

論文内容の要旨

論文題目

Representation Learning with Neural Networks for Structural World
(ニューラルネットワークによる構造的外界の表現学習)

氏名 長野 祥大

深層学習は人工ニューラルネットワーク(NN)で機械学習の問題をモデリングし、勾配ベースの手法を用いてモデルのパラメーターを訓練する技術の総称である。人工 NN の研究は生物の神経回路を模した知的情報処理を目指して始まった。現在では、人工 NN は非線形関数の近似器として広く利用されており、背後の理論的側面も徐々に解明されつつある。近年、大規模な計算機と大量のデータセットを背景に、急速に深層学習の研究が進展し様々なタスクで人間を上回る知的情報処理が可能となった。特に、微分可能な計算グラフという共通基盤は、これまで様々な分野でドメインごとに培われてきた技術を相互に交換可能にした。

深層学習技術は幅広いタスクで利用されているが、その中でも確率分布の学習を行う深層生成モデルは重要である。機械学習は現実世界の課題を解決するという工学的な実用性を追求する学問であるが、我々人間を含む知的システムの情報処理のモデルとして捉えることも可能である。ラベルを用いずに学習する生成モデルはこのような目的に即している。特に、深層生成モデルの研究の進展は非線形で複雑な確率分布のモデリングを可能にし、自然画像や音声などの高品質な生成を実現した。

このように、非線形関数の近似器として利用されるニューラルネットワークであるが、実用上はデータセットの背後に何らかの構造を仮定することで、探索すべき空間を制約し性能が向上する。例えば高次元のデータセットに関して次元間の相関を仮定したネットワーク構造は多数研究されてきた。自然画像データセットに用いられる二次元畳み込みは画像データの隣接するピクセルは相関する・データ内に含

まれる物体は座標移動に対する不変性を持つという経験的事実を反映する。次元畳み込みは音声などの時間的に変動するデータの時間相関を利用する。更に畳み込みだけでは捉えられないような長期の相関のための再帰結合や Attention, 周期的な長距離相関を捉えるための Dilated Convolution などが提案されている。

一方で、1つのデータ点の中での次元間の構造の他に、データとデータの関係という意味での構造も考えることができる。深層学習を含め、多くの機械学習手法ではデータ分布について i.i.d.の仮定を置くが、現実の世界には特定の相関構造が多々見られる。例えばデータセットが所属するラベルを識別するタスクは典型的に機械学習が対象とする問題であるが、人間が認識する物事概念には階層構造がある。このような明示的な状況以外でも、データを収集する時間や環境が異なれば、データ分布に影響を与えるだろう。このように、特定の構造を有するデータセットに対して、NN を用いた表現学習のアプローチは近年注目を集めている。例えばクラスター構造など想定している構造に対して有用な確率分布を利用する、潜在空間に特定の構造を仮定する、目的関数を工夫することで潜在空間に有益な情報を抽出しやすくするなどのアプローチなどが存在する。本論文では、深層生成モデルの代表的なモデルである VAE を対象に、クラスター構造・階層構造・局所構造が仮定できるデータセットに対する推論と学習について述べる。

第3章では、クラスター構造を持つデータセットに対して、深層生成モデルの推論のダイナミクスを数値的に解析した。VAE はエンコーダーとデコーダーを用いて、画像データからのノイズ除去が可能である。ノイズを含む画像をエンコーダーで潜在空間に射影し、デコーダーでデータ空間に戻すことでノイズが除去されることに加え、出力された画像を再度入力として用いこの操作を繰り返す逐次的推論によって徐々にノイズが減少することが知られている。しかし、この逐次的な推論の効果は経験的な報告にとどまる。本章では、このような現象のメカニズムに迫るため、VAE の推論過程を典型的なデータセットを用いて数値的に解析した。

クラスター構造は我々が関心を持つデータセットで幅広く見られる構造である。本章ではクラスター構造を有するデータセットとして自然画像データセットの MNIST を用いた。MNIST で訓練した VAE で逐次的推論を行った際、潜在空間における活動パターンは訓練データのクラスターの中心に急速に接近することを示した。また、入力に加えるノイズ量に依存して潜在空間上のダイナミクスは変化することを明らかにした。これらの結果は逐次的推論のダイナミクスがデータセットの背後にある構造を反映することを示唆する。これまで報告されてきた逐次的推論の

効果は、このクラスター中心への接近に起因すると考えられ、同一のモデルにおいても、推論を繰り返す回数によって出力の抽象度を操作できることが示唆された。

第4章では、階層的な構造を持つデータセットに対して有用な確率分布と、これを利用した表現学習手法の提案を行った。近年、自然言語処理の分野において木構造のような階層性のあるデータの双曲空間への埋め込みが重点的に研究されている。

双曲空間への階層的なデータの埋め込みの有用性が報告されている一方で、これまでの提案はすべて決定論的な埋め込みでデータの不確実性を表現できていなかった。本章では曲がった空間である双曲空間上で勾配ベースの学習に利用しやすい確率分布を提案した。双曲空間に代表される、非ユークリッド空間上での確率分布の構成は、大きく分けて対象の空間上に直接分布を定義するアプローチ(Riemannian分布)と、接空間からの射影によって定義するアプローチ(Wrapped分布)がある。前者はガウス分布に相当する分布の最大エントロピー性が多様体上で満たされる一方で、複雑な分布を用いたい場合に尤度が閉じる保証がなく、また、その分布からのサンプリングは一般に困難である。対して、射影ベースのアプローチでは、接空間で定義した任意の確率分布を射影することが可能であり、容易にサンプリングできる。また、射影を通して勾配を伝搬できるため深層学習に適している。以上の理由から、本研究では双曲空間上の確率分布を射影ベースで構成し、表現学習に適用した。数値実験としてVAEを用いた自然画像データセットの学習と、自然言語の単語埋め込みであるWord Embeddingタスクの双方で双曲空間の有用性を検証した。特に、階層構造が仮定できる画像データセットを持ちいた実験では、双曲空間上のノルムとデータセット内の階層構造が関連した。

第5章では、局所的な構造を持つデータセットに対して有用な深層生成モデルの学習法を提案する。機械学習が対象とする多くのデータセットは物理的法則に支配されているため、高次元な観測も少数の操作変数を使って表されるという仮定は妥当である。この操作変数の抽出を目指すdisentanglement representationは近年注目を集め、数多くの研究が行われている。一方でそのような表現はデータセットやモデルへのなんらかの帰納バイアスなしには実現できないことが明らかになっているが、帰納バイアスの具体的形式はこれまで積極的に議論されていない。

我々が関心を持つ操作変数は大抵の場合典型的なスケールを持つ。例えば顔画像のデータセットにおいて、個々人の顔の違いは表情の違いに比べて変化のスケールが大きく、表情の違いは個人ごとに変わりうる。つまり、大域的な操作変数に局所的な操作変数が依存するような局所構造が存在する。このような局所構造を持つデ

ータセットに対して、メタ学習は効果的なアプローチである。Model-Agnostic Meta-Learning (MAML)に代表されるメタ学習は、データセットにタスクラベルが与えられている状況で、未知のタスクに早く適用できるようにモデルを訓練する。本章では、データセットに含まれるそれぞれの局所構造をメタ学習のタスクと捉え、近傍関係を補助的なラベルとして与えることで、各局所構造間で転用可能な表現が抽出できることを示す。特定データの局所近傍ごとにタスク固有のパラメーターを持つ生成モデルとして、MAMLの階層ベイズ的解釈を用いたLocal VAEを提案する。潜在表現の獲得の文脈で用いられるデータセットである3D Shapes, Cars3D, SmallNORBを用いた実験で、Local VAEによって生成画像・潜在表現ともに品質が向上することを示した。

本研究では、VAEを中心技術としてデータセットのインスタンス間の構造を捉える枠組みについて広く研究を行った。現在の深層学習技術は大規模なデータセットと大量の計算資源のもとで様々なタスクで人間を上回る性能を獲得した。一方で、環境の変動の吸収や少ない事例からの汎化などの観点から見れば、現時点では人間に劣っているといえる。工学的応用にとどまらず、知性を明らかにするという理学的観点からも我々はこれらの問題に対処する必要がある。特にデータ同士の構造は人間が日々の意思決定で暗黙的に利用している普遍的知識である。効率的な情報処理システムの実現のためには、そのシステムが接する外界についての考察が必要不可欠であると考え、本論文の提案がこのような今後の知能研究の一端となることを期待する。