を選んで、v とエッジで結ぶ。u が選ばれる確率は u の次数に比例する。
- Edge-step: 既存のグラフからノード u と v を選び、これらの間にエッジ $\{u,v\}$ を追加する。u と v が選ばれる確率はそれぞれ u と v の次数に比例する。

PA では、ノードが 1 個で自己ループを持つグラフ G_0 を起点として、各時刻で確率 p で vertex-step、それ以外で edge-step を行うことでグラフを成長させる。時刻 t で次数が k のノード数を $m_{k,t}$ とする。次の定理が成り立つ。

定理 1.1 [1]　PA $G(p)$ において、$t \to \infty$ のとき $m_{k,t}$ は確率 1 で

$$m_{k,t} = M_k t + O\left(2\sqrt{k^3 t \log t}\right)$$

となる。ここで M_k は次で与えられる：

$$M_1 = \frac{2p}{4-p}, \qquad M_k = \frac{2p}{4-p}\frac{\Gamma(k)\Gamma\big(1+2/(2-p)\big)}{\Gamma\big(k+1+2/(2-p)\big)} \quad (k \geqq 2)$$

ただし $\Gamma(x)$ はガンマ関数 $\Gamma(x) = \displaystyle\int_0^\infty e^{-t}t^{x-1}\,dt, x>0$ である。

定理 1.1 は期待値との乖離度合いの上界の一つのオーダーを与えている。本講演では、この評価式のより精密な評価を与えることを目標に、定理 1.1 の t や k に関する評価を中心に述べる。

2　定理 1.1 の精密化

定理 1.1 で与えられる誤差の評価を改善するために、まず k を固定したときの $m_{k,t}$ の時間変化を追い、その最大値と最小値をシミュレートすることで、t に関する評価 $\sqrt{t\log t}$ が妥当であるかを見る。

$k=2$ を例にとる。まず、$p=0.5, t=10^6$ までの PA グラフを $n=9500$ サンプルとる。そして、500 サンプルごとの $m_{1,t}-M_1 t$ の最大値と最小値の平均を求めた。この値を $\sqrt{t\log t}$ や、これよりも増加速度の遅い関数の例である $\sqrt{t\log\log t}$ で割ったものの最大値・最小値を見ることで、最大値・最小値が $\sqrt{t\log t}$ のオーダーであるか、または $\sqrt{t\log\log t}$ のオーダーであるかを判別することを試みる。結果、左側の $(m_{k,t}-M_k t)/\sqrt{t\log t}$ よりも右側の $(m_{k,t}-M_k t)/\sqrt{t\log\log t}$ が一定値に近づいているように見える。

次に、t を固定したときの k に対する評価を見る。まず、$t=10000$、サンプル数 $n=20000$ とした PA ランダムグラフを作成し、各 k についての $m_{k,t}-M_k t$ の最大値・最小値を計算した。そして、20000 サンプルを 500 サンプル 40 セットに分け、500 サンプルごとの最大値と最小値の平均をとった（図 2）。この図は、サンプル数を 500 としたときに平均してどれくらいの最大（最小）値をとるかを示

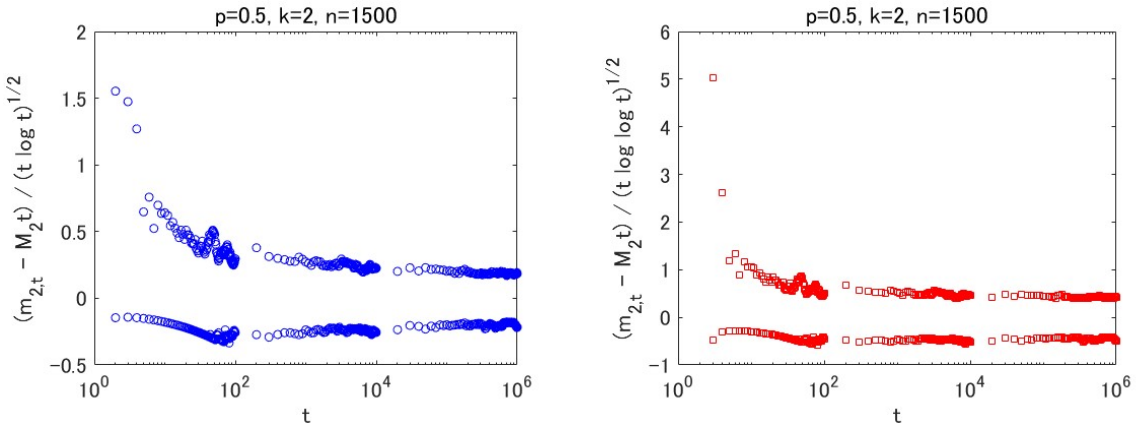


図 1　$(m_{k,t} - M_k t)/\sqrt{t \log t}$（左）と $(m_{k,t} - M_k t)/\sqrt{t \log\log t}$（右）の 100 サンプルごとの最大値と最小値の平均（$p = 0.5, k = 2, n = 4000$）

している。両対数プロットで直線傾向が表れていることから、k はべき乗型減衰をすると分かる。これは定理 1.1 の理論的上限・下限が k に関して増加関数であったこととは対照的である。

また、これらのプロットについて $c_t k^{a_t}$ の形のフィッティングも行った。フィッティング範囲は線形から外れる小さな k は除外し、最大値では $5 \leqq k \leqq 50$、最小値では $5 \leqq k \leqq 15$ とした。フィッティング結果は

$$\text{最大値：} c_t = 176.0665,\ a_t = -1.0062,$$
$$\text{最小値：} c_t = 249.5350,\ a_t = -1.2105$$

であった。最大値と最小値で指数 a_t が異なることから、このシミュレーションからは最大と最小では k に関する減衰の速さが異なるという結果になった。

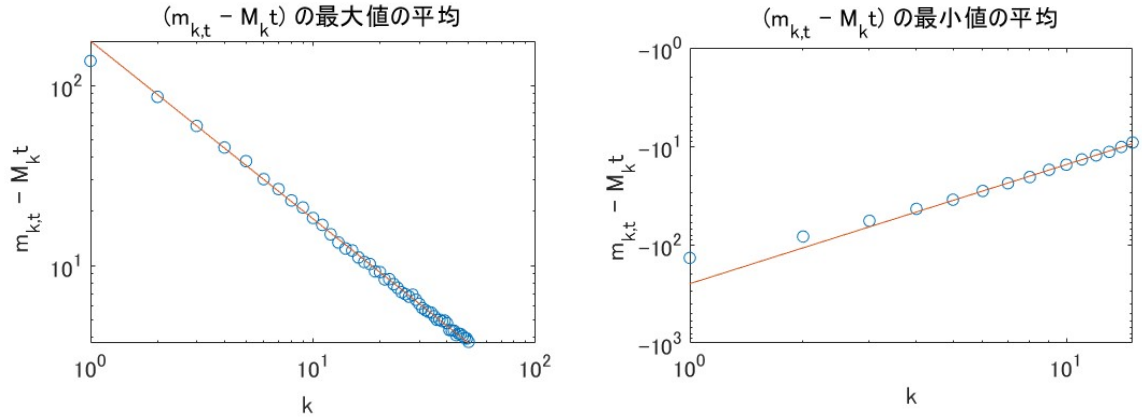


図 2　各 k に対する $m_{k,t} - M_k t$ の最大値・最小値（$p = 0.5$、$t = 10000$、$n = 10000$）

参考文献

[1] F.Chung and L.Lu, *Complex Graphs and Networks*, American Mathematical Society, 2006.

[2] 池田信行・小倉幸雄・高橋陽一郎・眞鍋昭治郎、確率論入門 I、培風館、2006.

[3] William F. Stout, *The Hartman-Wintner law of the iterated logarithm for martingales*, The Annals of Mathematical Statistics 1970, Vol.41, No.6, 2158-2160.

GMANOVA モデルにおける直接的罰則付最尤推定量と最適化等 (報告書)

永井 勇[1]

[1] 中京大学 教養教育研究院
(一部は講演スライドのみに記載)

2025 年 9 月 21 日

複雑・高次元データの統計科学の新展開：深化と融合

本研究の概要

対象 経時測定データ; 各個体で時間と共に測定したデータ
目的 データに潜む時間による変動を捉える
前提 全ての個体の測定時点が共通
→ GMANOVA モデルが分析によく使われている

従来の推定手法で起きる問題

(1) 各個体の (測定時点と無関係な) 説明変数の間に相関の高い組がある場合, 推定量が不安定になる
(2) 時間による変動を捉えるために用いる関数が柔軟すぎる場合, 目的の時間による変動ではなく分析に用いるデータに過剰に適合する

提案手法; (1) と (2) を回避する手法

罰則付推定量と罰則パラメータの最適化のための C_p 型情報量規準の構築とその先のアイデア

目次

ここからの話

モデルと仮定

対象 経時測定データ ; n 個体で時点と共に測定したデータ
目的 データに潜む時間による変動 (経時変動) を捉える
前提 全個体で測定時点は共通 (p 回測定)
→ GMANOVA モデル (Potthoff & Roy, 1964) で分析される;

$$\boldsymbol{Y} = \mathbf{1}_n \boldsymbol{\mu}' \boldsymbol{X}' + \boldsymbol{A}\boldsymbol{\Xi}\boldsymbol{X}' + \boldsymbol{\mathcal{E}},$$

- $\boldsymbol{Y}$; 各行が各個体の経時測定データからなる $n \times p$ 行列
- $\mathbf{1}_n$; n 次元の 1 ベクトル, $\mathbf{0}_h$; h 次元ゼロベクトル
- $\boldsymbol{X}$; 経時変動を捉えるために使う $p \times q$ 行列 (解析者が決める)
- $\boldsymbol{A}$; $\boldsymbol{A}'\mathbf{1}_n = \mathbf{0}_k$ (各列で中心化後) の説明変数を表す $n \times k$ 行列
- $\boldsymbol{\mu}$, $\boldsymbol{\Xi}$; q 次元未知ベクトル, $k \times q$ 未知行列
- $\boldsymbol{\mathcal{E}}$; $\boldsymbol{\mathcal{E}} \sim \mathcal{N}_{n,p}(\mathbf{0}_n\mathbf{0}_p', \boldsymbol{\Sigma} \otimes \boldsymbol{I}_n)$ (仮定) の $n \times p$ 誤差行列, $\boldsymbol{\Sigma}$; $p \times p$ 未知正定値行列

→ $\boldsymbol{Y}$, $\boldsymbol{A}$, $\boldsymbol{X}$ などから $\boldsymbol{\mu}$, $\boldsymbol{\Xi}$ (必要なら $\boldsymbol{\Sigma}$) を推定
仮定 $\mathrm{rank}(\boldsymbol{A}) = k$, $\mathrm{rank}(\boldsymbol{X}) = q$, $n - k - p - 2 > 0$.

$\boldsymbol{\mu}$, $\boldsymbol{\Xi}$ の推定と経時変動の推定について

例 $\boldsymbol{X} = \begin{pmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{q-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_p & t_p^2 & \cdots & t_p^{q-1} \end{pmatrix}$ とする ($\mathrm{rank}(\boldsymbol{X}) = q$ の範囲内で)
($t_1, \ldots, t_p$ は経時測定データの測定時点 ($t_1 < \cdots < t_p$))
→ この $\boldsymbol{X}$ をモデル ($\boldsymbol{Y} = \mathbf{1}_n\boldsymbol{\mu}'\boldsymbol{X}' + \boldsymbol{A}\boldsymbol{\Xi}\boldsymbol{X}' + \boldsymbol{\mathcal{E}}$) で用いると,
$\boldsymbol{\mu}$; 全個体で共通の多項式の各次数の係数と切片
$\boldsymbol{\Xi}$; 各説明変数に対応した多項式の各次数の係数と切片
→ $\boldsymbol{\mu}$, $\boldsymbol{\Xi}$ の推定 $\Leftrightarrow$ 多項式の各次数の係数と切片を推定
$\Leftrightarrow$ $\boldsymbol{Y}$ の経時変動 ($= E[\boldsymbol{Y}]$) を ($\boldsymbol{X}$ の形で) 推定

- $\boldsymbol{X}$ の中の関数を変える (例; $\boldsymbol{X}$ の (i,j) 成分 $= (\sin(t_i))^j$)

$\Leftrightarrow$ 使う関数の重み付き和で $\boldsymbol{Y}$ の経時変動の推定する形
$\boldsymbol{\mu}$, $\boldsymbol{\Xi}$ はそれぞれの関数への重みに対応

よく使われる推定量と問題

- $\boldsymbol{Y} = \mathbf{1}_n\boldsymbol{\mu}'\boldsymbol{X}' + \boldsymbol{A}\boldsymbol{\Xi}\boldsymbol{X}' + \boldsymbol{\mathcal{E}}$ のモデルで $\boldsymbol{\mu}$ と $\boldsymbol{\Xi}$ の推定

→ $\boldsymbol{\mathcal{E}} \sim \mathcal{N}_{n,p}(\mathbf{0}_n\mathbf{0}_p', \boldsymbol{\Sigma} \otimes \boldsymbol{I}_n)$ の下での MLE (最尤推定量);

$$\begin{cases} \hat{\boldsymbol{\mu}} = (\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{Y}'\mathbf{1}_n/n, \\ \hat{\boldsymbol{\Xi}} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{Y}\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}, \end{cases}$$

ここで,
$\boldsymbol{S} = \boldsymbol{Y}'\{\boldsymbol{I}_n - \mathbf{1}_n\mathbf{1}_n'/n - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\}\boldsymbol{Y}/(n-k-1)$.

- $\mathrm{rank}(\boldsymbol{S}) = p$ を追加で仮定
- $E[\boldsymbol{S}] = \boldsymbol{\Sigma}$ ($\boldsymbol{S}$ は $\boldsymbol{\Sigma}$ の不偏推定量)

これらの推定量の問題

(1) $\boldsymbol{A}$ の列の間 (説明変数の組) に相関の高い組がある場合, $\boldsymbol{\Xi}$ の推定量 ($\hat{\boldsymbol{\Xi}}$) が不安定になる (多重共線性)
(2) $\boldsymbol{X}$ に使う関数が柔軟すぎる場合, $\boldsymbol{Y}$ に過剰に適合する $\Leftrightarrow$ 経時変動を上手く捉えられない (過剰適合)

本研究の目的とアイデア

- $\boldsymbol{Y} = \mathbf{1}_n\boldsymbol{\mu}'\boldsymbol{X}' + \boldsymbol{A}\boldsymbol{\Xi}\boldsymbol{X}' + \boldsymbol{\mathcal{E}}$ のモデルで $\boldsymbol{\mu}$ と $\boldsymbol{\Xi}$ の推定

$$\begin{cases} \hat{\boldsymbol{\mu}} = (\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{Y}'\mathbf{1}_n/n, \\ \hat{\boldsymbol{\Xi}} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{Y}\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}. \end{cases}$$

- 問題点; $\boldsymbol{A}$ に相関が高い列の組がある → 不安定になる
  $\boldsymbol{X}$ に使う関数が柔軟すぎる → 過剰適合する

本研究の目的

これらの問題点を回避したい

アイデア; リッジ型の推定量の拡張

不安定になる⇒Hoerl and Kennard (1970) の推定量の流用
過剰適合する⇒Nagai (2011) の推定量の流用

→ 導入したパラメータの最適化のための C_p 型情報量規準の構築とその先のアイデアについて

ここからの話

提案する推定量

- $\boldsymbol{Y} = \mathbf{1}_n\boldsymbol{\mu}'\boldsymbol{X}' + \boldsymbol{A}\boldsymbol{\Xi}\boldsymbol{X}' + \boldsymbol{\mathcal{E}}$ のモデルで,
  不安定性・過剰適合を回避しつつ $\boldsymbol{\mu}$ と $\boldsymbol{\Xi}$ の推定 (= 経時変動の推定)
  - 仮定; $\boldsymbol{A}'\mathbf{1}_n = \mathbf{0}_k$ (各列で中心化されている), $\mathrm{rank}(\boldsymbol{A}) = k$, $\mathrm{rank}(\boldsymbol{X}) = q$, $\boldsymbol{\mathcal{E}} \sim \mathcal{N}_{n,p}(\mathbf{0}_n\mathbf{0}_p', \boldsymbol{\Sigma} \otimes \boldsymbol{I}_n)$, $n - k - p - 2 > 0$
  - $\boldsymbol{\mu}$ と $\boldsymbol{\Xi}$ の MLE; $\hat{\boldsymbol{\mu}} = (\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{Y}'\mathbf{1}_n/n$,
    $\hat{\boldsymbol{\Xi}} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{Y}\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}$.
    ここで, $\boldsymbol{S} = \boldsymbol{Y}'\{\boldsymbol{I}_n - \mathbf{1}_n\mathbf{1}_n'/n - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\}\boldsymbol{Y}/(n-k-1)$.
- Hoerl and Kennard (1970) などのアイデアを使うと

$$\begin{cases} \hat{\boldsymbol{\mu}}(\lambda) = (\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X} + \lambda\boldsymbol{I}_q)^{-1}\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{Y}'\mathbf{1}_n/n, \\ \hat{\boldsymbol{\Xi}}(\theta,\lambda) = (\boldsymbol{A}'\boldsymbol{A} + \theta\boldsymbol{I}_k)^{-1}\boldsymbol{A}'\boldsymbol{Y}\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X} + \lambda\boldsymbol{I}_q)^{-1}, \end{cases}$$

ここで $\theta \geq 0$, $\lambda \geq 0$ は罰則パラメータ.

