

審査の結果の要旨

氏 名 大野 健太

IT技術の急速な進歩に伴い、機械学習は大量データから有用な知見を得るための方法論としてその重要性を増している。効率的な学習を実現するためには、先験的知見をモデルに反映させた「帰納バイアス」が重要な役割を果たす。特に、深層学習においてはそのネットワーク構造の工夫に帰納バイアスを反映させることで各種パターン認識タスクにおいて高い性能を実現させている。しかし、このネットワーク構造の工夫が精度に与える影響については、あくまで経験的な評価によってのみ調べられており、理論的な解明は進んでいない。本論文では、深層学習の構造が引き起こす帰納バイアスを理論的に解明することを目標とし、特にResNet型の畳み込みニューラルネットワーク (CNN)、グラフニューラルネットワーク (GNN)、スキップ接続を有するマルチスケールGNNに焦点を当て、その理論解析とともに数値実験によって理論および提案手法の有効性を検証している。

本論文は「Analysis of Deep Learning from the Viewpoint of Model Structures」(モデル構造的見地からの深層学習解析) と題し5章からなる。論文は大きく分けて三つのパートに分かれており、それぞれ「ResNet型CNNの関数近似・推定理論」「GNNの過平滑化」「マルチスケールGNNの勾配ブースティング法的観点による解析」に主眼を置いている。

第1章「Introduction」では、深層学習における主要なネットワーク構造とその理論解析に関する研究背景と、本論文の構成および主結果の概要を与えている。

第2章「Background」では、本論文で扱うネットワーク構造および関連する話題の概観と統計的学習理論の基本事項を与えている。

第3章「Approximation and Non-parametric Estimation Analysis of ResNet-type Convolutional Neural Networks」では、ResNet型の構造に着目し、ResNet型CNNがいかなる帰納バイアスを生むかを理論的に解析している。ResNetは主に画像認識において標準的な構造であり、実用上大きな成功を収めている。本研究では、真の関数がブロックスパースな全結合ネットワーク (FNN) で効率的に近似できるのであれば、ResNet型のネットワークにおいても同様の近似性能を達成できることを示している。ブロックスパースなネットワークにより効率的な近似が実現できる関数クラスとしてHölderクラス、

Barronクラス, Besovクラスといった理論上有用な関数クラスが含まれるが, 理論解析においてはエッジがスパースであるという実用上不自然な制約が必要であった. しかし, ResNet型のネットワークを考えることでそのようなスパース性を排除することができ, 理論的に最適な誤差をスパース性の制約なしで達成できることが示された. また, 近似精度の理論から, 有限データにおける推定誤差の収束レートも導出し, 上記の主要な関数クラスにおいてほぼミニマックス最適な収束レートを達成することが示された.

第4章「Over-smoothing of Non-linear Graph Neural Networks」では, グラフ型データに対する深層学習手法であるGNNを考察し, その中でも重要なグラフ畳み込みニューラルネットワーク (GCN) の帰納バイアスについて解析している. GCNは各ノードに付与された特徴量を線形変換と非線形活性化関数による変換する「特徴量変換」の操作と, 周辺ノードの特徴量を平均して自分の特徴量に足しこむ「集約」の操作を繰り返し適用する. GCNはGNN手法の中でも高い性能を出しているが, 通常の高層学習と違い層を深くするにつれて大幅に性能が落ちる特性があった. この一つの理由として集約操作によりノード特徴量が急速に情報を失う過平滑化があると考えられている. 過平滑化がおきることは線形なGCNに対しては理論的に示されていたが, 非線形活性化関数ありのネットワークについては示されていなかった. 本研究では, ReLU活性化関数を用いている限り, この過平滑は非線形活性化関数を用いても起きることを示し, 無情報な特徴量の部分空間への収束レートがグラフラプラシアンの特値ギャップにより決まることを示した. また, Erdős-Rényiモデルにおける特値ギャップを解析することでノード数が多く密なグラフでは過平滑化が起きやすいことを示した. 理論の有効性は種々の数値実験により実証されている.

第5章「Optimization and Generalization Analysis of Multi-scale Graph Neural Networks through Gradient Boosting」では, 第4章で解析した過平滑化を防ぐ方法としてマルチスケールGNNを考察している. GNNを各層で勾配ブースティング法によって弱学習機を加える学習方法としてとらえなおし, その観点に基づいた勾配ブースティング法を用いた学習方法を提案し, そのアルゴリズム的な収束解析と汎化誤差の解析を与えている. マルチスケールGNNは, 各層で集約操作を行いグラフの大域的情報を得るが, 各中間層からの出力を直接出力層へスキップ接続することで過平滑化する前のノード特徴量の情報を利用した推論を可能とする. これは見方を変えれば, 集約操作を複数回適用した後の各層の特徴量を用いた弱学習機を足し合わせた勾配ブースティング法とみなせる. 各層で弱学習可能条件が満たされていれば層を増やすごとに単調に訓練誤差を減少させることができ, また学習されたネットワークのパラメータがある条件を満たしていれば層を増やしても汎化ギャップが悪くならないことが示される. このことより, マルチスケールGNNのように複数のスケールの情報を組合せて用いることで過平滑を防ぎつつ, 適切な汎化性能を保證することができることが示唆される. 数値実験により提案法の性能を評価し, 実用上も有用な手法であることが実証されている.

以上を要するに，本論文では深層学習の構造がもたらす帰納バイアスに着目し，種々の応用上有用なモデルにおいてそのネットワーク構造がもたらす帰納バイアスを理論的に解明した．また，これらの理論解析は，十分な数値実験によってその有効性が検証されている．応用上有用なモデルへの理論的解析の付与は深層学習手法の科学的発展という意味でも重要な貢献である．帰納バイアスの理論解明という機械学習の本質的な問題に取り組み，理論と実用の両面へ有用な知見をもたらし，数理情報学の発展に大きく寄与している．

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる．