



深層学習によるデータ固有のフラクタル構造などへの適応を証明

1. 発表者

中田 竜明 (アメリカ・ラトガース大学 統計学科 博士課程)

今泉 允聡 (東京大学大学院総合文化研究科 附属先進科学研究機構 准教授

／同研究科 広域科学専攻 准教授)

2. 発表のポイント

- ◆深層学習が、データ中に暗黙的に存在するフラクタル構造などの低次元構造に適応し、その結果高い予測精度を達成することを数学的に証明した。
- ◆シンプルな低次元構造に適応できる深層学習以外の手法もあるが、深層学習ならばフラクタルなどの滑らかではない構造に適応できることを初めて明らかにした。
- ◆データ固有の構造に即したニューラルネットワーク設計法などへの拡張が期待される。

3. 発表概要

アメリカ・ラトガース大学統計学科博士課程の中田竜明氏、東京大学大学院総合文化研究科附属先進科学研究機構の今泉允聡准教授による研究チームは、深層学習による回帰問題（教師あり学習の一種）の汎化誤差（注1）を数学的に解析しました。その結果、深層学習がデータの中に隠れているフラクタル（注2, 図1）などの低次元構造に適応し、その次元の大きさに応じて誤差を減少させられることを数学的に証明しました。なお、深層学習以外のデータ解析手法も、データのシンプルな低次元構造に適応できることが既に証明されていますが、本研究が構築した理論では、深層学習ならば通常の構造だけでなく、フラクタル構造などの複雑な形状にも適応できることが明らかになりました（図2）。

本研究成果は、国際科学雑誌「Journal of Machine Learning Research」にオンライン掲載されました。

4. 発表内容

<研究の背景>

深層学習は高い性能を発揮するデータ解析法ですが、その原理には不明な点が多く残っています。この原理を明らかにする上で重要なのが、データの持つ「隠れた低次元構造」です。深層学習が扱うデータの多くは、表面上からは発見できない低次元構造を隠し持っています。これは「低次元多様体仮説」などの学説によって解釈され、現代的なデータ解析手法の多くはこの低次元構造を抽出し、これに適応した特徴量設計などを通して予測精度を向上させていることが経験的に知られています。

この低次元構造への適応ですが、理論的にはカーネル法やk平均法（注3）などの深層学習以

外のデータ解析手法も、可微分多様体^(注5)などのシンプルな低次元構造には適応できることが過去の研究により証明されています。しかし深層学習は、これらの方法より高い性能を実験的に発揮するため、この実験結果を説明するにはより発展的な理論が必要でした。

<研究内容>

アメリカ・ラトガース大学統計学科博士課程の中田竜明氏、東京大学大学院総合文化研究科附属先進科学研究機構の今泉允聡准教授による研究チームは、深層学習による回帰問題（教師あり学習の一種）の汎化誤差を数学的に解析し、データ数の増加に応じた誤差の減少スピードを解析しました。

結果として、この誤差の減少スピードは、データの表面的な次元には影響されず、代わりにデータ固有の低次元構造の次元（固有次元）によって決まっていることが数学的に証明されました。すなわち、深層学習が高解像度画像のような高次元データを分析する場合、このデータに隠れている構造が低次元であるなら、深層学習はデータが低次元の場合と同様の高い精度を発揮することが証明されました。

さらに、深層学習はこの低次元構造がフラクタルなどの複雑な形状を持っている場合でも、それに応じた固有次元に適応し、誤差を早く減少させることが明らかになりました。これは、回帰の設定では深層学習以外の方法では示されていないことです。よって、深層学習が低次元構造への適応という文脈で、他手法に対する優位性を持つ可能性が示唆されました。

<社会的影響>

深層学習は優れたデータ解析手法ですが、計算コストが大きいなどの実用上の問題点が残っています。本研究の成果は、データの構造に適応したコンパクトなニューラルネットワーク設計を行うなどの、計算コストを削減する運用法の開発に役立つと考えられます。

5. 発表雑誌

雑誌名：*Journal of Machine Learning Research* (2020年9月20日掲載)

論文タイトル：“Adaptive Approximation and Generalization of Deep Neural Network with Intrinsic Dimensionality”