ここからの話

θ と λ の最適化← 予測の観点で良い値を選ぶ

C_p 型情報量規準の構築 (1/3) -目的のもの-

- $\boldsymbol{Y} = \mathbf{1}_n\boldsymbol{\mu}'\boldsymbol{X}' + \boldsymbol{A}\boldsymbol{\Xi}\boldsymbol{X}' + \boldsymbol{\mathcal{E}}$ のモデルで,
  不安定性や過剰適合を回避しつつ,
  経時変動を推定するために, $\hat{\boldsymbol{\mu}}(\lambda)$, $\hat{\boldsymbol{\Xi}}(\theta,\lambda)$ を使う
  仮定 $\boldsymbol{A}'\mathbf{1}_n = \mathbf{0}_k$, $\mathrm{rank}(\boldsymbol{A}) = k$, $\mathrm{rank}(\boldsymbol{X}) = q$. $\boldsymbol{\mathcal{E}} \sim \mathcal{N}_{n,p}(\mathbf{0}_n\mathbf{0}_p', \boldsymbol{\Sigma} \otimes \boldsymbol{I}_n)$, $n - k - p - 2 > 0$.
  - それぞれの推定量; $\hat{\boldsymbol{\mu}}(\lambda) = (\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X} + \lambda\boldsymbol{I}_q)^{-1}\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{Y}'\mathbf{1}_n/n$,
    $\hat{\boldsymbol{\Xi}}(\theta,\lambda) = (\boldsymbol{A}'\boldsymbol{A} + \theta\boldsymbol{I}_k)^{-1}\boldsymbol{A}'\boldsymbol{Y}\boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X} + \lambda\boldsymbol{I}_q)^{-1}$.
    ここで, $\boldsymbol{S} = \boldsymbol{Y}'\{\boldsymbol{I}_n - \mathbf{1}_n\mathbf{1}_n'/n - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\}\boldsymbol{Y}/(n-k-1)$.

→ θ (≥ 0), λ (≥ 0) (罰則パラメータ) の最適化が必要

- 次の PMSE (予測平均二乗誤差) が小さくなるように選ぶ;

$\mathrm{PMSE}[\hat{\boldsymbol{Y}}] \overset{\mathrm{def.}}{=} E[E_{\boldsymbol{Y}_F}[\mathrm{tr}\{(\boldsymbol{Y}_F - \hat{\boldsymbol{Y}})\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}_F - \hat{\boldsymbol{Y}})'\}]]$,
ここで, $\boldsymbol{Y}_F$ は $\boldsymbol{Y}$ と独立に同一の分布から得られるもの,
$\hat{\boldsymbol{Y}}$ は今の $\boldsymbol{Y}$ などからの予測値, $E_{\boldsymbol{Y}_F}[\cdot]$ は確率変数 $\boldsymbol{Y}_F$ に基づいた期待値 ($E[\cdot]$ は $\boldsymbol{Y}$ に基づいた期待値).

C_p 型情報量規準の構築 (2/3) -変形と結果-

- $\boldsymbol{Y} = \mathbf{1}_n\boldsymbol{\mu}'\boldsymbol{X}' + \boldsymbol{A}\boldsymbol{\Xi}\boldsymbol{X}' + \boldsymbol{\mathcal{E}}$ のモデルで,
  提案する推定量 ($\hat{\boldsymbol{\mu}}(\lambda)$, $\hat{\boldsymbol{\Xi}}(\theta,\lambda)$) のために,
  予測値 ($\hat{\boldsymbol{Y}}(\theta,\lambda) = \mathbf{1}_n\hat{\boldsymbol{\mu}}'(\theta,\lambda)\boldsymbol{X}' + \boldsymbol{A}\hat{\boldsymbol{\Xi}}(\theta,\lambda)\boldsymbol{X}'$) に対し,
  $\mathrm{argmin}_{\theta\geq 0,\lambda\geq 0}\mathrm{PMSE}[\hat{\boldsymbol{Y}}(\theta,\lambda)]$ が欲しい.

変形 $\mathrm{PMSE}[\hat{\boldsymbol{Y}}(\theta,\lambda)]$ の定数を除いた部分;

$$E[\mathrm{tr}\{(\boldsymbol{Y} - \hat{\boldsymbol{Y}}(\theta,\lambda))\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \hat{\boldsymbol{Y}}(\theta,\lambda))'\}] + 2\,\mathrm{tr}(\boldsymbol{H}_\theta)E[\mathrm{tr}(\boldsymbol{G}_\lambda)],$$

ここで $\boldsymbol{G}_\lambda = \boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X} + \lambda\boldsymbol{I}_q)^{-1}\boldsymbol{X}'$,
$\boldsymbol{H}_\theta = \mathbf{1}_n\mathbf{1}_n'/n + \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A} + \theta\boldsymbol{I}_k)^{-1}\boldsymbol{A}'$.

→ **C_p 型情報量規準 (C_p 規準)**

$$C_p(\theta,\lambda) = \mathrm{tr}\{(\boldsymbol{Y} - \hat{\boldsymbol{Y}}(\theta,\lambda))\boldsymbol{S}^{-1}(\boldsymbol{Y} - \hat{\boldsymbol{Y}}(\theta,\lambda))'\} + 2\,\mathrm{tr}(\boldsymbol{H}_\theta)\mathrm{tr}(\boldsymbol{G}_\lambda).$$

C_p 型情報量規準の構築 (3/3) -lMC_p 規準へ-

目的 PMSE に基づいて $\hat{\boldsymbol{\mu}}(\lambda)$, $\hat{\boldsymbol{\Xi}}(\theta,\lambda)$ の θ と λ を最適化

→ PMSE の推定量として C_p 規準を作った

- C_p規準での θ, λ の最適化; $\mathrm{argmin}_{\theta\geq 0,\lambda\geq 0} C_p(\theta,\lambda)$

現実 $E[C_p(\theta,\lambda)] \neq \mathrm{PMSE}[\hat{\boldsymbol{Y}}(\theta,\lambda)] - 定数$

→ バイアス補正が必要

- Yangihara et al. (2023) の手法でバイアスが補正できるとすると

like Modified C_p (lMC_p) 規準

$$lMC_p(\theta,\lambda) = c \times \mathrm{tr}\{(\boldsymbol{Y}-\hat{\boldsymbol{Y}}(\theta,\lambda))\boldsymbol{S}^{-1}(\boldsymbol{Y}-\hat{\boldsymbol{Y}}(\theta,\lambda))'\} + 2\,\mathrm{tr}(\boldsymbol{G}_\lambda)\left\{\mathrm{tr}(\boldsymbol{H}_\theta) + \frac{\mathrm{tr}(\boldsymbol{R}_{\theta,\lambda})}{n-k-1}\right\},$$

ここで $c=1-(p+1)/(n-k-1)$, $\boldsymbol{R}_{\theta,\lambda}=\boldsymbol{Y}'\boldsymbol{H}_\theta\boldsymbol{Y}(\boldsymbol{I}_p-\boldsymbol{G}_\lambda)\boldsymbol{S}^{-1}$.

- lMC_p 規準での最適化; $\mathrm{argmin}_{\theta\geq 0,\lambda\geq 0} lMC_p(\theta,\lambda)$

ここからの話

数値実験結果の怪しい部分と先の話へ

- 数値実験の結果で怪しかった部分

$(n=30,q=5$ の場合)				従来法	提案手法		提案手法 ($\theta=0$)		提案手法 ($\lambda=0$)	
p	k	ρ_y	ρ_u	MLE	C_p	lMC_p	C_p	lMC_p	C_p	lMC_p
20	5	0.25	0.25	86.62	86.43	4495.16	86.43	4814.19	86.61	86.62
			0.50	85.84	85.70	6518.66	85.70	6801.80	85.82	85.92
			0.95	85.68	85.30	14817.17	85.81	15043.83	85.32	84.96
		0.95	0.25	86.51	86.50	86.53	86.50	86.51	86.51	86.53
			0.50	86.43	86.41	86.47	86.41	86.43	86.42	86.47
			0.95	86.27	86.00	86.27	86.26	86.27	86.00	86.27

→ lMC_p 規準ではバイアスが補正し切れていないのでは？

- $E[lMC_p(\theta,\lambda)] \neq \mathrm{PMSE}[\hat{\boldsymbol{Y}}(\theta,\lambda)] -$ 定数になっている？
- Yanagihara et al. (2023) の推定量に対する条件は満たしている
- λ を入れた影響が残っている？
  - $\lambda=0$ にした場合は上手くいっている
- $n=50$ では上手くいっているので条件に見落としがある？
  - このときはバイアスが小さくなり lMC_p 規準が C_p 規準と同じに？

→ λ の影響を考慮したバイアス補正が必要

※この先の話はアイデア段階の話

なぜバイアスが残っているのか？

- 途中までは Yanagihara et al. (2023) と同じ計算で変形可
  - $\boldsymbol{H}_\theta$ の方しか関連しないため
  - 再掲 $\boldsymbol{H}_\theta = \mathbf{1}_n\mathbf{1}_n'/n + \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A}+\theta\boldsymbol{I}_k)^{-1}\boldsymbol{A}'$
    - 具体的には, Yangihara et al. (2023) での期待値の計算の直前まで

→ 期待値を計算する箇所で $\boldsymbol{G}_\lambda$ の一部の項が暴れている？

再掲 $\boldsymbol{G}_\lambda = \boldsymbol{S}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1}\boldsymbol{X}'$,
$\boldsymbol{S} = \boldsymbol{Y}'\{\boldsymbol{I}_n - \mathbf{1}_n\mathbf{1}_n'/n - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\}\boldsymbol{Y}/(n-k-1)$.

補足 この $\boldsymbol{H}_\theta$ と $\boldsymbol{G}_\lambda$ を使うと $\hat{\boldsymbol{Y}}(\theta,\lambda) = \boldsymbol{H}_\theta\boldsymbol{Y}\boldsymbol{G}_\lambda$ と書ける

- $\hat{\boldsymbol{Y}}(\theta,0)$ については Yanagihara et al. (2023) の数値実験でバイアスが補正できていることが確認されている
- $\lambda=0$ でバイアスが補正できている (Yanagihara et al. (2023)) のは, Wishart 分布の性質をうまく使っているため
  - Wishart 分布に従う確率変数行列に関する定理を使って, 期待値計算

→ $\lambda\boldsymbol{I}_q$ を加えると直接 Wishart 分布 (やその派生分布) にならない

⇒ バイアスが残る？

λ 入りでのバイアス補正のアイデア (1/5)

- バイアス補正のための期待値の計算 (特に $\boldsymbol{G}_\lambda$ の一部) が厄介
- アイデア 1; 期待値の計算で厄介な項の一部 $((\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1})$ を Wishart 分布系に変形してから期待値の計算
  - 糸口 1.1 $n=50$ で上手く行っていた (n が大きくてバイアスが小さくなってるだけ?)
  - 糸口 1.2 $(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1}$ を非心 Wishart 分布などに変形する？
    - 定義 非心 Wishart 分布 $W_s(\mathfrak{u},\boldsymbol{\Omega},\boldsymbol{J})$ の定義 (RMLE Gupta & Nagar (2000) Th. 3.5.1)
      $\boldsymbol{V}\sim\mathcal{N}_{u,s}(\boldsymbol{M},\boldsymbol{\Omega}\otimes\boldsymbol{I}_u)$ ($\boldsymbol{\Omega}$ は $s\times s$ 正定値行列) のとき,
      $\boldsymbol{V}'\boldsymbol{V}$ が従う分布で $\boldsymbol{J}-\boldsymbol{\Omega}^{-1}\boldsymbol{M}'\boldsymbol{M}$ (非心度)
    - 補足 $\boldsymbol{M}=\mathbf{0}_u\mathbf{0}_s'$ のとき (つまり $\boldsymbol{J}=\mathbf{0}_s\mathbf{0}_s'$ のとき) が Wishart 分布
    - 注意 1 本や流派により s, $\mathfrak{u}$, $\boldsymbol{\Omega}$, $\boldsymbol{J}$ の順番が異なる場合有
    - 注意 2 行列型の多変量正規分布の定義も異なる場合も有
    (本研究では $E[\boldsymbol{V}]=\boldsymbol{M}$, $\mathrm{Cov}[\mathrm{vec}\{\boldsymbol{V}\}]=\boldsymbol{\Omega}\otimes\boldsymbol{I}_u$ となる正規分布を表す.)
- アイデア 2; 対応した項などをマクローリン展開して, Wishart 分布の高次の項の期待値の計算
  - 糸口 2.1 $\lambda=0$ の場合は Yanagihara et al. (2023) に対応
  - 糸口 2.2 Wishart 分布の高次の期待値の計算は可能

アイデア 1; Wishart 分布系へ (1/3) (2/5)

- $(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1}$ を変形 (注意; そのままでは非心 Wishart 分布でない)
  - $\lambda=0$ のときは Yanagihara et al. (2023) の話になるので, そこでの MC_p 規準との比較で $\lambda=0$ か否かを判断可能

→ $\lambda>0$ として変形し, $\ell=\lambda^{-1}$ とすると
$(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1} = \ell(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}(\ell\boldsymbol{I}_q+(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1})^{-1}$.

- $\boldsymbol{S} = \boldsymbol{Y}'\{\boldsymbol{I}_n - \mathbf{1}_n\mathbf{1}_n'/n - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\}\boldsymbol{Y}/(n-k-1)$ で, $\{\boldsymbol{I}_n - \mathbf{1}_n\mathbf{1}_n'/n - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\}$ が s べき等対称行列なので, $(n-k-1)\boldsymbol{S}\sim W_p(n-k-1,\boldsymbol{\Sigma})$ (Schott (2017), Theorem 11.25), よって, $(n-k-1)(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1} \sim W_q(r,(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1})$, ここで $r=n-k-p+q-1$ (Muirhead (1982), Theorem 3.2.11).

→ Wishart 分布の定義より, $(n-k-1)(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1} = \boldsymbol{T}'\boldsymbol{T}$ となる行列 $\boldsymbol{T}$ ($r\times q$ で $\mathrm{rank}(\boldsymbol{T})=q$ となる行列) があり, $\boldsymbol{T}\sim\mathcal{N}_{r,q}(\mathbf{0}_r\mathbf{0}_q', (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\otimes\boldsymbol{I}_r)$ とできる.

アイデア 1; Wishart 分布系へ (2/3) (3/5)

- $(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1} = \ell(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}((\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}+\ell\boldsymbol{I}_q)^{-1}$

再掲 $\ell=\lambda^{-1}$, $\boldsymbol{T}\sim\mathcal{N}_{r,q}(\mathbf{0}_r\mathbf{0}_q', (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\otimes\boldsymbol{I}_r)$ で $\mathrm{rank}(\boldsymbol{T})=q$ の行列で $(n-k-1)(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1} = \boldsymbol{T}'\boldsymbol{T}$, $r=n-k-p+q-1$.

