

令和 3 年度 【 学園研究費助成金 < B > 】 研究成果報告書

学部名 現代マネジメント

フリガナ イイ マサル
氏 名 石井雅治

研究期間 令和 3 年度

研究課題名 深層神経回路における特異値と関数空間の部分空間対応

研究組織

	氏 名	学 部	職 位
研究代表者	石井雅治	現代マネジメント	教授
研究分担者			
研究分担者			

1. 本研究開始の背景や目的等 (200 字~300 字程度で記述)

近年における深層神経回路の学習（以下、深層学習）の広範囲の成功の核は、データから自律的に高い汎化性能を達成することにある。しかしそのメカニズムには依然として謎が残っており、その解明は、今日における機械学習の最も重要な課題の 1 つであるといつてよい。

深層学習のスタンダードとなっている最急降下に関し、我々は先の研究でその勾配の特異値を通して関数空間の部分空間ごとに別々の速度で収束することを示唆したが、このことによつて早期停止が高い汎化能力を与える可能性がある。本研究は、この部分空間の構造と速度差の解析を通して、この課題の少なくとも一部を解き明かすことを目的とする。

2. 研究の推進方策 (300 字程度で記述)

- (1) 深層学習において、最急降下について、その勾配を適当な関数空間における正規直交系（以下、この基底を元基底）で表し、特異値分解を理論的に表し、数値的にも検証した。
- (2) 特異値分解に対応する、関数空間の正規直交基底と（以下、特異値基底）と元基底の関係を、単に数値的にではなく、解析的に関係を調べた。特に元基底として、汎化能力と密接に関わる複雑度基底に関わるものを利用した。
- (3) 各特異値に対応する基底の張る部分空間ごとに、上の結果を利用して、相対的に先に学習が進行する部分空間と汎化能力の関係を解析することによつて、非線形ダイナミクスをモデルの複雑度のダイナミクスに関係付けた。

3. 研究成果の概要 (600字~800字程度で記述)

勾配降下を用いる深層神経回路は高性能ではあり、学習を適当なところで終了する“早期終了”が高い汎化能力を与えることが知られている。しかし、なぜ高い汎化能力性能が生じるかは十分に理解されてこなかった。また、従来の AIC や WAIC 等の統計理論では説明できない、“2重降下”と呼ばれる奇妙な現象が観察されている。回帰問題に関する通常の学習では、同じ学習データに対し、パラメータ数を増やしていくと、最初は汎化誤差が減少するが、パラメータ数がデータサイズに近づくと増大し、それを超えると再び減少する。ところが、早期終了すると、多くの場合、パラメータ数の増大について汎化誤差もほぼ単調に減少するというものである。我々の研究で得られた次の結果によって、このような現象を理解する力学の枠組みが与えられる。

まず、勾配降下の勾配の空間に空間周波数で定義された”複雑度(関数)基底”を導入し、この空間を特異値分解した。以下、勾配の非退化なパラメータ数を m で、サンプル数を n で表す。このとき、モデルを複雑度基底で近似展開するによって、モデルの成分の複雑度に m 個の順位が存在し、これが特異値の m 個の順位に近似的に対応することを証明した。

これによって、勾配降下の非線形ダイナミクスをモデルの複雑度のダイナミクスに關係付けられる。すなわち、勾配降下では、モデルに複雑度の低いパラメータ初期値をとると、概ね、複雑度の低い構造から先に学習が進み、複雑度の高い構造はその後に学習されることが主張できる。特に、複雑度の順位について、低い順に残差が削られるが n までで停止し、 n を超える ($m > n$ の場合) 順位については、残差はほとんど削られない。

この結果とモデルの複雑度基底による展開を用いることによって、上記のような現象についての理解や、少なくとも従来の個別的な研究結果に統一的視点からの解釈を与えることができる。

4. キーワード (本研究のキーワードを1項目以上8項目以内で記載)

①最急降下	②早期終了	③深層学習	④汎化能力
⑤特異値分解	⑥関数空間	⑦複雑度基底展開	⑧

5. 研究成果及び今後の展望 (公開した研究成果、今後の研究成果公開予定・方法等について記載すること。既に公開したものについては次の通り記載すること。著書は、著者名、書名、頁数、発行年月日、出版社名を記載。論文は、著書名、題名、掲載誌名、発行年、巻・号・頁を記載。学会発表は発表者名、発表標題、学会名、発表年月日を記載。著者名、発表者名が多い場合には主な者を記載し、他〇名等で省略可。発表数が多い場合には代表的なもののみ数件を記載。)

研究成果

石井雅治, 深層神経回路における勾配降下の複雑度展開と勾配の退化性, 力学系理論の最近の進展とその応用, 数理解析研究所講究録, 京都大学数理解析研究所, 2022年(掲載予定 No, Page 未定).

今後の展望

以上の結果の適用範囲の全貌は未だ明らかになっていないので、これをまとめ、神経回路網の主要学会である NeurIPS 等に発表を申込みたい。また、最急降下の汎化能力とダイナミクスの関係がかなり明らかになったが、勾配の退化次数がこの関係に強い影響を与えるのでその推移の状況を詳しく解析したい。