＜研究レポート＞

# どれが原因？どれが結果？ 教師あり学習による時系列因果推論

京都大学大学院・情報学研究科
（NTTコミュニケーション科学基礎研究所・協創情報研究部）
**近原　鷹一**

時々刻々と変化する時系列から変数間の因果関係の有無・方向を明らかにする因果推論は，多くの応用が期待できる．本研究では因果関係が既知の様々なデータを用いて，因果関係を高精度に推定するための機械学習技術を提案した．

## どれが原因？どれが結果？

時々刻々と変化する時系列データから，変数間の因果関係を発見することは時系列解析の重要なタスクの一つである．例えば，研究開発（R&D）に対する投資額$X$は売上額$Y$に影響を与えるが，$Y$は$X$に影響を与えないという因果関係($X\rightarrow Y$)は企業における重要な意思決定の一助となる．自然科学の例で言えば，時系列マイクロアレイデータから遺伝子間の制御関係を明らかにすることはバイオインフォマティクス及び創薬研究における最も重要なタスクの一つである．本研究では，Granger causality [1]と呼ばれる因果関係の定義に基づき，時系列データから高精度に因果関係の有無・方向を発見するための推定技術を提案する．

## 既存技術の問題点

数ある因果関係の定義の中でよく用いられるのがGranger causality [1]である．これは，「変数$X$の過去の値が変数$Y$の未来の値を予測するのに役立つ」ならば因果関係は$X\rightarrow Y$であるとするものである．Granger causalityの有無・方向を推定するためには，予測の良し悪しを評価するための予測式（自己回帰モデル）が必要である．このとき，予測式がデータにうまくあてはまるものであれば正しく因果関係の有無・方向を推定できるが，そうでない場合，誤った推定結果になる．一般に，個々の時系列データに対して適切な予測式を選択するためには，時系列の長さやノイズの性質など，種々の情報を考慮する必要があり，データ解析の深い専門知識が要求される．本研究では，このような問題を解決するため，予測式の選択が不要でデータ解析の専門知識を必要としない，Granger causality推定の新たなアプローチを提案する．

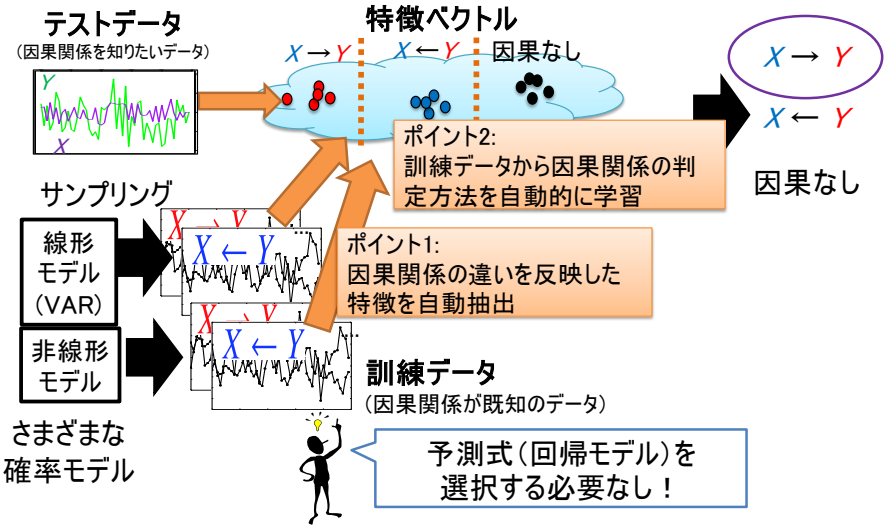


図1: 提案した時系列因果推論技術の概要

## 提案技術の概要

提案技術[2, 3]の概要を図1に示す．提案技術では，予測式の選択を不要とするため，因果関係を推定したい時系列データ(テストデータと呼ぶ)とは別に，様々な性質を持つ時系列データで，かつ因果関係の有無・方向が既知であるような訓練データを大量に用いることを考える．ここで，因果関係が既知の実データを用意することは一般に困難であるが，線形モデル・非線形モデルなどから，人工的に生成した時系列データを用意するのは容易である．提案技術では，こうした訓練データを用いて，(1) 因果関係が$X\rightarrow Y$, $X\leftarrow Y$, 因果なしの時系列データにどのような特徴があるかを抽出し，(2) 抽出した特徴に基づいて分類器を学習することで，因果関係が未知のテストデータからその因果関係の有無・方向を判定する問題を教師あり学習の問題(分類問題)として解く．以下では．(1)の特徴抽出の方法について，述べる．

## 因果関係を反映した特徴の抽出

時系列データからうまく特徴を抽出し，Granger causalityの有無・方向を推定することを考える．図2に示すように，Granger causalityを判断するうえで重要な「過去の変数値が予測に役立つか否か」は，数学的には「過去の変数値($S_X$, $S_Y$)を与えたときに，未来の時刻における変数値の(条件付き)分布が異なるか否か」によって定義される．提案技術では，このような(条件付き)分布を，カーネル平均によって再生核ヒルベルト空間中の点として表し，その点の間の距離(MMD [4])を，分布間の差異を表す特徴とすることで，Granger causalityの有無・方向を反映した特徴をうまく設計した．このような特徴抽出の過程で，予測式の選択などは一切不要であるため，提案技術ではデータ分析の専門知識は何ら必要としない．