→ $(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1} = \ell\boldsymbol{T}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T}+(n-k-1)\ell\boldsymbol{I}_q)^{-1}\ldots$
同じような形で面倒になっただけでは？

- $\boldsymbol{T}$ の特異値分解 (SVD) より, $(\boldsymbol{K}_1',\boldsymbol{K}_2')'\boldsymbol{T}\boldsymbol{L}=(\boldsymbol{D}^{1/2},\mathbf{0}_q\mathbf{0}_{r-q}')'$, ここで, $\boldsymbol{K}_1$ は $q\times r$ 行列, $\boldsymbol{K}_2$ は $(r-q)\times r$ 行列, $\boldsymbol{L}$ は $q\times q$ 直交行列, $\boldsymbol{D}$ は $\boldsymbol{T}'\boldsymbol{T}$ の固有値を対角に並べた $q\times q$ 行列.

→ $\boldsymbol{L}$ が直交行列より $\boldsymbol{K}_2\boldsymbol{T}=\mathbf{0}_{r-q}\mathbf{0}_q'$

→ $r\geq 2q$ ($\Leftrightarrow n-k-p-1\geq q$) として, $\boldsymbol{U}=\boldsymbol{K}_2'(\boldsymbol{I}_q,\mathbf{0}_q\mathbf{0}_{r-2q}')'$ と置くと, (SVD で直交行列の一部として $\boldsymbol{K}_2$ を得ているので $\boldsymbol{K}_2\boldsymbol{K}_2'=\boldsymbol{I}_{r-q}$.)
$\boldsymbol{T}'\boldsymbol{T}+\delta\boldsymbol{I}_q = (\boldsymbol{T}+\sqrt{\delta}\boldsymbol{U})'(\boldsymbol{T}+\sqrt{\delta}\boldsymbol{U})$, ここで $\delta=(n-k-1)\ell$.

←$r\geq 2q$ なら $\boldsymbol{T}'\boldsymbol{T}+\delta\boldsymbol{I}_q\sim W_q(r,(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1},\delta\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})$

アイデア 1; Wishart 分布系へ (3/3) (4/5)

- $(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1} = \ell\boldsymbol{T}'\boldsymbol{T}\{(\boldsymbol{T}+\sqrt{\delta}\boldsymbol{U})'(\boldsymbol{T}+\sqrt{\delta}\boldsymbol{U})\}^{-1}$ で, $\boldsymbol{T}\sim\mathcal{N}_{r,q}(\mathbf{0}_r\mathbf{0}_q', \boldsymbol{\Sigma}_*^{-1})$ より
$W_q(r,\boldsymbol{\Sigma}_*^{-1})$ や $W_q(r,\boldsymbol{\Sigma}_*^{-1},\delta\boldsymbol{\Sigma}_*)$ に従う確率変数行列の話で期待値の計算ができそう.
⇔ バイアス補正ができそう. ただし $\boldsymbol{\Sigma}_*=\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}$.

再掲 $\ell=\lambda^{-1}$, $\delta=(n-k-1)\ell$, $r=n-k-p+q-1$

問題 1 $n-k-p-1\geq q$ ($\Leftrightarrow r\geq 2q$) が必要
- $n=50$ で上手く行っていたのは, これが成立していたから？

問題 2 $W_q(r,\boldsymbol{\Sigma}_*^{-1})$ に従う確率変数行列 や
$W_q(r,\boldsymbol{\Sigma}_*^{-1},\delta\boldsymbol{\Sigma}_*)$ に従う確率変数行列の逆行列の積やべき乗の期待値などが必要

注意 それぞれの確率変数行列同士は独立ではない

⇐ このアイデアでの補正は無理なのでは？！

アイデア 2; マクローリン展開利用 (5/5)

- λ を入れた怪しい部分 $(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1}$ を マクローリン展開してから期待値計算に？！

→ 実際に (行列の関数であることに注意して) マクローリン展開すると,

$$(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}+\lambda\boldsymbol{I}_q)^{-1} = \sum_{i=1}^{\infty}(-1)^{i-1}\lambda^{i-1}\{(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1}\}^i.$$

- $(n-k-1)(\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X})^{-1} \sim W_q(n-k-p+q-1,(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1})$

→ Yanagihara et al. (2023) の手法が応用できるかも？！
理由 Yanagihara et al. (2023) では Wishart 分布の性質を利用している

→ 無限個の期待値が必要 アイデア 1 (Wishart 分布系にする) よりこっちの方針がよさそう

→ どこかで打ち切る？
- 展開する項が増えれば増えるほど, バイアスを小さくできる ⇔ 期待値の計算が大変 & 補正の項が増える
- 各項で期待値を計算し, オーダーで表現する？

まとめ

- 経時測定データの分析を考える

前提 全ての個体で測定時点が共通

実例 カナダの各都市の月別平均気温と緯度経度 (Kokoszka & Reimher, 2017)

→ $\boldsymbol{Y} = \mathbf{1}_n\boldsymbol{\mu}'\boldsymbol{X}' + \boldsymbol{A}\boldsymbol{\Xi}\boldsymbol{X}' + \boldsymbol{\mathcal{E}}$ (GMANOVA モデル) が分析によく使われる
- $\boldsymbol{X}$ に使う関数の重み付き和で経時変動を推定する形

問題点 従来の推定量の問題点;
$\boldsymbol{A}$ の列に相関の高い組がある場合, 不安定になる.
$\boldsymbol{X}$ に用いる関数が柔軟すぎる場合, 過剰適合する.

本研究 これらの問題点を回避する罰則付推定量の提案
- Hoerl & Kennard (1970), Nagai (2011) のアイデアの流用

→ 導入したパラメータ (θ,λ) の最適化が必要

⇒ C_p 型情報量規準を構築 (λ を入れるとバイアスが？！)

今後 アイデア 1・2 などで λ の影響込みバイアス補正など

参考文献

1. Gupta, A. K. and Nagar D. K. (2000). *Matrix Variate Distributions*. Chapman and Hall/CRC.
2. Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: applications to nonorthogonal problems. *Technometrics*, **12**, 69–82.
3. Kokoszka, P. & Reimherr, M. (2017). *Introduction to Functional Data Analysis*, Chapman and Hall/CRC.
4. Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley-Interscience.
5. Nagai, I. (2011). Modified C_p criterion for optimizing ridge and smooth parameters in the MGR Estimator for the nonparametric GMANOVA model. *Open Journal of Statistics*, **1**, 1–14.
6. Potthoff, R. F. & Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, **51**, 313–326.
7. Schott, J. R. (2017). *Matrix Analysis for Statistics* (Third Edition), John Wiley & Sons, Inc.
8. Yanagihara, H., Nagai, I., Fukui, K. & Hijikawa,Y. (2023). Modified C_p criterion in widely applicable models. *Smart Innov. Syst. Tec.*, **352**, Intelligent Decision Technologies 2023: Proceedings of the 15th KES International Conference on Intelligent Decision Technologies KES-IDT-2023 (eds. I. Czarnowski, R. J. Howlett & L. C. Jain), 173-182.

Shrinkage estimators in multivariate normal distributions with block compound symmetry covariance structures

今野　良彦　(大阪公立大学)*

多変量正規分布の分散共分散行列が block compound symmetry と呼ばれる制約下にあるときに, 平均ベクトルの推定問題を統計的決定理論の立場から考える. 適当な直交行列で標本空間の基底を変換することで, 標本空間の座標に作用する置換群の部分群に関して不変な構造をもつ分散共分散行列は, 成分が実, 複素, 四元数である Hermite 行列の直和で書けるできること [3] を援用して, 平均ベクトルの縮小型推定量を導出できることを報告した.

問題設定: block compound symmetry 構造をもつモデルの中で最も簡単なモデル

$$\boldsymbol{Y} \sim \mathsf{N}_d\big(\widetilde{\boldsymbol{\xi}},\, \boldsymbol{\Sigma}\big) \qquad \text{with} \qquad \boldsymbol{\Sigma} = \sigma^2\big\{(1-\rho)\mathbf{I}_d + \rho\mathbf{1}_d\mathbf{1}_d^\top\big\} \tag{1}$$

を考える. ただし, $\widetilde{\boldsymbol{\xi}} \in \mathbb{R}^d$, $\mathbf{I}_d$ は d 次の単位行列, $\mathbf{1}_d = (1,\, 1,\, \ldots,\, 1)^\top \in \mathbb{R}^d$, $0 < \sigma < \infty$, $-\dfrac{1}{d-1} < \rho < 1$ である. また, $(\,\cdot\,)^\top$ は転置作用素である. このモデルは intraclass covariance structures とも呼ばれる分散共分散構造をもつ多変量正規モデルである.

$n \geq 1$ とする. $\boldsymbol{Y}$ の $n+1$ 個の独立複製を $\boldsymbol{Y}_1,\, \boldsymbol{Y}_2,\, \ldots,\, \boldsymbol{Y}_{n+1}$ に対して

$$\overline{\boldsymbol{Y}} = \frac{1}{n+1}\sum_{i=1}^{n+1}\boldsymbol{Y}_i, \quad \boldsymbol{S} = \sum_{i=1}^{n+1}(\boldsymbol{Y}_i - \overline{\boldsymbol{Y}})(\boldsymbol{Y}_i - \overline{\boldsymbol{Y}})^\top$$

とおく. σ と ρ を未知としたとき, 平均ベクトル $\widetilde{\boldsymbol{\xi}}$ の推定問題を不変な損失関数

$$\widetilde{\mathsf{L}}\big(\widetilde{\boldsymbol{d}},\, \widetilde{\boldsymbol{\xi}} \,|\, \boldsymbol{\Sigma}\big) = (n+1)\big(\widetilde{\boldsymbol{d}} - \widetilde{\boldsymbol{\xi}}\big)^\top \boldsymbol{\Sigma}^{-1}\big(\widetilde{\boldsymbol{d}} - \widetilde{\boldsymbol{\xi}}\big)$$

のもとで考える. ただし, $\widetilde{\boldsymbol{d}}$ は, 観測 $\big(\overline{\boldsymbol{Y}},\, \boldsymbol{S}\big)$ に基づく $\widetilde{\boldsymbol{\xi}}$ の推定量である. さらに, $\widetilde{\mathsf{R}}(\widetilde{\boldsymbol{d}},\, \widetilde{\boldsymbol{\xi}} \,|\, \boldsymbol{\Sigma})$ を $\widetilde{\boldsymbol{d}}$ の危険関数とする. すなわち

$$\widetilde{\mathsf{R}}\big(\widetilde{\boldsymbol{d}},\, \widetilde{\boldsymbol{\xi}} \,|\, \boldsymbol{\Sigma}\big) = \mathsf{E}\big[\widetilde{\mathsf{L}}\big(\widetilde{\boldsymbol{d}},\, \widetilde{\boldsymbol{\xi}} \,|\, \boldsymbol{\Sigma}\big)\big]$$

である. ただし, $\mathsf{E}[\,\cdot\,]$ は $\big(\overline{\boldsymbol{Y}},\, \boldsymbol{S}\big)$ の同時分布に関する期待値である.

主結果: この推定問題の通常の canonical form を導出するために

$$\boldsymbol{X} = \sqrt{n+1} \times \overline{\boldsymbol{Y}} \sim \mathsf{N}_d\big(\boldsymbol{\xi},\, \boldsymbol{\Sigma}\big) \quad \text{and} \quad \boldsymbol{S} \sim \mathsf{W}_d\big(n,\, \boldsymbol{\Sigma}\big) \tag{2}$$

とおく. ただし, $\mathsf{W}_d\big(n,\, \boldsymbol{\Sigma}\big)$ は自由度 n, スケール行列 $\boldsymbol{\Sigma}$ の d 次の Wishart 分布で, $\boldsymbol{\xi} := \sqrt{n+1}\widetilde{\boldsymbol{\xi}}$ で, $\boldsymbol{\Sigma}$ は (1) で与えられる. ここで, $\boldsymbol{\Sigma}$ は正定置であるが, $n < d$ の場合

* 〒 558-8585 大阪市住吉区杉本 3-3-138　大阪公立大学 大学院理学研究科
e-mail: konno@fc.jwu.ac.jp
web: https://mcm-www.jwu.ac.jp/~konno/index.html

には $\mathbf{S}$ は正則ではない (その逆行列が存在しない) ことに注意する. 平均ベクトル $\widetilde{\boldsymbol{\xi}}$ の推定問題は, 平均ベクトル $\boldsymbol{\xi}$ を不変な損失関数

$$\mathsf{L}\big(\widehat{\boldsymbol{\xi}}, \boldsymbol{\xi} \big| \boldsymbol{\Sigma}\big) = \big(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}\big)^\top \boldsymbol{\Sigma}^{-1} \big(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}\big) \tag{3}$$

のもとで考えると同じである. ただし, $\widehat{\boldsymbol{\xi}} = \sqrt{n+1}\widetilde{\boldsymbol{d}}$ は観測 $\big(\boldsymbol{X}, \boldsymbol{S}\big)$ に基づく推定量である.

Canonical form

$\boldsymbol{X}$ と $\mathbf{S}$ は独立に (2) の分布に従い, 分散共分散行列は (1) で与えらえる. ただし, $\sigma > 0$, $-\dfrac{1}{d-1} < \rho < 1$ である. このとき, $\big(\boldsymbol{X}, \mathbf{S}\big)$ に基づき $\boldsymbol{\xi}$ の推定問題を不変な損失関数 (3) のもとで考える.

平均ベクトル $\boldsymbol{\xi}$ の James-Stein 型推定量とその positive-part 推定量は

$$\widehat{\boldsymbol{\xi}}^{\mathrm{JS}} = \left(1 - \frac{d-2}{(d-1)n\boldsymbol{X}^\top \left\{ (\operatorname{Tr}\boldsymbol{S} - \frac{1}{d}\mathbf{1}_d^\top \mathbf{S} \mathbf{1}_d)\boldsymbol{I}_d + \frac{1}{d}\big(\mathbf{1}_d^\top \boldsymbol{S} \mathbf{1}_d - \operatorname{Tr}\boldsymbol{S}\big)\mathbf{1}_d\mathbf{1}_d^\top \right\}^{-1} \boldsymbol{X}}\right)\boldsymbol{X},$$

$$\widehat{\boldsymbol{\xi}}^{\mathrm{PJS}} = \left(1 - \frac{d-2}{(d-1)n\boldsymbol{X}^\top \left\{ (\operatorname{Tr}\boldsymbol{S} - \frac{1}{d}\mathbf{1}_d^\top \mathbf{S} \mathbf{1}_d)\boldsymbol{I}_d + \frac{1}{d}\big(\mathbf{1}_d^\top \boldsymbol{S} \mathbf{1}_d - \operatorname{Tr}\boldsymbol{S}\big)\mathbf{1}_d\mathbf{1}_d^\top \right\}^{-1} \boldsymbol{X}}\right)_+ \boldsymbol{X}$$

で与えられる. ただし, $\operatorname{Tr}(\,\cdot\,)$ は行列のトレースで, $(a)_+ = \max\{0, a\}$ for $a \in \mathbb{R}$ である.

定理

任意の $\boldsymbol{\xi} \in \mathbb{R}^d$ と $\sigma > 0$, $-\dfrac{1}{d-1} < \rho < 1$ に対して

$$\mathsf{R}\big(\widehat{\boldsymbol{\xi}}^{\mathrm{PJS}}, \boldsymbol{\xi} \big| \boldsymbol{\Sigma}\big) \le \mathsf{R}\big(\widehat{\boldsymbol{\xi}}^{\mathrm{JS}}, \boldsymbol{\xi} \big| \boldsymbol{\Sigma}\big) \le \mathsf{R}\big(\boldsymbol{X}, \boldsymbol{\xi} \big| \boldsymbol{\Sigma}\big).$$

モデル (2) のもとでの数値実験による提案した James-Stein 型推定量の改良率（最尤推定量に対するもの）と block compound symmetry covariance structures モデルへ拡張した結果について簡単に報告した.

References

[1] Andersson, S. (1975). Invariant Normal Models. *Ann. Statist.* **3** 132-154.

[2] Faraut, J., Korányi, A.(1994). Analysis on symmetric cones. The Clarendon Press, Oxford University Press, New York.

