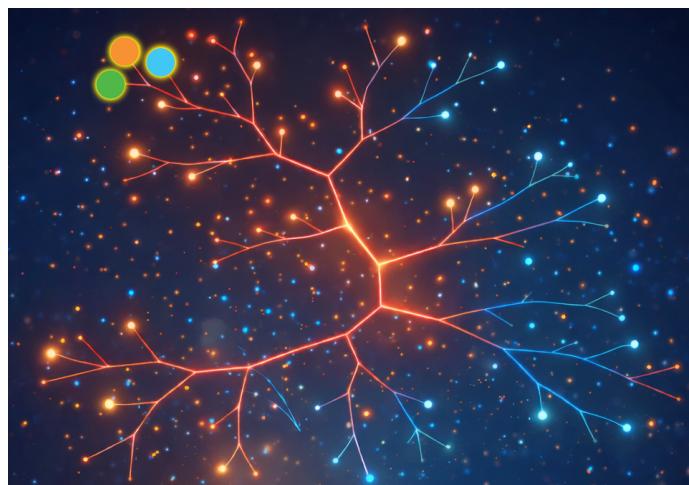


## テンソルツリーによる生成モデル構築のスキーム

——生物系統樹など、因果関係の解明が可能——

## 発表のポイント

- ◆生成モデリングの代替パラダイムとして、対象確率分布をモデル化する単層非負値適応テンソルツリー (NATT) の構造最適化スキームを提案した。
- ◆本手法では、ネットワーク構造そのものが確率のネットワークという意味を持つため解釈が容易である。
- ◆社会のさまざまな面に応用されつつある人工知能や機械学習のベースになる生成モデリングの枠組みとして、現在主流になっているものとは大きく異なる方法となっており、特に多数の要因が絡み合っているような因果関係の解明などへの応用が期待される。



テンソルツリー生成モデルによる因果構造の推定

## 発表内容

東京大学物性研究所 赤松克哉特任研究員(研究当時:大学院理学系研究科博士後期課程)、同 川島直輝教授、京都大学大学院情報学研究科 原田健自助教、東京大学大学院理学系研究科 大久保毅特任准教授の共同研究チームは、生成モデリングの代替パラダイムとして、対象確率分布をモデル化する単層非負値適応テンソルツリー (NATT) の構造最適化スキームを提案した。

NATTスキームは、テンソルツリー(注1)形式で表現できる確率分布関数のうち、与えられたサンプルの集合に最も適合するものを求めるためのものである。個々のサンプルは多数の確率変数(特徴量)の観測値のリストであり、たとえば画像認識の問題であれば、一枚の画像が一つのサンプルに相当する。一般的な生成モデルと同様に、NATTにおいても「もっともらしい」サンプルを新たに生成することが可能である。しかし、本研究で注目したのは、得られたテン

ソルツリーの構造が、対象となるデータセットの情報論的な構造を反映したものになるという点であり、これを利用して、確率変数間の因果関係などを推定することができる。

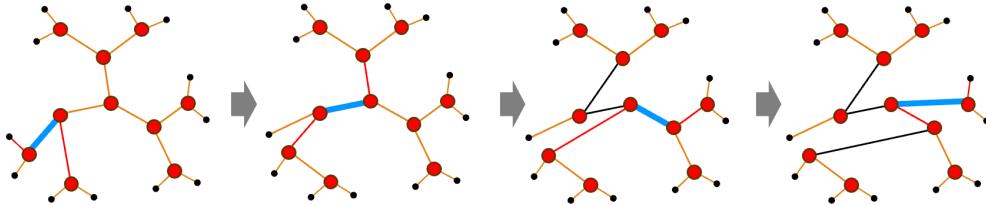


図1:テンソルツリーの構造最適化。ツリー内の枝を巡回的に、注目されている枝が担う相互情報量が最小になるように繋ぎ変えていくことで、全体として最適な構造を求める。

NATTでは、ツリー内の枝を巡回的に調べて、その枝が媒介する相互情報量が最小になるように枝の繋ぎ変えを行うことで、ツリー構造を自動的に最適化している。また、このツリー自体が、条件付き確率で多くの確率変数を結合した確率ネットワークを表現している。このため、探索で得られたツリー構造が、対象となっている変数の間の因果関係として解釈できる。ベンチマークとして、人工的なデータと実データの両方についてNATTの性能を検証した。人工的なデータとしては、ビット演算ネットワークの構造の推定や、サンプルに含まれない変数を持つランダムベイジアンネットワーク（注2）で非提示領域の内部構造の推定などを行った。実データの例としては、犬や猫の仲間のミトコンドリアDNA配列からそれらの系統樹推定のもととなるクラドグラムを構築した。この結果、現在知られているクラドグラム（注3）とほぼ正確に一致するツリー構造が得られた。