## 評価実験

因果関係が既知であるような人工データ・実データを用いて，提案技術による推定精度を評価した．人工データ実験(図3)では，因果関係を推定したいテストデータが線形である場合には線形な予測式(VAR)によって，非線形である場合には非線形な予測式(GAM, kernel)によって高い精度を達成したが，異なる予測式を選んだ場合は推定精度が低くなった．一方提案技術では，因果関係を推定したいテストデータとは別に，因果関係が既知の様々な訓練データを用いているため，テストデータの時系列長が不十分な場合においても，線形・非線形の両方の場合において十分高い推定精度を達成した．

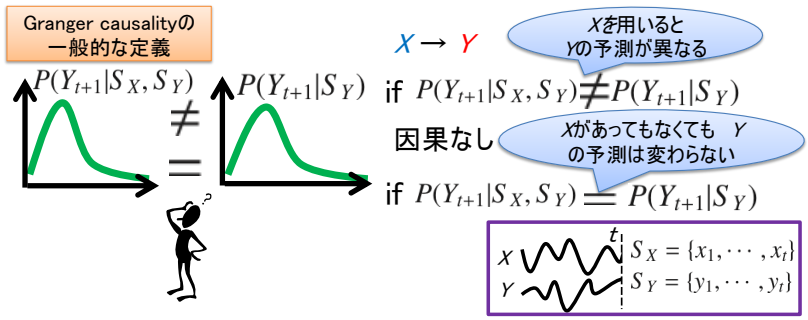


図2: Granger causalityの一般的な定義

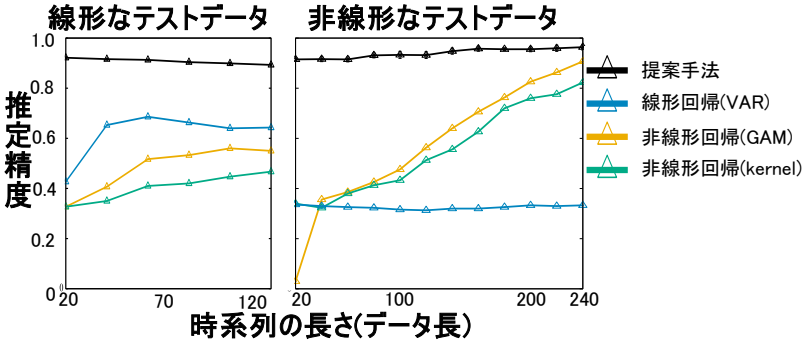


図3: 人工データを用いた推定精度の評価結果

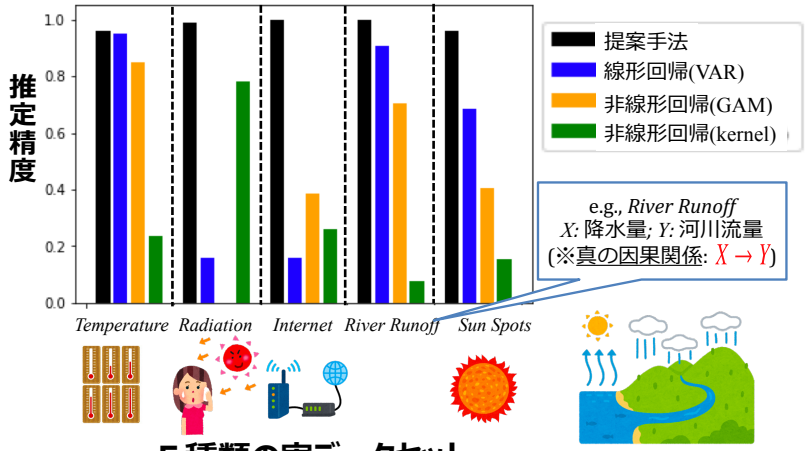


図4: 実データを用いた推定精度の評価結果

ここで用いている訓練データは既に述べた通り，人工的に生成したものであるが，興味深いことにこの傾向は因果関係を推定したいテストデータが実データの場合においても同様であった(図4)．これらの結果は，提案技術がGranger causalityの有無・方向に基づいてうまく特徴抽出できていることを示唆している．

## おわりに

本研究では，時系列データから高精度に因果関係の有無・方向を推定するための機械学習技術を提案した．複雑な様相を示す時系列から因果関係を正しく推定することは一般に困難であり，既存技術ではデータに対して強い仮定を置くことが多い．より幅広い時系列データに対して適用できる実用的な因果推論技術を提案することは今後の課題である．

＜参考文献＞


[1] C. W. Granger. Investigating causal relations by econometric models and cross-spectral methods. Journal of the Econometric Society, pages 424–438, 1969.

[2] Y. Chikahara and A. Fujino. Causal inference in time series via supervised learning. In IJCAI, pages 2042–2048, 2018.

[3] 近原鷹一，藤野昭典, "教師あり学習に基づくGranger causalityの推定," 情報処理学会論文誌：数理モデル化と応用(TOM), Vol.11, No.3, 58-73, 2018.

[4] A. Gretton, K. M. Borgwardt, M. Rasch, B. Sch¨olkopf, and A. J. Smola. A kernel method for the two-sample-problem. In NeurIPS, pages 513–520, 2007.