[3] Graczyk, P., Ishi, H., Kołodziejek, B., Massam, H. (2022). Model selection in the space of Gaussian models invariant by symmetry. *Ann. Statist.* **50** 1747-1774.

[4] Jensen, S.T. (1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance. *Ann. Statist* **16** 302-322.

Survival Analysis in Astrophysics: from Univariate to Multivariate

Tsutomu T. Takeuchi

1. Division of Particle and Astrophysical Science, Nagoya University,Furo-cho, Chikusa-ku, Nagoya 464–8602, Japan
2. The Research Center for Statistical Machine Learning, the Institute of Statistical Mathematics, 10–3 Midori-cho, Tachikawa, Tokyo 190–8562, Japan

Observational data in astronomy are invariably subject to truncation due to the detection limits of observation instruments. When estimating distribution functions of astronomical statistical quantities from such data, it is appropriate to apply survival analysis of truncated data. However, survival analysis, which has been developed in the field of statistics, was practically unknown in the field of astronomy until the 1980s, and astronomers have developed estimation methods independently of the framework in mathematical statistics. Although many of the methods devised by astronomers finally converge with survival analysis, they have been mathematically unorganized, and statistical discussions became systematic only after the 21st century.

In this work, we introduced a method to estimate a luminosity function of galaxies from astronomical surveys. This type of analysis is known to be a truncated data analysis. We start from the discussion for the case on a univariate luminosity function. The estimator of the univariate luminosity function from truncated data was introduced by Lynden-Bell (1971) in the astronomical community. Woodroofe (1985) reformulated the Lynden-Bell's estimator. He introduced it to the statistical community, and provided a mathematically rigorous proof on its convergence properties.

However, bivariate counterpart of this problem has been less studied in astronomy. Bivariate luminosity function in astronomy is constructed based on a univariate luminosity function estimated form one primary selection wavelength band. Then, the conditional luminosity function from the secondary selection band, estimated with the Kaplan–Meier estimator, is combined. However, this method leads to a biased estimate for the bivariate function, due to the single-band selection. van der Laan (1996) provided an estimator for a bivariate distribution function of a doubly-truncated data. Though this is an appropriate method for bivariate truncation, it drops the information from censored data, which are inevitably introduced in a 2-band selection. Then, we propose a merged method that can also incorporate the censored information. This can be achieved by setting the truncation and censoring variables. Performance verification and application to actual data will be presented with numerical examination in our future work.

We should note that, in all the related discussions, independence of X and Y is supposed. However, this independence should be examined carefully. Particularly, in recent deep surveys, the time evolution effect cannot be negligible in the sample. This problem has been pointed out by some authors (e.g., Efron and Petrosian, 1992). Since this issue will become more important in future projects, then more mathematically rigorous discussions will be desired.

References

Efron, B. and Petrosian, V. (1992). A Simple Test of Independence for Truncated Data with Applications to Redshift Surveys, ApJ, **399**, p. 345, November, DOI: http://dx.doi.org/10.1086/171931.

Lynden-Bell, D. (1971). A method of finding distances to double galaxies, *Monthly Notices of the Royal Astronomical Society*, **155**, p. 95, DOI: http://dx.doi.org/10.1093/mnras/155.1.95.

van der Laan, M. J. (1996). Nonparametric Estimation of the Bivariate Survival Function with Truncated Data, *Journal of Multivariate Analysis*, **58** (1), 107–131, URL: https://www.sciencedirect.com/science/article/pii/S0047259X96900421, DOI: http://dx.doi.org/https://doi.org/10.1006/jmva.1996.0042.

Woodroofe, M. (1985). Estimating a Distribution Function with Truncated Data, *The Annals of Statistics*, **13** (1), 163–177, URL: http://www.jstor.org/stable/2241151.

クローナル植物に関する地上データと地下情報の紐づけ統合と統計モデル

島谷健一郎（統計数理研究所）

植物の中には、種子繁殖に加え、クローンを作って拡大する種がある。イチゴのように地上部にランナーを飛ばす種はクローンの広がり方が目に見えるが、ササのように地下茎を伸長させる種では、掘らないとクローン繁殖の実態が見えないため、その実態は意外なほど知られていない。実際、以下のような素朴な疑問が未解決だった。

1. 各クローンは先端を毎年伸ばしているか。
2. クローナル植物は、地上部（株、ラメット）を増やすクローナル成長と、各地上部の大きさ（葉の大きさ、茎の長さ、等々）の拡大からなる。では、遺伝子の単位としての 1 クローン（個体、ジェネット）の成長率はどう定義され、どう推定できるか。
3. 地下茎は、分枝（枝分かれ、1 つの地上部が複数の地上部を作る）は、どのくらいの頻度で行われ、クローン全体（ジェネット）の拡大に定量的にどのくらい貢献しているか。
4. 個々の地上部（ラメット）の寿命は何年くらいか。。
5. 地下茎の寿命は何年くらいか。堀り起こすと、地上部と地上部を結ぶ地下茎が出てくるが、地下茎が 1 本も出ていない地上部もある。一方、地下茎はあるが地上部はすでに枯死した古株も掘り起こされる。
6. 例えば 10m 広がったクローン（ジェネット）は何年（以上）前の実生から始まったか。

2005 年から 2007 年の 6　7 月、北海道十勝地方のスズラン集団について、2m x 28m の調査プロットを設置し、地上部（株、ラメット）のモニタリング調査を行った。さらに遺伝子（マイクロサテライト、中立マーカー）分析によりクローン識別も行った。2007　2009 年には、調査プロット内の面積にして約半分を掘り起こし、地上部と地下茎のつながりを記録した。

こうした地上及び地下部のデータ・情報から、個体群行列モデルを用いてクローン全体（ジェネット）の成長率を定式化し、クローナル繁殖率や分枝率の推定と合わせて論文として公開した。

Araki SK., Shimatani IK, Ohara M. 2022. Genet dynamics and its variation among genets of a clonal plant Convallaria keiskei. OIKOS. doi: 10.1111/oik.09367

この論文の中で、上記の未解決問題のうち、（北海道十勝地方のスズランについては）１　３の答えを提供した。鍵となったのは、個体群行列モデルにおけるクラス分けを、クローンの広がりの先端か否かで分けた（　葉の枚数と長さ）点にある。しかし、4　6は未解決のまま残った。

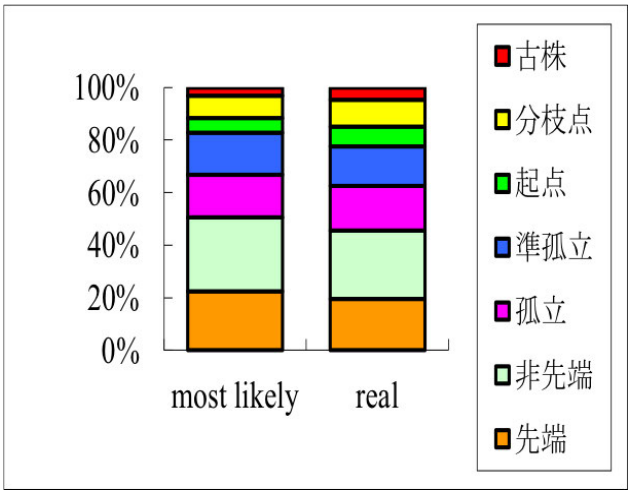
この論文の個体群行列モデルのクラス分けを、先端･非先端と（実測できない株が多いが）齢にし、死亡率と繁殖率を齢依存ロジスティック式などで与えることにより、数理モデルとしてクローンごとの株（ラメット）動態をシミュレーションできる。実データとの照合では、モニタリング調査により生誕１２年目であることがわかっている株しか齢は不明なので、齢依存死亡率の実証は直接には無理である。しかし、掘り起こした株は地下茎とつながっていたり、途中で切れていたりする。株の３つ組　地下茎２本のつながりかたを分類し、適当なカテゴリーに分ける（右図、生は生きている株、c は地下茎が切れていた、e は消えていた、d は古株、b は翌年の株の新芽なので e と同値)。一方、適当な事前分布のもとでクローンの広がりをシミュレーションで作成し、同じように５つ組単位で分類する。両者を比べて近い結果を残したパラメータセットを選抜することで、ベイズ統計の事後分布のランダムなサンプルを得ることができる（近似ベイズ法)。

事前分布は、株（ラメット）の死亡率関数（と繁殖率関数）と地下茎の死亡率関数の両方に与えるが、前者は、上記の個体群行列モデルにおいてクローン全体（ジェネット）が平衡状態に到達していることを仮定することで、齢について平均した死亡率や繁殖率が求められ、それと地上部と地下部の実データからの推定値を比べることでも選抜できる。

推定に成功すれば、冒頭の ４　５の答えを与えられる。さらに、こうした数値（事後分布）に、地上部（株）間距離や、地下茎間の角度を合わせてランダムウオークモデルを作ることで、クローン全体の広がる早さも推定でき、冒頭の問題６にも答えられる。

本発表時点では、この方法で事後分布のランダムサンプルを得る所には至らなかった（右図)。その原因について考察し、改善策を提示するにとどまった。

e生d	起点	6
e生生	起点	79
e生e	孤立	156
e生b	孤立	39
c生c	準孤立	23
e生c	準孤立	72
c生b	準孤立	12
c生e	準孤立	65
生生b	先端	68
d生e	先端	21
生生e	先端	133
b生e	先端	2
c生d	非先端	6
c生生	非先端	103
d生c	非先端	1
d生生	非先端	15
生生c	非先端	28
生生d	非先端	3
生生生	非先端	141
生d生	古株	6
cd生	古株	5
生dc	古株	5
cdd	古株	6
dd生	古株	8
生dd	古株	3
ddd	古株	4
ddc	古株	2
ed生	古株	11
edd	古株	3
生de	古株	1
		1027



Mobility analyses of geospatial agents

藤原 直哉
東北大学大学院情報科学研究科

地理情報システム（Geographical Information Systems, GIS）技術の発展により地理空間に関連するデータ分析は大きく進歩してきたが、従来の GIS においては静的な情報の分析に重点が置かれ、時間的な変化を伴うデータに関しては十分に研究が行われてこなかった。しかし、近年では動的な情報への関心が高まりを見せている。そのような動的なデータの代表的なものとして、モノや人の流動に関するデータがある。その背景には、携帯電話の GPS などによってデータが大量に取得が可能となったことなどが挙げられる。GPS では、自動車のような移動体のデータも取得可能であり、データの特性や目的を考慮した分析手法の開発が強く求められている。本稿では、このような流動データの分析に関連し、主に著者が近年関わった研究を紹介する。なお、人流データ分析一般の研究動向については、レビュー論文 [1] などを参照されたい。

地理空間における流動の中でも、最も注目されてきた対象の一つは人流である。都市間の人口流動をはじめとして、人流は古くから地理学、交通工学、経済学など多くの分野で研究されてきた。日本においては、パーソントリップ調査にみられるような大規模調査が長期に渡って行われ、これに基づいて都市内における人の移動が詳細に分析されてきた。人流分析には、アンケートを中心として多様な種類のデータが活用されてきたが、González ら [2] によって携帯電話由来の大規模データが導入されて以降、人流研究の方法論は大きく変貌を遂げつつある。

人流データの分析には、大きく分けて非集計の個別データを直接用いる方法と、何らかの形で集計したデータを利用する方法がある。近年の革新は、非集計データにおいて高精度かつ大規模な観測が可能となった点にあるが、データ集計手法とその分析手法の開発は依然として重要な課題である。この点に関して我々はいくつかの貢献を行ってきた。例えば、文献 [3] においては、人流データの新しい集計手法を提案した。人流データの集計方法としては、始点から終点に移動する人数をカウントする、Origin-Destination（OD）行列の方法が標準的であるが、人の行動には時間的周期性があるため、OD 行列を用いて連続する複数のトリップを発生させると、例えば人の移動距離などを過大推計してしまう。そこで、我々は、自宅位置も合わせて集計を行う方法（Home-Origin-Destination, HOD）を提案し、午前中に自宅を出発して夜間に帰宅するという、現実的な移動パターンを生成することに成功した。感染症の拡大シミュレーションなど、この集計方法が有効である場面は多いと考えられる。

また、集計データの取り扱いにネットワーク分析の手法を導入している。例えば、ネットワークにおけるコミュニティ分割アルゴリズムである Map Equation [4] を地域間通勤流動データに適用し、通勤流動の空間単位の検出を試みた [5]。その結果、流動データは距離などの地理的情報を明示的に含んでいないにもかかわらず、得られた階層的コミュニティ構造は近隣自治体をまたぎつつも概ね地理的に連続した形を示した。得られたコミュニティは、地域的な政策を実施する空間単位としての応用などが期待される。

地理空間における流動データの代表は人流であるが、近年、別種の流動データも注目されつつある。その一例が自動車の軌跡データである。文献 [6] においては、自動車の軌跡データから、仙台港のコンテナターミナル内において記録がある車両を、コンテナを輸送するトラクタヘッドと判断し、その行動の分析を行った。仙台港のコンテナターミナルには、空のコンテナを積載しているトレーラー用のゲートと、実入りコンテナを積載しているトレーラー用のゲートが存在しているが、それぞれのゲートの訪問前後にトラクタヘッドが訪れる場所の特性が、空コンテナ用ゲートと実入りコンテナ用ゲートで異なることを明らかにした。

位置情報（ジオタグ）付きツイートデータのように、位置情報以外の情報が付与されているデータも存在している。文献 [7] では、ジオタグ付きツイートデータにおいて、地域ごとに、単語が言及される回数を調べた。その結果、地名においては、使用頻度が有意に高い地域が存在し、その地名が指し示す場所と解釈できることを示した。一方、一般名詞においては、このような使用頻度の地域依存性は観測されなかった。これらの結果は、SNS データによって、人間の空間認識や地名の使用傾向について理解できる可能性を示唆している。

以上のように、地理空間における流動データは急速に普及しつつあり、新たな分析手法の確立が求められる。今後も使用できるデータの模索と手法開発を進めるとともに、地理空間における「流動」の理解を深めていく予定である。

References

[1] Y. Du, T. Aoki, and N. Fujiwara, Journal of Computational Social Science **8**, 90 (2025).

[2] M. C. González, C. A. Hidalgo, and A.-L. Barabási, Nature **453**, 779–782 (2008).

[3] Y. Du, T. Aoki, and N. Fujiwara, The European Physical Journal Plus **139**, 403 (2024).

[4] M. Rosvall and C. T. Bergstrom, Proceedings of the National Academy of Sciences **105**, 1118–1123 (2008).

[5] S. Fujishima, N. Fujiwara, Y. Akiyama, R. Shibasaki, and R. Sakuramachi, International Journal of Economic Theory **16**, 38–50 (2020).

[6] H. Zheng, C. Zhao, Y. Ogawa, R. Shibasaki, and N. Fujiwara, in 2024 IEEE International Conference on Big Data (BigData), 6823–6833 (2024).

[7] T. Hiraoka, T. Kirimura, and N. Fujiwara, PLoS ONE **20**, e0325022 (2025).

政治学研究における統計分析の応用動向

河村 和徳（拓殖大学）

1 はじめに

かつて政治学分野における統計分析と言えば、政治に参加するアクターに対するサーヴェイデータ（個票データ）分析と投票結果や経済・財政指標などを用いたアグリゲートデータ（集計データ）分析のいずれかであった。前者は行動科学の系譜に位置付けられ、後者は政治経済学の系譜から研究が進められてきた。しかしながら、近年の欧米のメジャーな研究雑誌では、政治学心理実験的なアプローチによる論文や大規模自然言語モデルを活用した論文、また位置情報などを利用したビッグデータを用いた論文などが掲載されるようになり、自然科学系の研究者との共著論文も広く見られるようになった。こうした動きの背景には、政治学の「科学志向」があると考えられる。政治学の権威と呼ばれる研究者の言説をそのまま受け入れるのではなく、実証的に政治現象を把握しようという動きがこうした研究動向につながっている。

本報告では、簡単ではあるが、研究事例を紹介するとともに、統計分析に偏りすぎることの課題について述べたいと思う。

2 応用動向

2.1 実験的アプローチと fMRI の利用

実験的アプローチは、現実世界をそのまま観察する統計分析とは異なり、研究者が観察対象に「介入する」点が大きく異なる。そして、独立変数を無作為割り当てするこのアプローチは、観察的な研究と比べ因果推論上の利点があるとされる（肥前 2016）。なお、実験に用いる刺激は言語情報が中心ではあるが、Asano & Patterson（2024）のような映像情報や岡田ほか（2025）のような音声情報のように、非言語情報を刺激に用いる研究も行われるようになっている。心理実験をより進め、谷口（2020）に見られるように与えられた政治情報を fMRI を利用し、有権者の認知の中でもアンケートでは扱いにくい反応を分析する試みもなされている。アンケート調査では、社会的望ましさバイアスなどの影響が生じる可能性が少なくない。Political Neuroscience として、より精緻な議論をしていこうという動きと捉えることができる。

2.2 大規模言語処理の応用

SNS の普及によって大規模コーパスを活用して政治現象を分析しようという動きもある。政治的デマの拡散などの研究がこれにあたる。池田ほか（2018）などが該当する。ただ、これらの研究の中には「政治・社会を分析対象にしている」だけで、政治学的な議論が少なくない。言い換えると、「やったらこうなりました」という事象の説明だけに終わってしまっているのである。統計的な分析から導かれる結果に対し社会科学的に妥当な解釈を行うためには、李・河村・木村（2024）のように社会科学の研究者と自然科学分野の研究者による共同チームによる研究が不可欠であり、そうした要請から政治学でも学際的な枠組みでの研究が増えつつある。

2.3　位置情報データを用いた研究

投票所にいつ、どのようなタイミングで赴くか。近年、人々の政治行動を位置情報データを用いて分析する動きもある。たとえば、Harada, Ito, & Smith（2024）の試みを挙げることができる。

3　生じている課題

政治学の分野でも、他分野での研究動向を刺激として受けつつ、より研究手法の高度化、複雑化が進んでいる。しかしながら、新しい試みを進める上での課題も見えつつある。

その大きなものの 1 つが、データを提供する環境の未整備である。統計分析をするためには何らかの形でデータが提供される必要がある。政治学の分野では、政府からのデータ提供がどうしても主となろう。しかしながら、わが国では統計分析に利用できるデータセットの整備が進んでいないところがある。地方議会の議事録然り、政治資金データ然りである。データ収集に割くエネルギーを減らすためには、制度設計を働きかけ、その構築に誰かがコミットする必要がある。しかし、これにコミットするとなると、研究者の誰かが犠牲にならなければならない。

また、近年、ヨーロッパを中心に個人情報の管理の厳格化が進んでいることも研究を難しくさせる流れを生んでいる。多くの人が情報提供を拒否すれば、サンプルの偏りは避けられず、研究遂行を難しくする。政治学の分析の中には、政治家や官僚を特定することで成り立つものもある。公人としてどこまで情報を提供すべきか、より高度な統計利用の一方で、データを集める根本部分が課題となっている点を忘れてはならない。

また前述したように、研究者側に結果を妥当な解釈をするための知識も求められる。

文献

肥前洋一. 2016.『実験政治学』勁草書房.

ASANO Masahiko and PATTERSON Dennis. 2024. "The Multiple Patterns of Factional Influence: Evidence from the 2021 LDP Presidential Election", *Asian Politics & Policy*, 16(4).

岡田陽介・後藤心平・戸田香・遠藤勇哉・河村和徳. 2025.「非言語情報としての声の違いがアナウンサーの印象およびニュースの認知に与える影響」日本公共政策学報告。

谷口尚子. 2020.「ｆMRI を用いた有権者の脳活動の計測」『法学研究』93(1).

池田圭佑・榊剛史・鳥海不二夫・栗原 聡. 2018.「口コミに着目した情報拡散モデルの提案およびデマ情報拡散抑制手法の検証」『情報処理学会論文誌数理モデル化と応用（TOM）』11(1).

李昕翮, 河村和徳, 木村泰知. 2024.「BERTopic モデルを利用した震災復興研究：宮城県議会議員は何を発言してきたのか」『行動計量学』51(2) 181-191.

Harada, Masataka, Gaku Ito, & Daniel M. Smith. 2024. "Using Cell-phone Mobility Data to Study Voter Turnout," *Political Behavior*.

選挙研究における非言語情報の測定と分析――政治家の声の高低と印象形成

岡田陽介（拓殖大学）

1　はじめに

政治家から有権者に提供される情報には，言語情報（選挙公報やビラの内容，演説内容など）と，非言語情報（顔の表情や，話し方や身振りなど）とが存在する．選挙研究においては演説や公約の内容分析など，言語情報研究では一定の蓄積があるが，非言語情報研究では，ポスターの笑顔の程度や顔の美顔度などに基づいた試みがあるものの，言語情報研究と比較して研究蓄積に乏しい．本報告では，非言語情報研究の新たな試みとして政治家の声に着目する．そして，収集した政治家の音声から周波数を測定し，声の高低が印象形成や投票選択，得票率に与える効果についての分析方法の検討と選挙研究おける非言語情報の意義の検討を行う．

2　候補者の声のデータ収集と測定

現状，政治家の声を網羅したアーカイヴは存在しない．したがって，声の収集にあたっては，街頭演説の録音や web 上の動画，政見放送の録画などを用いることになる．ただし，短い選挙運動期間中に全候補者の街頭演説を録音することは物理的に困難であり，また，ノイズを避けなければならない点など困難な場合もある．

本報告で検討する一連の研究では，主として党首討論会や公開討論会の動画，政見放送の録画を用い，音声冒頭の文意の通る数秒間の基本周波数（F_0）測定している．通常，発言冒頭では名前を名乗ることが多く，測定対象を揃えることが可能になる．

図 1 は 2021 年衆院選の政見放送での岸田文雄元首相の測定例である．「岸田文雄です」（kishida fumio desu）との発言のうち，周波数は最大値 154.4Hz（H で示した「shi」）から最小値 82.8Hz（L で示した「su」）まで幅があるが，平均周波数は 106.1Hz であった．

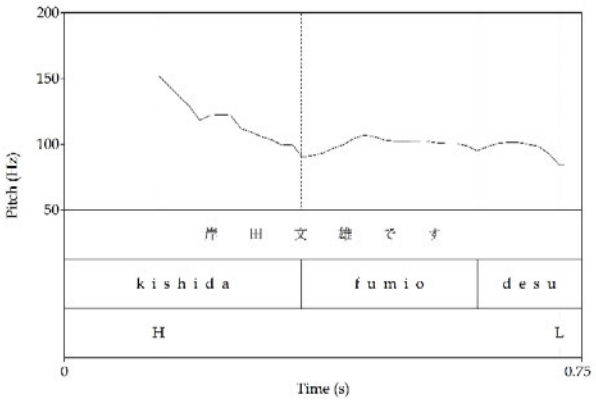


図 1　岸田文雄の音声（2021 年衆院選政見放送）

3　政治家の声と分析方法

候補者の「声」の高低と有権者の印象形成や投票選択・得票率の関連についてはいくつかの分析方法が考えられる．第 1 に，音声の周波数を（高く／低く）変更した実験刺激を用い，実験室実験やサーベイ実験にて好感度や信頼度などを測定する方法，第 2 に，周波数と現実の得票率や

当落などのアグリゲートデータとの関連みる方法，第 3 に，周波数と世論調査データとを結合した疑似実験データによって，政治家への感情温度や投票選択との関連をみる方法である．

4　候補者の声と有権者の印象形成・投票選択・得票率

図 2 は，実在の地方議員の音声（157.6Hz）から作成した，高条件（197.8Hz）と低条件（112.3Hz）の実験刺激を元に，都内私立大学生 54 名(男性 34 名, 女性 20 名：平均年齢 20.7 歳（SD = 2.5)）を対象に好感度・信頼度（それぞれ 0-100）を測定した実験結果である．好感度の低条件で好感度が高く有意な差が確認された（M_high = 44.9 *vs.* M_low = 57.9, t (35) = 2.05, p = .048)．また，図 3 は，2014 年衆院選小選挙区立候補者 (945 人／959 人) の音声から測定した周波数と得票率の関連である．弱いながらも低い声で得票率が高い傾向が確認された（r = −.09, p = .007)．

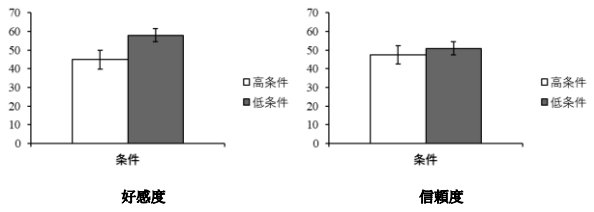


図 2　地方議員の音声と好感度・信頼度

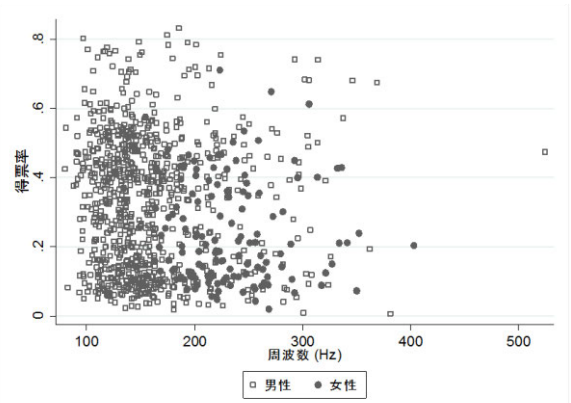


図 3　2014 年衆院選の候補者周波数と得票率

さらに図 4 は, 政党党首（2009 年衆院選・2010 年参院選）の周波数と有権者に対する世論調査データ（JES IV）を結合し，諸要因を統制した上で，党首の声と比例区での当該政党への投票との関連を示したものであるが，周波数の上昇に伴って，投票選択確率が減少することが示されている．

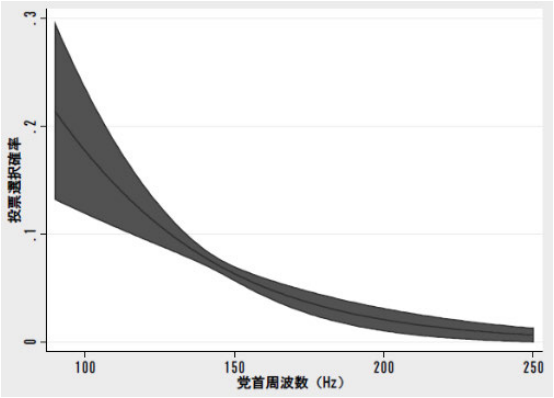


図 4　党首周波数と比例区投票

5　選挙おける非言語情報研究の意義

一連の分析では，政治家の声の高低が好感度や信頼度などの印象形成だけでなく，実際に得票率や投票選択に効果をもたらすことが示された．これは，仮に政治家が言語情報として同一の内容を提示していたとしても，声など非言語情報の声質によっては同一の内容が伝達されるわけでないことを示唆している．また，政治活動・選挙運動において，演説は最も古典的かつ基礎的な手段であるが，それを媒介する「声」という非言語情報が，政治家の印象やイメージ形成において機能しうることを示唆するものである．

テキスト分析から見える自民党県連の組織構造

高知県立大学
吐合大祐

1. 問いの所在

本研究の目的は、都道府県議会議員の選挙公報データを通じて、地方議員レベルの政党組織の特徴を可視化することである。近年の政党研究は、地方レベルの政党組織の特徴を実証的に論じるものが増加している（例：Detterbeck 2012）。国政レベルの政治家の再選活動や政策決定において影響をもたらすのは、普段から国政政治家の活動を支え、同じ政党に所属する地方政治家の存在が大きい。このように、政党組織の実態を「中央」「地方」それぞれのレベルから多層的に捉えようするアプローチを「マルチレベル政党組織論」と呼ぶ。これに連なる研究群は、政党組織の特徴を中央レベル・地方レベルにそれぞれ分類した上で、各レベルで活動する政治家（候補者）の特徴が政党組織の構造にどう影響をもたらすのかを実証してきた（例：建林編 2013）。

これら研究は政党組織を多層的に捉えることで、これまで見落とされてきた地方レベルの持つ政治的インパクト（例：選挙協力、地方からの政策要望）を描こうと試みる。しかしいずれの研究も、地方レベルの政治的/政策的特徴を十分に掬えていない。すなわち、地方を政党政治の分析枠組みに加えようとはしているが、地方自治体それぞれで形成されるはずの地方レベルの政治家の行動をうまく測定できていない。

本研究は、こうした課題を克服するために、政党組織を構成する地方レベルの政治家の政策立場を、地方議員の関心から測定することを試みるものである。

2. 本研究の分析

地方レベルの政策関心はどのように測定できるか。本研究は、地方議員が選挙時に公開する「選挙公報のテキスト分析」を方法として用いる。地方分権改革が浸透した今日の政治環境に目配りした場合、地方議員は自治体の政策決定の権限を有する知事の存在を考慮する事が重要である（例：馬渡 2010; 辻 2015）。

本研究では、地方議員の政策関心を測定するために、テキスト分析の一手法である「Wordscore」を用いる。Wordscore[1]とは、選挙公報や議会議事録といった政治的文書を一次元上にスケーリングするための指標の一つであり、Laver et al.（2003）をはじめとする研究によって展開された、教師あり学習モデルに分類される分析手法の一つである。この手法では、分析者が単語辞書を任意に設定した上で、本来観察したい文書の政策位置を設定した辞書をもとに学習した上で測定できる手法である。知事に注目する理由は、（１）選挙状況的に「自民党 v.s.立憲民主党」の対立構造が現れているため分類しやすいこと、（２）分権改革以降の知事の政策的影響力が非常に大きいと評価されていること、の 2 点である。知事を辞書として設定した上で、2019 年選挙に立候補した自民党公認候補の政策立場のバリエーションを測定し、各県の地方議員の政策関心を定義づける。

用いるデータは、2019 年の地方議会議員の選挙公報である。サンプルサイズは 2302 で、今回の分析では自民党公認ならびに推薦を得た議員のみを図表にまとめる。また今回は、地方議員の政策関心を県別に可視化した上で、自民党地方議員の政策関心のばら

[1] 詳細は Lowe（2008）を参照。

つき具合とその県別の多様性を論じることにしたい。選挙公報のテキスト分析には、R の quanteda 内に格納されている日本語の前処理機能を利用する。加えて、特定する単語は出現単語のうち上位 50%から 99%の範囲内に位置付けられる単語を用いることにしたい。そうすることで、自治体を跨く形で議員の政策関心を比較できるからである。

3. 分析結果

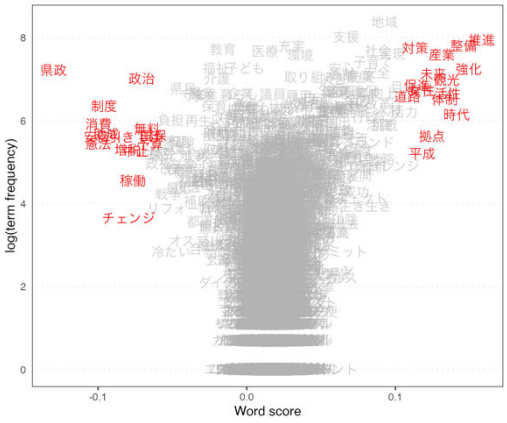


図１：単語スコアの分布

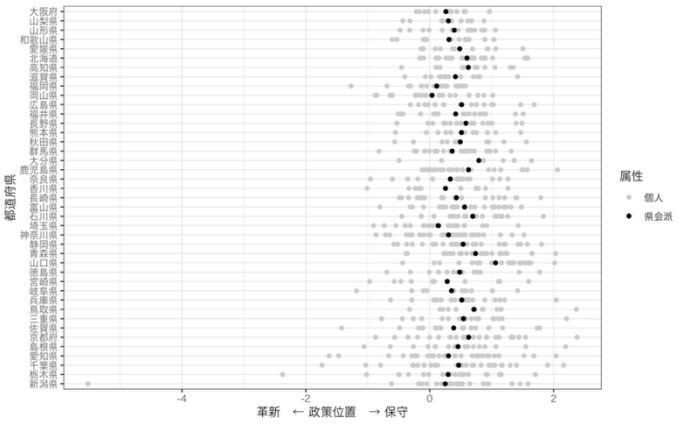


図２：県別の自民党県議の政策立場

そこで、2 つの分析結果を見てみよう。図 1 は、知事選挙の公報を辞書として用いた場合の、選挙公約に登場する単語のスコア分布を可視化したものである。これにより、選挙公報で用いられる単語イデオロギー位置の推定が可能となる。今回は、正の値に近づくほど「自民党・公明党公認の知事候補が用いやすい単語」、負の値に近づくほど「立憲民主党・国民民主党・共産党公認の知事候補が用いやすい単語」であることを意味している。図 1 の結果を元に、各議員の文書（すなわち選挙公約）の文言から県議個人のイデオロギーを推定したのが図 2 である。これを見ると、県議個人の政策位置がばらついていること、県によって平均値にも多様性が見られること、そして県単位で見た場合だとバリエーションにも差があるため、県ごとに政策関心さらには会派ごとの意思決定システムに違いが生じていることが示唆される。

4. まとめ

以上の分析により、都道府県ごとの地方議員の政策立場、さらには県議を束ねる地方組織の構造にも違いが見られると予測される。県議会議員は地方議会内で統一的な行動を取る必要があるためである。今後は、この分析をもとに、近年頻発する「保守分裂選挙」、また政党組織の重要な活動である「国会議員の新人候補のリクルートメント」との関係を明らかにしていきたいと考えている。

【参考文献】

Detterbeck, Klaus. 2012. *Multi-Level Party Politics in Western Europe.* Palgrave Macmillan

Laver, Michael, Kenneth Benoit and John Garry. 2003. “Exctacting Policy Positions from Political text Using Words as Data.” *American Political Science Review*, 97(2): 311-331.

Lowe, Will. 2008. “Understanding Wordscores.” *Political Analysis* 16: 356-371.

建林正彦編. 2013. 『政党組織の政治学』東洋経済新報社.

感情分析を用いた議員態度の可視化：北海道議会におけるオリンピック議論の事例

小樽商科大学商学部社会情報学科・李昕翮

キーワード：北海道議会、地方議会コーパス、感情分析、オリンピック、ソーシャル・ライセンス

一、 はじめに

オリンピックは、典型的な「政治的メガ・イベント」として位置づけられてきた（Grix 2013）。先行研究は、国家的叙事や社会的ライセンス、多層ガバナンスの観点からオリンピックの政治性を論じてきた。中谷・河村（2020）によると、中央からの丸投げや自治体内部の意見分岐が、議会発言を慎重で条件付きの支持へと導くことを示唆している。本研究は、この観点から 2011〜2023 年の北海道議会発言を対象とする。方法はテキスト計量的手法である。分析では「政策立場」と「感情」を組み合わせて検討する。ここから二つの仮説を提示する。

H1（叙事の失効−情感の負転）：国家的叙事が地域経験と乖離する場合、議員の情感は中立・期待からネガティブ・懸念へ移行する。

H2（責任下移−条件付き支持）：財政や実施の責任が地方に転嫁される場合、議員の発言は積極的支持よりも中立・検討、あるいは条件付き支持に傾く。

二、 研究方法

本研究のデータは、日本地方議会会議録コーパスに基づき、北海道議会定例会における一般質問のうち、「オリンピック／パラリンピック」に関する発言を対象としたものである。本研究では、辞書とルールを組み合わせた弱教師ありアノテーション手法（Lexicon- and rule-based weakly supervised annotation）を採用した。この手法は、政治テキストの感情や立場を測定する研究において広く実践されてきた系譜に位置づけられる（Taboada et al. 2011；Young and Soroka 2012）。

具体的には、推進・拡充・強化・増額といった方向語や肯定評価語を「支持」と定義した。一方、見直し・削減・縮小・撤回・中止などの方向語や否定評価語を「反対」とした。該当しない場合は「中立・検討」と分類した。

離散感情は、多ラベル付与で抽出する。怒り・懸念・悲しみ・嫌悪・驚き・喜び・期待・信頼・感謝の語彙群を基礎とした。該当しない場合は、一般的な極性（ポジ／ネガ／中立）へフォールバックさせる。分類の枠組みについては、NRC-Emotion- Lexicon（Mohammad and Turney 2013）を参照した。

計量化は行単位を最小観測単位とする。各行iに立場スコア$s_i \in \{-1、0、1\}$と感情スコア $v_i \in \{-1、0、1\}$ を付与する。文長 $\log(1 + 文字数)$ と強度語の有無に基づく重みw_iを構成し、加重集計に用いる（Taboada et al. 2011）。集約はボトムアップ型で行った。すなわち、行単位のデータを発言単位にまとめ、さらに議員×年度単位に集計し、最終的に年度単位に集約した。推移は月次平均で示し、不確実性は自助法（ブートストラップ）により 95%信頼区間を推定する（Efron and Tibshirani 1994）。

$$S = \frac{\sum_i w_i s_i}{\sum_i w_i} \qquad V = \frac{\sum_i w_i v_i}{\sum_i w_i} \tag{1}$$

ここでSは集計単位における政策立場スコア、Vは感情スコアであり、iは当該単位に含まれる行を表す。重みw_iは以下のように定義した。

$$w_i = \log(1 + \text{字数}_i) \times \text{強度係数}_i \times \text{モデル信頼度}_i \tag{2}$$

以下では、算出された立場スコア S の構成要素と重み付けの根拠について述べる。まず、行レベルで付与した立場スコアの集計にあたり、まず文長に基づく正規化を導入した。具体的にはlog (1 + 文字数)を重みとし、長文の情報量を反映させる一方で、極端に長い発言が平均値を支配しすぎることを抑制した。このような対数的補正は、情報検索分野でも広く妥当性が確認されている（Salton and Buckley 1988; Schütze et al. 2008）。

三、 研究結果

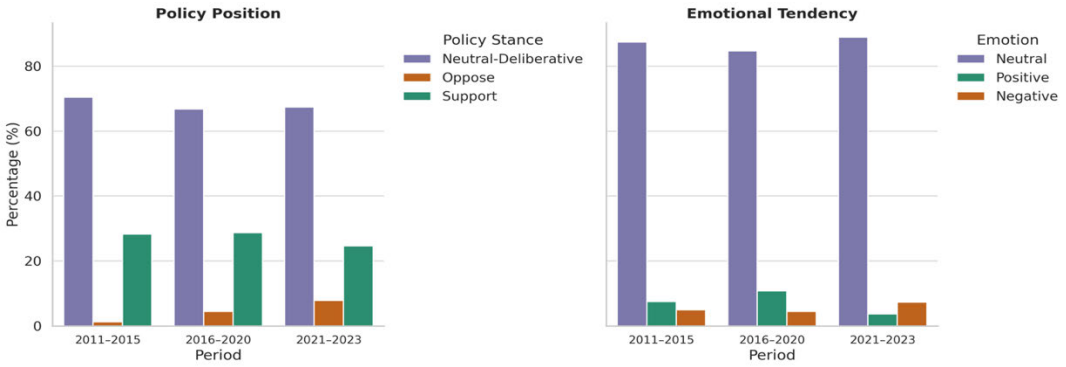


図 1 、政策立場・感情傾向の構成比（全体像 2011–2023）

図 1 は 2011～2023 年における北海道議会のオリンピック関連発言の政策立場と感情傾向を示す。左が政策立場、右が感情傾向であり，いずれも「中立／検討」が多数を占める。支持は時間とともに減少し、2021 年以降には反対が顕在化した。感情次元でも中立が主流だが、2016～2020 年に一時的に正の期待が増え，2021 年以降は否定的表現が大幅に増加した。これらは仮説 1（叙事の失効による感情の負転）と仮説 2（責任下移による条件付き支持）を裏付ける結果である。

参考文献

Efron, B., & Tibshirani, R. J. (1994). An introduction to the bootstrap. Chapman and Hall/CRC.

Grix, J. (2013). Sport politics and the Olympics. Political studies review, 11(1), 15-25.

Mohammad, S. M., & Turney, P. D. (2013). Nrc emotion lexicon. National Research Council, Canada, 2, 234.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. Information processing & management, 24(5), 513-523.

Schütze, H., Manning, C. D., & Raghavan, P. (2008). Introduction to information retrieval (Vol. 39, pp. 234-265). Cambridge: Cambridge University Press.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), 267-307.

Young, L., & Soroka, S. (2012). Affective news: The automated coding of sentiment in political texts. Political communication, 29(2), 205-231.

中谷昌弘， 河村和徳. (2020). 奈良県におけるオリンピックキャンプ地誘致に関する一考察. 研究紀要 「姫路日ノ本短期大学」, 42, 69-81.

Harnessing Kolmogorov–Arnold Networks for Accurate and Efficient Time Series Forecasting

Lingyu Jiang[1], Fangzhou Lin[1,2], Michael Zielewski[1], Kazunori D Yamada[1]

[1]Tohoku University [2]Worcester Polytechnic Institute

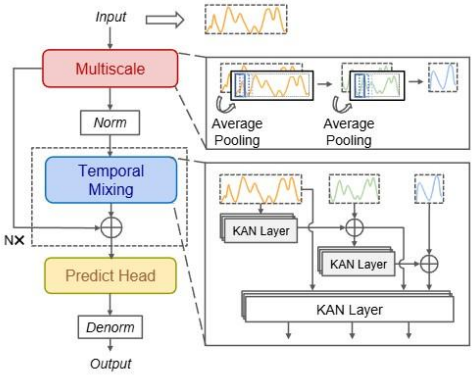
Introduction

Long-term time series forecasting (LTSF) plays a critical role in many real-world applications such as energy management, traffic planning, and weather prediction. Over the past decade, deep learning has revolutionized this task. Early approaches relied on recurrent neural networks (RNNs), particularly Long Short-Term Memory networks (LSTMs), which captured temporal dependencies but suffered from training inefficiency and limited scalability. Later, Transformer-based architectures, empowered by self-attention mechanisms, became dominant by modeling long-range dependencies more effectively. Recently, however, simple multilayer perceptrons (MLPs) have re-emerged as competitive alternatives, demonstrating that lightweight architectures can sometimes outperform more complex Transformer-based models.

Despite this progress, most state-of-the-art MLP-based methods rely on stacking additional handcrafted modules, such as frequency decomposition or patch mixing layers, to compensate for the flat structure of MLPs. While these extensions initially boosted performance, recent benchmark studies highlight diminishing returns, indicating the marginal benefit of adding complexity is limited. To address this challenge, we propose Kolmogorov–Arnold Networks (KAN) as a next-generation modeling core. Unlike conventional MLPs, KAN leverages adaptive spline-based basis functions, enabling fine-grained local modulation of nonlinearities. This unique property suggests that KAN can serve as a more powerful and flexible foundation for advancing LTSF models.

Proposed Method: KANMixer

To rigorously evaluate the capability of Kolmogorov–Arnold Networks (KAN) as a modeling core for LTSF, we designed a concise, yet effective architecture named KANMixer. Unlike existing approaches that rely on heavy modules and intricate designs, our goal was to keep the model simple and isolate the contribution of KAN itself.



KANMixer integrates two representative paradigms widely used in LSTF: an explicit modeling component, which captures multi-scale temporal dynamics through downsampling and an implicit modeling component, which hierarchically fuses temporal dependencies. The overall ar chitecture is illustrated in Figure 1.

This design not only allows us to examine KAN's standalone effectiveness in LTSF but also reveals how KAN performs under explicit and implicit modeling strategies that dominate current practice. Extensive experiments on

seven benchmark datasets demonstrate that KANMixer achieves performance on par with state-of-the-art models, and even surpasses them on several datasets, despite its minimalistic design.

Experimental Results

We conducted extensive experiments on seven widely used benchmark datasets, including the ETT series, Exchange Rate, Weather, and Electricity. KANMixer was compared against recent state-of-the-art models such as Transformer-based, MLP-based, and KAN-based approaches. Our results show that KANMixer consistently achieves performance on par with the best existing methods, and in several cases, surpasses them. Due to space limitations, detailed numerical results and comparisons will be presented in the accompanying presentation slides.

Discussion

To further understand the modeling capability of KAN, we conducted a series of controlled experiments within the KANMixer framework. First, we compared depth-wise variants of KAN and MLP. The results show that KAN consistently outperforms MLP across all depths, but stacking beyond three layers fails to provide additional gains and often leads to gradient explosion, suggesting that both KAN and MLP face optimization challenges at excessive depths.

We also performed component-wise ablation studies. Among the different modules, the KANbased prediction head contributed the most significant improvements in performance. This can be attributed to the fact that the final mapping from latent features to forecasts is highly complex, and the adaptive basis functions of KAN offer superior flexibility in approximating such mappings compared to conventional MLPs.

Furthermore, we evaluated different basis functions within KAN, including Chebyshev, Fourier, and Wavelet. Only the B-spline basis function consistently provided superior performance and stability, highlighting that KAN's effectiveness fundamentally stems from the strong approximation capacity of B-spline functions.

Finally, we investigated the impact of decomposition priors, a commonly used paradigm in time series forecasting. Interestingly, while decomposition improved performance when MLP served as the backbone, it degraded performance when combined with KAN. This suggests that the adaptive representation power of KAN may be hindered by overly strong handcrafted structural priors, which restrict its flexibility to learn directly from raw data.

Conclusion and Future Work

In this work, we presented KANMixer, a simple yet effective architecture that employs Kolmogorov–Arnold Networks as the modeling core for LTSF. Experiments show that KANMixer achieves accuracy comparable to, and in some cases surpassing, state-of-the-art models, while remaining structurally concise.

KANMixer remains an efficient model. Compared to large Transformer-based approaches, it requires far fewer parameters and computations, though its complexity is moderately higher than the most minimal MLP designs. This balance allows KANMixer to retain practicality while offering stronger modeling capacity.

The main limitation is computational overhead: training time and GPU memory increase due to the lack of optimized CUDA kernels for spline operations. These are engineering, not fundamental, challenges. Future work will focus on efficiency improvements such as kernel optimization and model compression to broaden the applicability of KAN-based models.