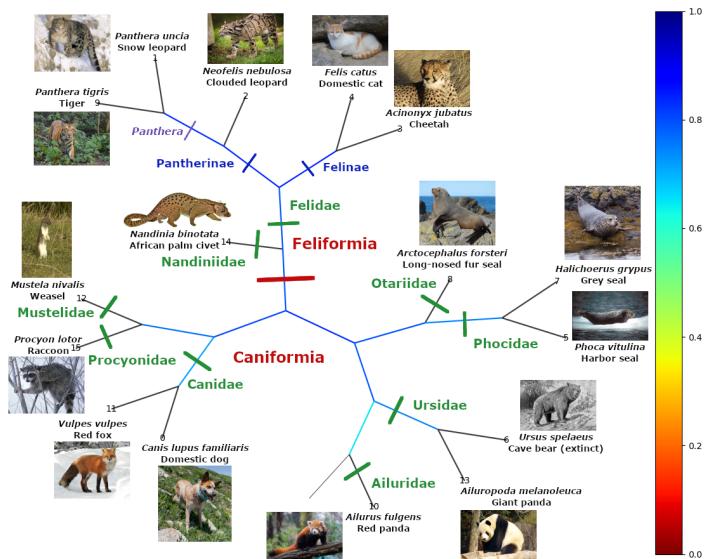


図2:16種の食肉目動物のミトコンドリアDNAに含まれる1140塩基対の情報に基づいてNATTが生成したクラドグラム。レッサーパンダ(Ailuridae)とイヌ科(Canidae)の分岐位置を除くと一般に正しいとされている構造に一致している。線の色は線が担っている相互情報量の大きさを表している。

昨今、社会的に大きなインパクトを与えている人工知能の実態である生成モデルはほとんどの場合ニューラルネットワークがベースとして使われているが、対象の情報論的な構造を反映

したネットワーク構造の最適化が容易な、テンソルツリーの有用性が認識され始めている。テンソルツリーを応用した先行研究では、ツリーが波動関数を表現するという形式を取っていたため、部分的なネットワークを切り出しても、それに対して確率的な解釈ができなかった。これに対して、本研究では、ツリー 자체が確率分布関数を表すように設計されているため、ネットワークは局所的にも各変数間の確率的関係を与えるものとなっている。本研究で提案しているテンソルツリーの構造最適化に基づく生成モデル構築のスキームは、社会のさまざまな面に応用されつつある人工知能や機械学習のベースになる生成モデリングの枠組みとして、現在主流になっているものとは大きく異なる方法となっており、特に多数の要因が複雑に絡み合っているような因果関係の解明などへの応用が期待される。

#### ○関連情報：

プレスリリース：テンソルネットワークによる生成モデル 一株式騰落パターンから相関構造が発現— (2025/4/7)

<https://www.kyoto-u.ac.jp/ja/research-news/2025-04-07>

#### 発表者・研究者等情報

東京大学

物性研究所

赤松 克哉 特任研究員（研究当時：東京大学大学院理学系研究科博士後期課程）

川島 直輝 教授

大学院理学系研究科

大久保 育 特任准教授

京都大学大学院情報学研究科

原田 健自 助教

#### 論文情報

雑誌名：Machine Learning: Science and Technology

題名：Plastic tensor networks for interpretable generative modeling

著者名：Katsuya O. Akamatsu\*, Kenji Harada, Tsuyoshi Okubo and Naoki Kawashima

DOI:10.1088/2632-2153/ae3048

URL: <https://doi.org/10.1088/2632-2153/ae3048>

#### 研究助成

本研究は、京都大学とトヨタ自動車の共同研究プロジェクト「モビリティ基盤数理の研究」、グローバルサイエンス国際卓越 大学院コース、科学研究費補助金（課題番号 20K03766, 21H05182, 21H05191, 22K18682, 23H01092, 23H03818, 24K06886）、国立研究開発法人科学技術振興機構の共創の場形成支援プログラム(COI-NEXT)「量子ソフトウェアと HPC・シミュレーション技術の共創によるサステイナブル AI 研究拠点 (JST Grant No. JPMJPF2221)、東京大学物性研究所共同利用スーパーコンピュータの支援により実施されました。

## 用語解説

(注 1) テンソルツリー :

テンソルネットワークのうち、ネットワークがツリー状になっているもの。(つまり、閉じたループを含まないもの。) テンソルネットワークとは、グラフ上の各ノードにテンソルが定義されており、ノード間をつなぐ線がテンソルの添字を表すもの。テンソルネットワークは、近年、物性理論の分野において多体問題の新しい理論や計算手法の基礎として研究が進んでいる。

(注 2) ランダムベイジアンネットワーク :

多くの確率変数が互いに因果関係で結ばれたネットワークをランダムに生成したもの。

(注 3) クラドグラム :

生物同士の進化的な近縁関係を表した樹形図で、系統樹に似ているが、時間的な情報を含んでいない。

## 問合せ先

(研究内容については発表者にお問合せください)

東京大学物性研究所

教授 川島 直輝 (かわしま なおき)

Tel : 04-7136-3260 E-mail : kawashima@issp.u-tokyo.ac.jp

東京大学物性研究所 広報室

Tel : 04-7136-3207 E-mail : press@issp.u-tokyo.ac.jp

東京大学大学院理学系研究科・理学部 広報室

Tel : 03-5841-8856 E-mail : media.s@gs.mail.u-tokyo.ac.jp

京都大学 広報室国際広報班

Tel : 075-753-5729 Fax : 075-753-2094 E-mail : comms@mail2.adm.kyoto-u.ac.jp