著者：Ryumei Nakada, Masaaki Imaizumi

URL: <https://www.jmlr.org/papers/v21/20-002.html>

6. 問い合わせ先

東京大学大学院総合文化研究科附属先進科学研究機構 准教授

理化学研究所革新知能統合研究センター 客員研究員

科学技術振興機構 さきがけ研究者（兼任）

情報・システム研究機構統計数理研究所 客員准教授

今泉 允聡 (いまいずみ まさあき)

本研究成果は、国立研究開発法人科学技術振興機構 (JST) の戦略的創造研究推進事業 さきがけ「新しい社会システムデザインに向けた情報基盤技術の創出」研究領域 (研究総括：黒橋禎夫 京都大学大学院 情報学研究科 教授) における「深層学習の高速化にむけた適応ネットワークの数学的発見と学習法開発」(研究代表者：今泉允聡 准教授) (JPMJPR1852) の一環として得られました。

7. 用語解説

(注1) 汎化誤差：

データをもとに構成した予測手法の損失 (誤りの大きさ) の期待値。データ解析手法の予測性能の尺度の一つとして用いられる。

(注2) フラクタル：

幾何的な形状を表現する概念のクラス。自己相似形をもつものが多い。代表的な例はカントール集合やコッホ曲線で、無限小の細かさを持つ刺状の構造を非常に多く持つ。リアス式海岸など、自然界にはフラクタル構造を持つ複雑な構造が多く存在する。

(注3) カーネル法・k近傍法：

共にデータ解析手法の一種。カーネル法はカーネル関数と呼ばれる所与の関数の和で予測を構成する。k近傍法は、データそれぞれの近傍から代表点を発見して予測を構成する。

(注4) 可微分多様体

幾何学的な形状を表現する概念。集合に対して、局所的にユークリッド空間とみなせる滑らかな構造を導入したもの。

8. 添付資料

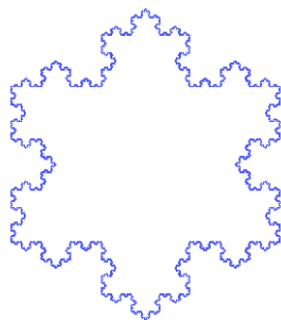


図1：フラクタルの一例、コッホ曲線。直線に刺を生やすという操作を無限回繰り返すことで作られる。無限に細かい刺を無限個持ち、得られる曲線はほとんど至る所で微分不可能で

ある。

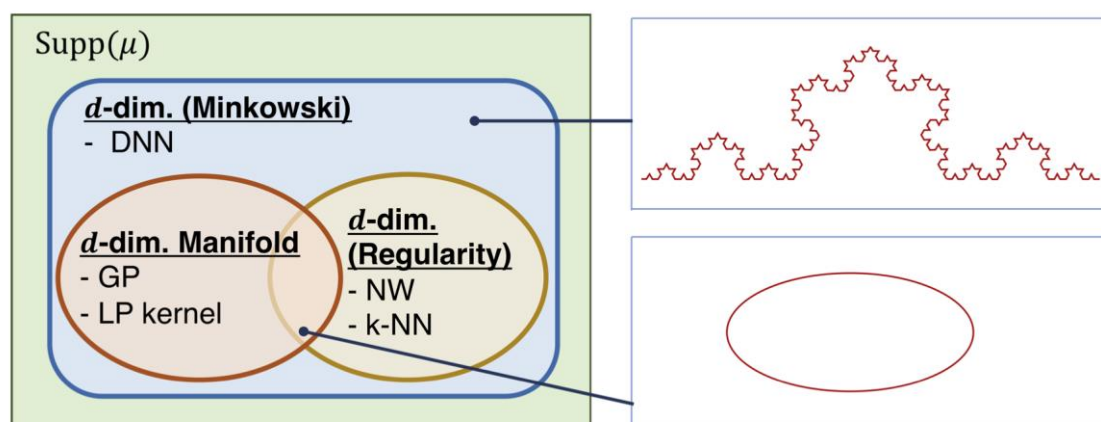


図 2：扱えることが判明している低次元構造の包含関係を表した図。深層学習（DNN）はフラクタル（右上：コッホ曲線）を扱えるが、カーネル法などの他手法（GP, LP kernel, NW, k-NN）が明示的に扱えるのは滑らかな低次元構造（右下：楕円）である。