クロスオーバー試験における持続血糖データの階層ガウス混合モデリング

新川裕也[1)]，原めぐみ[2)]，川口淳[3)]

[1)]佐賀大学医学系研究科　博士課程，[2)]佐賀大学医学部社会医学講座，

[3)]佐賀大学医学部 生物統計学・生物情報学分野

1. 背景・目的

持続血糖測定器（CGM; Continuous Glucose Monitoring）は従来の自己血糖測定器と異なり，血糖値データを持続的に取得でき血糖値の変動を評価できるため，臨床・研究の双方で普及してきている．血糖値の変動に関して一般的に，変動係数や振幅, Time in Range（TIR）/Time Above Range（TAR）/Time Below Range（TBR）などの指標が評価項目として用いられている[1]．これらの指標はスケール不変で比較が容易であるという利点がある一方で，計測期間内で 1 つの代表値を使用するため，どの状態（通常時，食後，就寝中など）のときに血糖値がどの程度変化するのか，という分布上の位置付けまでは十分に示せない．本研究で扱う，桑菱茶の食後血糖上昇抑制への効果を測るためのプラセボ対照クロスオーバー試験[2]のような状況では，食後高血糖領域でどの程度血糖値が下がるのかという介入効果をより具体的に示すことが理想である．本研究では, CGM によって得られた血糖値の分布を GMM(Gaussian Mixture Model)[3]でモデル化することで血糖値の状態を分けて考え，さらに，クロスオーバー試験における個人差，介入効果，時期効果，持ち越し効果を同時に考慮した混合効果モデル[4]で GMM の各コンポーネントの平均値を表す階層構造を考えることで，食後高血糖領域での介入の影響を評価することを目的とする．

2. 方法

本研究で分析されたデータは，桑菱茶の継続摂取による食後血糖上昇抑制効果を調査したプラセボ対照 2 剤 2 期クロスオーバー試験の一環で計測された CGM による 15 分間隔の血糖値データである．モデルと表記法について説明する．被験者識別子を$i = 1, 2, \dots, N$，クロスオーバー試験の period を $j = 1,\ 2$ とする. sequence について $s_i = 1$ は被験者iが$j = 1$に treatment1 を受け、$j = 2$で treatment2 を受けることを示し，$s_i = 2$だとその逆の順序を示す．観測時点を$t = 1, 2, \dots, T_{ij}$として，観測された血糖値を y_{ijt} とする. GMM のコンポーネント数を$h = 1,2, \dots K$とする．ここで，各コンポーネントの平均値が混合効果モデルに従うと仮定するとモデル式は次のように記述できる．

$$P\left(y_{ijt}\middle|\Theta\right) = \Sigma_{h=1}^{K}\pi_{ij,h}N\left(\mu_{ij,h}, \sigma_h^2\right)$$
$$\mu_{ij,h} = m_h + b_{i,h} + \beta_h^{(trt)}x_{ij}^{(trt)} + \beta_h^{(per)}x_{ij}^{(per)} + \beta_h^{(seq)}x_{ij}^{(seq)}$$
$$b_{i,h} \sim N\left(0, \sigma_{b,h}^2\right)$$

ここで，$\pi_{ij,h}$，$\mu_{ij,h}$は被験者i，時期jにおける GMM の各コンポーネントの重みと平均で，σ_h^2は分散を示す．m_hはコンポーネントhのベースラインを表し，$b_{i,h}$は被験者ごとのランダム効果を示し，$\sigma_{b,h}^2$はランダム効果に正規分布に仮定したときの分散を示す．$x_{ij}^{(trt)}, x_{ij}^{(per)}, x_{ij}^{(seq)}$はそれぞれ treatment, period, sequence を表すバイナリデータで，$\beta^{(trt)}, \beta^{(per)}, \beta^{(seq)}$はそれらの固定効果を示す．$\Theta$はパラメータ集合 $\left\{\pi_h, m_h, \sigma_h^2, b_{i,h}, \beta_h^{(trt)}, \beta_h^{(per)}, \beta_h^{(seq)}, \sigma_h^2, \sigma_{b,h}^2\right\}$を示している．本研究の主要な推定はベイズ推定に基づき，未知パラメータに事前分布を与え，観測データの尤度とベイズの定理より事後分布を導出し事後平均と 94%最高密度区間(HDI)で要約した．本研究では$K = 3$として，混合モデルのラベルスイッチング問題を避けるため，順序制約（$\mu_1 < \mu_2 < \mu_3$）を課しサンプル後に再ラベリングを行った．計算には Python と PyMC5 を用い NUTS による MCMC で近似した．

3. 結果

本モデルにより,個人差，時期効果，持ち越し効果を同時に考慮した上で, CGM から取得された血糖値の状態ごとの平均及び滞在確率を推定できた．事後平均は$\mu_1 = 91.2, \mu_2 = 112.3, \mu_3 = 144.0$で、$\beta_3^{(trt)} = -3.8$ (94%HDI： $-5.7 \sim -1.9$) だった．桑菱茶飲用期間中は高血糖領域での平均値が下がる結果となった．

4. 考察

本モデルでは個体内ランダム効果と状態（潜在クラス）を同時に扱うことで，従来の代表値での解析では捉えにくい CGM データの複雑な構造，分布の形状・滞在確率の差を治療効果として定量化できることが示された．今回の手法では，クラスタ別の介入効果が推定されるので，介入による多様な反応を評価することができ，今後の CGM を用いた臨床試験に応用できると考えられる．

5. 参考文献

[1] Danne, Thomas, et al. "International consensus on use of continuous glucose monitoring." *Diabetes care* 40.12 (2017): 1631-1640.

[2] Shinkawa, Yuya, et al. "Evaluation of the Postprandial-Hyperglycemia-Suppressing Effects and Safety of Short-Term Intake of Mulberry Leaf and Water Chestnut Tea: A Randomized Double-Blind Placebo-Controlled Crossover Trial." *Nutrients* 17.14 (2025): 2308.

[3] Reynolds, Douglas. "Gaussian mixture models." *Encyclopedia of biometrics*. Springer, Boston, MA, 2015. 827-832.

[4] 折笠秀樹. *クロスオーバー試験の計画および解析*. Diss. University of Toyama, 2016.

関数合成コントロール法とその拡大

岡野遼（一橋大学・理研 AIP）

栗栖大輔（東京大学）

1. はじめに

合成コントロール法は, 単一の処置ユニットと複数の対照ユニットが存在するパネルデータの設定で, 処置ユニットにおける因果効果を推定する方法であり, 主に社会科学の分野で広く用いられている [1, 2]. この方法では, 合成コントロールと呼ばれる対照ユニットの加重平均を用いて処置ユニットの潜在結果変数を推定し, 因果効果の推定を行う. また, 合成コントロール法は近年様々な形に拡張されており, 中でも Ben-Michael, et al. [3] は, 拡大合成コントロール法と呼ばれる, オリジナルの推定量にバイアス補正を施す方法を提案している.

一方で, 合成コントロール法では通常結果変数がスカラーないしはベクトルである状況を想定するが, 応用によっては関数, 分布, ネットワークなど, より複雑な構造を持つデータが結果変数として現れることがある. 例えば, 妊娠中絶の規制が年齢別出生率曲線に与える因果効果や, 最低賃金の上昇が所得分布に与える因果効果に関心がある場合などである. 合成コントロール法でこのような複雑な構造を持つ結果変数を扱うためには, 何らかの形で既存の枠組みを一般化する必要がある.

本研究では, 合成コントロール法とその拡大の枠組みを, 結果変数が（可分な）ヒルベルト空間に値をとる場合に一般化し, 推定量の推定誤差の理論解析等を行う. 我々の枠組みでは, L^2 空間に値をとる関数データが結果変数として扱える他, 分布やネットワークも適切な変換を施すことで結果変数として扱うことが可能である. また, Kurisu et al. [4] では, 結果変数が測地距離空間に値をとる場合での合成コントロール法を提案しているが, 我々のアプローチは, それと比べて, ヒルベルト空間というより構造がある空間に舞台を限定しているため, 推定誤差などに関する理論解析がしやすいというメリットがある.

2. 設定

$i = 1, ..., N$ で添字付けられる N 個のユニットの結果変数が, $t = 1, ..., T$ で添字付けられる T 個の期間にわたって観測されるとする. ユニット i の時期 t における, 処置を受けた時と受けなかった時の潜在結果変数をそれぞれ Y_{it}^I, Y_{it}^N と書く. 合成コントロール法では, $1 < T_0 < T$ なるある時期 T_0 が存在して, 時期 $t = 1, ..., T_0$ では全てのユニットは処置を受けず, 時期 $t = T_0 + 1, ..., T$ では $i = 1$ なるユニットのみが処置を受ける状況を仮定する. 従って, 観測される結果変数 Y_{it} は,

$$Y_{it} = \begin{cases} Y_{it}^N & \text{if } i \geq 2 \text{ or } t \leq T_0 \\ Y_{it}^I & \text{if } i = 1 \text{ and } T_0 + 1 \leq t \leq T. \end{cases}$$

となる.

本研究では, 結果変数が可分な実ヒルベルト空間 $\mathcal{H}$ に値をとると仮定する. $\mathcal{H}$ の典型例としては, コンパクト区間 $\mathcal{I}$ 上の二乗可積分関数全体がなす空間 $L^2(\mathcal{I})$ がある. また, 一次元分布の空間やネットワークの空間も, その要素を適切に変換することによって, 可分な実ヒルベルト空間として扱うことが可能である.

以上の設定の下で, 処置ユニット $(i = 1)$ の処置後の期間 $t = T_0 + 1, ..., T$ における個人因果効果

$$\tau_t = Y_{1t}^I - Y_{1t}^N = Y_{1t} - Y_{1t}^N$$

に注目し, その推定を行う. 特に, 潜在結果変数 Y_{1t}^N は観測されないため, これをデータから推定する必要がある.

3. 提案手法の概要

まず, 基本となる関数合成コンロール法を導入する. Δ^{N-1} を $\mathbb{R}^{N-1}$ における単体とする：

$$\Delta^{N-1} = \left\{ \gamma = (\gamma_2, ..., \gamma_N) \in \mathbb{R}^{N-1} : \sum_{i=2}^{N} \gamma_i = 1, \gamma_i \geq 0 \text{ for } i = 2, ..., N \right\}.$$

関数合成コントロール法では, 処置前の期間 $t = 1, ..., T_0$ において, 処置ユニットの結果変数を対照ユニット $(i = 2, ..., N)$ の結果変数の加重平均で近似することを考え, その近似誤差が最も小さくなるような重みを計算する. すなわち,

$$\hat{\gamma}^{\text{scm}} = \underset{\gamma \in \Delta^{N-1}}{\arg\min} \sum_{t=1}^{T_0} \left\| Y_{1t} - \sum_{i=2}^{N} \gamma_i Y_{it} \right\|_{\mathcal{H}}^2$$

なる重み $\hat{\gamma}^{\text{scm}} \in \Delta^{N-1}$ を計算する ($\|\cdot\|_{\mathcal{H}}$ は空間 $\mathcal{H}$ に与えられているノルムである). そして, 処置後の期間 $t = T_0 + 1, ..., T$ に対し, 処置ユニットの潜在結果変数 Y_{1t}^N を, 対照ユニットの結果変数の加重平均

$$\hat{Y}_{1t}^{N,\text{scm}} = \sum_{i=2}^{N} \hat{\gamma}_i^{\text{scm}} Y_{it}$$

で推定する. これにより, 個人因果効果 τ_t の推定量として, $\hat{\tau}_t^{\text{scm}} = Y_{1t} - \hat{Y}_{1t}^{N,\text{scm}}$ が得られる.

処置前の期間の近似誤差 $\sum_{t=1}^{T_0} \|Y_{1t} - \sum_{i=2}^{N} \hat{\gamma}_i^{\text{scm}} Y_{it}\|_{\mathcal{H}}^2$ が 0 でない場合, 標準的なデータ生成過程の下で, 推定量 $\hat{Y}_{1t}^{N,\text{scm}}$ は Y_{1t}^N に対してバイアスを持つ. そこで, Ben-Michael, et al. [3] に従って, 推定量にバイアス補正を施した拡大推定量の構成を考える. 処置後の期間の潜在結果変数 $Y_{it}^N (i = 1, ..., N)$ に対し, 何らかのモデルを仮定し, そのモデルに基づく Y_{it}^N の推定値を $\hat{m}_{it}^N$ とする. そして, Y_{1t}^N に対する拡大推定量を

$$\hat{Y}_{1t}^{N,\text{aug}} = \sum_{i=2}^{N} \hat{\gamma}_i^{\text{scm}} Y_{it} + \underbrace{\left(\hat{m}_{1t} - \sum_{i=2}^{N} \hat{\gamma}_i^{\text{scm}} \hat{m}_{it} \right)}_{\hat{Y}_{1t}^{N,\text{scm}} \text{のバイアスの推定値}}$$

により定義する. これにより, 個人因果効果 τ_t の拡大推定量として, $\hat{\tau}_t^{\text{aug}} = Y_{1t} - \hat{Y}_{1t}^{N,\text{aug}}$ が得られる. Ben-Michael, et al. [3] では, 結果変数のモデルとして特にリッジ回帰モデルを考え, その正則化パラメータをうまく選ぶことで, 拡大推定量がオリジナルの推定量よりも小さい推定誤差を持ち得ることを示している. 本研究でも, 関数データ間のリッジ回帰に基づいて結果変数の推定値 $\hat{m}_{it}^N$ を定め, 拡大推定量の構成を行う.

講演では, 拡大推定量の具体的な構成方法を紹介した後, 推定量の推定誤差に関する理論的な結果や, 数値実験・実データ解析の結果等を報告する予定である.

参考文献

[1] Alberto Abadie and Javier Gardeazabal. The economic costs of conflict: A case study of the basque country. *American economic review*, Vol. 93, No. 1, pp. 113–132, 2003.

[2] Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, Vol. 105, No. 490, pp. 493–505, 2010.

[3] Eli Ben-Michael, Avi Feller, and Jesse Rothstein. The augmented synthetic control method. *Journal of the American Statistical Association*, Vol. 116, No. 536, pp. 1789–1803, 2021.

[4] Daisuke Kurisu, Yidong Zhou, Taisuke Otsu, and Hans-Georg Müller. Geodesic synthetic control methods for random objects and functional data. *arXiv preprint arXiv:2505.00331*, 2025.

Modeling Structural Changes in Distributional Data

Jeng-Min Chiou and Yu-Ting Hsu
Institute of Statistics and Data Science, National Taiwan University

Statistical analysis of distributional data has become increasingly important in diverse fields such as demography, economics, and the social sciences. Such data often evolve over time or with respect to covariates and may undergo structural changes driven by demographic transitions, economic crises, or policy reforms. Identifying these changes is essential for understanding how distributions shift in both location and variability. Classical regression methods, designed for Euclidean responses, are not well suited for this task, particularly when responses lie in nonlinear metric spaces.

The Wasserstein space of probability distributions provides a natural setting for regression with distributional responses. It preserves the geometry of densities and offers meaningful notions of distance and averages. Building on this framework, we develop Fréchet piecewise regression (FPR), an extension of classical piecewise regression to the Wasserstein space. The model partitions the predictor domain into intervals separated by unknown breakpoints and fits regression functions defined via Fréchet means. This formulation allows the predictor–response relationship to adapt flexibly across segments while retaining the intrinsic geometry of the space. The framework is flexible enough to accommodate both continuous transitions, where regression functions remain smooth at breakpoints, and jump transitions, where abrupt changes occur.

Estimation proceeds by minimizing a piecewise Fréchet loss over the entire predictor domain. The procedure begins with an initial set of candidate breakpoints, obtains regression function fits, and then updates the breakpoints iteratively until the overall loss stabilizes. For discontinuous breakpoints, additional predictors are incorporated to capture jump effects at segment boundaries, extending the formulation used in the continuous case. This iterative procedure yields stable estimates of both regression functions and breakpoint locations. Because the number of breakpoints is typically unknown, we supplement estimation with a bootstrap-based selection strategy. Breakpoint significance is assessed through partial regression effect tests adapted to the Wasserstein setting, and insignificant candidates are removed using a backward elimination algorithm. To control false discoveries, we apply the Benjamini–Hochberg procedure. These steps are complemented by evaluation metrics such as mean integrated squared error and an adjusted Wasserstein coefficient of determination, which help determine the appropriate number of breakpoints.

Theoretical analysis shows that the estimated breakpoints converge in probability to their population counterparts and that the fitted regression functions converge to the true conditional

Fréchet mean. These results provide a rigorous foundation for the method. The assumptions underlying the consistency results are mild and allow for multiple breakpoints, ensuring that the framework applies to a broad range of structural change problems.

Simulation studies further demonstrate the effectiveness of the method. We consider scenarios where response distributions are generated from Beta, Gamma, truncated normal, and log-quantile models, with changes in both location and variability. Across these settings, the proposed FPR approach yields accurate estimates of regression functions and breakpoints. Compared with two functional principal component analysis (FPCA)–based competitors, one applied directly to densities and another based on log-quantile density transformations, our method demonstrates improved predictive performance under the Wasserstein metric and appears more sensitive to variability changes. While FPCA-based approaches can sometimes miss variability shifts or introduce estimation distortions, our simulations indicate that FPR more effectively detects both the locations and characteristics of structural changes.

The methodology is further illustrated through applications in demography and economics. In analyzing maternal age-at-childbearing distributions, the estimated breakpoints align with well-documented demographic transitions and policy changes. In a second application to U.S. housing value distributions, the estimated breakpoints correspond to periods of financial crisis and subsequent recovery. In both cases, the fitted regression functions offer interpretable descriptions of how societal events may be reflected in distributional changes. Compared with existing methods, which tend to either oversmooth or miss abrupt variability shifts, FPR captures structural dynamics more faithfully.

This work relates to several strands of current research. Fréchet regression in metric spaces has been developed in recent years, while Wasserstein geometry has emerged as a powerful framework for analyzing distributional data. Structural change detection is well studied for scalar and functional data, but methods tailored to distributional responses remain limited. The proposed FPR method bridges these areas by providing a principled, geometry-preserving approach to detecting structural changes in distributional regression.

In summary, we introduce a framework for regression with distributional responses that accommodates both continuous and discontinuous structural changes. The proposed approach combines theoretical guarantees with empirical validation, providing evidence of improved performance relative to existing alternatives, particularly in settings with variability shifts. Future directions include extending the methodology to high-dimensional predictors and developing scalable algorithms for large datasets. These contributions suggest that Fréchet piecewise regression offers a versatile and interpretable tool for analyzing distributional data subject to structural dynamics.

Nonlinear Trivariate Modal Interval Regression and Its Application to Spatial Precipitation Data

Sai Yao and Yuko Araki

Graduate School of Information Sciences, Tohoku University

1 Introduction

In regression analysis, mean, median, and mode regression have been widely studied. These three statistics represent central tendency. The mean is sensitive to outliers, while the median and mode are more robust. Besides central tendency, the spread of data is also important. Typical measures include the standard deviation, the interquartile range, and the modal interval (MI). MI is robust to skewness and outliers, and provides a direct description of the central region of the data. Research on modal interval regression is limited, since it requires direct estimation of intervals, which is more difficult than point estimation. MI estimation methods based on distributional assumptions lose generality, and nonparametric methods such as kernel density estimation (KDE) often give rough results.

In this study, we extend the nonlinear modal interval regression, originally proposed by Yao et al. (2023) and further developed in Yao et al. (under review), from bivariate data to three variables, and propose Trivariate Modal Interval Regression (TMIR). First, we estimate conditional MIs with KDE and convert the bounds into quantile levels. Next, we use bivariate spline surfaces to estimate the upper and lower bounds at the same time. We also propose a constraint to prevent the estimated bounds from crossing. The problem is written as a convex optimization problem and solved by the alternating direction method of multipliers (ADMM). Finally, we show heatmaps of the upper and lower bounds and the interval width. These plots show how the central region and the degree of concentration change in space.

2 Methodology: Trivariate Modal Interval Regression

2.1 Definition of the Modal Interval

For three variables, let X, Y be the explanatory variables and Z the response variable. Based on the conditional probability density function $f_{Z|X,Y}(z \mid x, y)$, the conditional $100\alpha\%$ MI (for $\alpha \in (0,1)$) is defined as:

$$\mathrm{MI}_{\alpha,Z} \coloneqq [m^{\mathrm{low}}_{\alpha,Z}, m^{\mathrm{up}}_{\alpha,Z}] \coloneqq \underset{[z_{\mathrm{low}}, z_{\mathrm{up}}]}{\arg\min}\ (z_{\mathrm{up}} - z_{\mathrm{low}}), \quad \text{s.t.} \quad \int_{z_{\mathrm{low}}}^{z_{\mathrm{up}}} f_{Z|X,Y}(z \mid x, y)\,\mathrm{d}z = \alpha \tag{1}$$

2.2 Step 1: Pointwise Modal Interval Estimation

The conditional cumulative distribution function can be estimated by KDE:

$$\hat{F}_{Z|x,y}(z) = \frac{\sum_{i=1}^{n} K_{h_1}(x_i - x)\, K_{h_2}(y_i - y)\, \mathcal{I}(z_i \le z)}{\sum_{i=1}^{n} K_{h_1}(x_i - x)\, K_{h_2}(y_i - y)}. \tag{2}$$

where $K_h(\cdot)$ is the Gaussian kernel and h_1, h_2 are bandwidths. At each observation point (x_i, y_i), the conditional MI is obtained as

$$\begin{aligned} \hat{\mathrm{MI}}_{\alpha,Z}(x_i, y_i) &\coloneqq \left[\hat{m}^{\mathrm{low}}_{\alpha,Z}(x_i, y_i), \hat{m}^{\mathrm{up}}_{\alpha,Z}(x_i, y_i)\right] \coloneqq \underset{[\hat{z}_i^{\mathrm{low}}, \hat{z}_i^{\mathrm{up}}]}{\arg\min} \left(\hat{z}_i^{\mathrm{up}} - \hat{z}_i^{\mathrm{low}}\right), \\ &\text{s.t.} \quad \hat{F}_{Z|x_i,y_i}(\hat{z}_i^{\mathrm{up}}) - \hat{F}_{Z|x_i,y_i}(\hat{z}_i^{\mathrm{low}}) \ge \alpha \end{aligned} \tag{3}$$

The upper and lower bounds are then converted into quantile levels:

$$\bar{p}_i = \hat{F}_{Z|x_i,y_i}(\hat{z}_i^{\mathrm{up}}), \qquad \underline{p}_i = \hat{F}_{Z|x_i,y_i}(\hat{z}_i^{\mathrm{low}}).$$

2.3 Step 2: Simultaneous Regression of Modal Interval Bounds

The asymmetric absolute loss

$$\mathcal{J}_p(t) := \begin{cases} pt, & t \geq 0, \\ -(1-p)t, & t < 0 \end{cases}$$

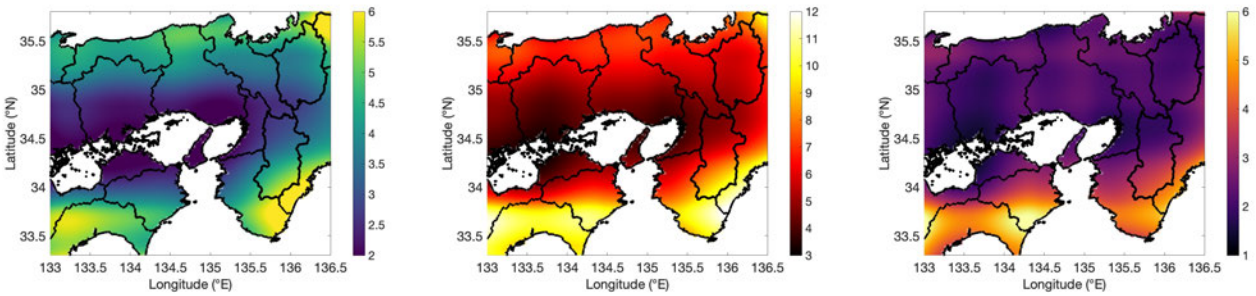
yields quantiles by expectation minimization. Using the quantile levels $\bar{p}_i, \underline{p}_i$ from Step 1, we simultaneously estimate the upper and lower bound functions $\bar{s}(x, y)$ and $\underline{s}(x, y)$ with bivariate splines:

$$\begin{aligned} \underset{\bar{s},\underline{s}\in S^{(2)}_{d,\rho}(\sqcup^{(2)})}{\text{minimize}} & \sum_{i=1}^{n} w_i \mathcal{J}_{\bar{p}_i}\left(z_i - \bar{s}(x_i, y_i)\right) + \lambda \int_\Omega \|\nabla^2 \bar{s}(x, y)\|^2 \mathrm{d}x\mathrm{d}y \\ & + \sum_{i=1}^{n} w_i \mathcal{J}_{\underline{p}_i}\left(z_i - \underline{s}(x_i, y_i)\right) + \lambda \int_\Omega \|\nabla^2 \underline{s}(x, y)\|^2 \mathrm{d}x\mathrm{d}y \\ \text{s.t.} \quad & \forall (x, y) \in \Omega,\ \bar{s}(x, y) \geq \underline{s}(x, y) \end{aligned} \tag{4}$$

Here, $S^{(2)}_{d,\rho}$ denotes the space of bivariate piecewise polynomial splines, w_i are weights based on local density, and λ is a smoothing parameter. In this study, we propose the non-crossing constraint $\bar{s}(x, y) \geq \underline{s}(x, y)$ to ensure that the estimated upper bound is not smaller than the lower bound.

3 Application: Spatial Precipitation Analysis

In this study, we analyzed annual mean daily precipitation data from 2000 to 2011 and estimated the conditional 80% MI in the Osaka and surrounding regions. The lower bound, upper bound, and interval width are shown as heatmaps. From the bounds, we can identify where precipitation is concentrated, and from the interval width, we can see how strongly it is concentrated.



(a) Lower bound (b) Upper bound (c) Interval width

Figure 1: Estimated 80% MI for precipitation in the Osaka and surrounding region (2000–2011).

References

[1] Yao, S., Kitahara, D., Kuroda, H., & Hirabayashi, A. (2023). Modal interval regression based on spline quantile regression. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E106.A(2), 106–123.

[2] Yao, S., Araki, Y., & Iwata, O. (under review). Nonlinear modal interval regression for bivariate data visualization. Manuscript submitted for publication.

Understanding Excess Return: An Investigation into the Applicability of CAPM, Fama-French, and Principal Component Analysis of Equity in Indonesia

Atina Husnaqilati

September 4, 2025

Abstract

This study investigates the performance of the Capital Asset Pricing Model (CAPM), Fama-French Three-Factor Model (FF3M), Fama-French Five-Factor Model (FF5M), and a Principal Component Analysis-enhanced Fama-French Model (FFPCA) in explaining stock returns in the Indonesian equity market. Using data from 100 stocks listed on the Indonesia Stock Exchange (IDX) between 2013 and 2022, the models' explanatory power was assessed through cross-sectional regression analysis. The results demonstrate that FF3M and FF5M consistently outperform the CAPM, with the size (SMB) and value (HML) factors playing critical roles in explaining return patterns. By contrast, the profitability (RMW) and investment (CMA) factors appear less significant in the Indonesian context, suggesting market-specific anomalies.

While the PCA-based Fama-French model (FFPCA) yields lower R^2 values than FF3M and FF5M, it exhibits higher expected returns, capturing broader market trends and extreme variations in returns. This outcome highlights the complementary nature of PCA as a data-driven technique. By aligning the number of principal components with traditional Fama-French factors through a clipping method, the analysis ensures consistency in interpretation while revealing latent structures that conventional regression may overlook.

A key innovation of this study lies in integrating insights from Random Matrix Theory (RMT) into the application of PCA. In financial econometrics, RMT provides a rigorous

framework for distinguishing genuine information from noise in large correlation matrices of asset returns. Eigenvalues that deviate significantly from the bounds predicted by RMT are typically associated with meaningful factors, whereas eigenvalues within the RMT bulk are regarded as statistical noise. By incorporating this perspective, the study enhances the robustness of PCA factor extraction, ensuring that identified components reflect true market dynamics rather than random fluctuations. This methodological addition reinforces the credibility of PCA as a tool for constructing data-driven extensions of the Fama-French model.

The findings underscore the adaptability of asset pricing models in emerging markets such as Indonesia, where unique structural features—including higher volatility, liquidity constraints, and concentrated sectoral exposures—shape return behavior. In particular, the dominance of size and value effects demonstrates that investors in such markets demand compensation for risks not fully captured by market beta alone. Furthermore, the integration of PCA and RMT offers a novel path for refining factor models in environments where traditional theories may fail to capture complexity.

In summary, this study contributes to both theory and practice. It confirms the relevance of multi-factor models in explaining stock returns in Indonesia, while also highlighting the promise of hybrid approaches that merge econometric theory with data-driven techniques such as PCA and Random Matrix Theory. The practical implication is clear: portfolio managers and policymakers can improve risk assessment and investment strategies by relying on models that emphasize size and value effects, complemented by RMT-based PCA to filter noise and identify robust latent factors.

Keywords: CAPM, Fama-French Models, PCA, Random Matrix Theory, Indonesia, Asset Pricing, Emerging Markets.

Practical Monte Carlo Methods Using Piecewise-Deterministic Markov Processes

Charly Andral (Université Paris Dauphine-PSL), Kengo Kamatani (ISM, JST CREST)

In this presentation, we report a new method to automate and accelerate Monte Carlo methods based on Piecewise-Deterministic Markov Processes (PDMPs). The focus is on representative PDMP samplers such as Zig-Zag, Bouncy Particle, Boomerang, and Forward Event-Chain (FEC). Existing automation methods combine maximisation of the event rate $\lambda(t) = \lambda(\phi_t(z_0))$ (via numerical optimisation) with Poisson thinning, but performance strongly depends on prior tuning of the optimisation horizon $t_{\max}$. Our contributions are: (i) automatic adjustment of $t_{\max}$ during runtime by a simple rule, (ii) thinning based on a piecewise-constant upper bound $\Lambda(t)$ constructed on a grid without numerical optimisation (e.g. Brent's method), (iii) a unified implementation strategy applicable to general PDMPs (given flow ϕ, rate λ, jump kernel Q), and (iv) a JAX implementation (`pdmp_jax`).

Framework: The extended state $Z_t = (X_t, V_t)$ evolves deterministically between events according to the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t} Z_t = F(Z_t), \qquad Z_t = \phi_t(Z_0),$$

and events occur according to an inhomogeneous Poisson process with rate $\lambda(Z_t)$. At event times, the state is updated via kernel Q. The next event time τ satisfies

$$\int_0^\tau \lambda\big(\phi_s(Z_0)\big)\,\mathrm{d}s \;=\; E, \qquad E \sim \mathrm{Exp}(1),$$

but since this is not solvable in closed form, thinning with an upper bound $\Lambda(t) \geq \lambda(t)$ is used.

Automatic adjustment of $t_{\max}$: If no event occurs in $[0, t_{\max}]$, set $t_{\max} \leftarrow \alpha_+ \, t_{\max}$; if too many proposed events are rejected, set $t_{\max} \leftarrow t_{\max}/\alpha_-$ ($\alpha_+, \alpha_- > 1$). As $t_{\max}$ is a numerical parameter, not part of the dynamics, the invariant distribution is preserved. This balances the tradeoff between the cost of constructing bounds and wasted thinning rejections.

Grid-based piecewise-constant bounds: Divide $[0, t_{\max}]$ into N intervals $0 = t_0 < \cdots < t_N$. On each interval, evaluate $\lambda(t)$ and use monotonicity or derivative signs to assign a safe constant bound $\Lambda(t_i)$. Then τ can be solved in closed form via

$$\sum_{j=1}^{i-1} \Lambda(t_j)\Delta t_j \leq E < \sum_{j=1}^{i} \Lambda(t_j)\Delta t_j.$$

This avoids optimisation at every step. Grid evaluation is vectorisable and robust to non-convex or multimodal λ.

Application to general PDMPs: For Zig-Zag ($\lambda_i(t) = [v_i\,\partial_i U(X_t)]_+$), Bouncy Particle ($\lambda_{\mathrm{BPS}}(t) = [\nabla U(X_t)\cdot V_t]_+ + \lambda_{\mathrm{ref}}$), Boomerang, FEC, etc., different strategies (componentwise, vectorized, signed) are combined to build safe bounds.

Correctness guarantee: If a rare violation $\Lambda(\tau) < \lambda(\tau)$ is detected, $t_{\max}$ is reduced and the bound rebuilt, ensuring exactness of the procedure.

Numerical behaviour (summary): Across anisotropic/multimodal distributions, locally mixing targets, and high-dimensional nearly-convex potentials, combining grid-based bounds with adaptive $t_{\max}$ consistently improved (a) thinning acceptance rate, (b) reduced optimisation cost, and (c) overall runtime.

Installation and minimal example

JAX and `pdmp-jax` installation:

```
pip install -U jax
pip install pdmp-jax
```

Example (Zig-Zag, 3D standard normal):

```
import jax, jax.numpy as jnp
jax.config.update("jax_enable_x64", True) # set BEFORE using pdmp_jax

import pdmp_jax as pdmp

U = lambda x: 0.5 * jnp.sum(x**2)
grad_U = jax.grad(U)
x0 = jnp.zeros(3); v0 = jnp.array([1., -1., 1.])

zz = pdmp.ZigZag(dim=3, grad_U=grad_U, grid_size=10)
samples = zz.sample(N_sk=20000, N_samples=5000,
               xinit=x0, vinit=v0, seed=0, verbose=False)
print("mean ~ 0:", samples.mean(0))
```

Conclusion: By combining adaptive horizon tuning with grid-based piecewise-constant bounds, PDMP samplers can be implemented robustly and efficiently without prior tuning. The same framework applies to multiple PDMPs, and the `pdmp_jax` implementation demonstrates practical benefits.

This talk is based on "Andral, C., & Kamatani, K. (2024). *Automated Techniques for Efficient Sampling of Piecewise-Deterministic Markov Processes.* arXiv preprint arXiv:2408.03682."

Kengo Kamatani (Institute of Statistical Mathematics)
10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
E-mail address: kamatani@ism.ac.jp

Charly Andral (Université Paris Dauphine–PSL)
Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France