

令和 5 年 6 月 19 日現在

機関番号：82626

研究種目：若手研究

研究期間：2019～2022

課題番号：19K20366

研究課題名（和文）ランダム深層ニューラルネットの数理的基盤の構築とその学習への応用

研究課題名（英文）Mathematical Foundations of Random Deep Neural Networks and their applications to machine-learning problems

研究代表者

唐木田 亮（Karakida, Ryo）

国立研究開発法人産業技術総合研究所・情報・人間工学領域・主任研究員

研究者番号：30803902

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究の目的は、ランダム結合をもつニューラルネットワークの解析に基づいて、深層学習に有用な数理的知見を獲得することである。この目的を目指して、まずパラメータ空間の幾何学的構造を定めるFisher情報行列の固有値を解析した。これにより、正規化層の効果や学習率の適切な設定に関して定量的な説明を与えた。さらにランダム神経回路まわりの摂動範囲での学習として特徴づけられるNTKレジームでは、自然勾配法の適切な近似を明らかにした。また記憶埋め込み型の連想記憶モデルに関連して、VAEにおける記憶想起過程やModern Hopfieldネットに対応するボルツマンマシンを解明にした。

研究成果の学術的意義や社会的意義

ランダム神経回路は古典的に理論的神経科学の枠組みで発展してきたが、近年は深層学習にその枠組みを拡張し、たとえば逆誤差伝播における解析が進みつつある。本研究課題もこの流れに沿うもので、特に、学習のプロセスに大きく影響を与えるFisher情報行列やNTK行列に着目し、各種モデルや学習手法の性質を明らかにした点に独自性があり学術的意義がある。本成果は様々な応用を支える基礎技術に理解を与えており、今後の深層学習技術の研究開発を進めるうえで有用となることが期待でき、その点で社会的意義もあるといえるだろう。

研究成果の概要（英文）：The purpose of this research is to gain mathematical insights for deep learning, based on the analysis of random neural networks. Toward this goal, we first analyzed the eigenvalues of their Fisher information matrix, which determine the geometric structure of the parameter space. This allowed us to provide a quantitative explanation of the effects of normalization layers and appropriate settings for learning rates. In the NTK regime, characterized by learning within the range of perturbation around initial random weights, we clarified the appropriate designs of approximated natural gradient methods. Related to associative memory models, we elucidated Boltzmann machines corresponding to Modern Hopfield networks and the memory recall process in VAEs.

研究分野：ソフトコンピューティング

キーワード：深層学習 機械学習 ニューラルネットワーク 統計力学 ランダム行列 数理工学

1. 研究開始当初の背景

学習素子を層状に並べた多層モデルである深層学習 (deep learning) は、柔軟な情報表現が可能で、画像・音声・テキストといった様々なデータの認識や生成で高い性能を発揮した (LeCun+, Nature, 2015)。深層学習では実用ベースでの開発が進み、多種多様なネットワークアーキテクチャ、学習アルゴリズム、学習の枠組みが提案されている。このため、対象の複雑性が高く、数理的な解析には困難も多い。それと同時に、数理的な背景がはっきりしないヒューリスティックな学習手法への依存が大きいともいえる。こうした現状に対し、深層学習の背後に潜んでいる普遍的な構造や現象を発見および理解できるか、というのが本研究の問いである。

2. 研究の目的

本研究の目的は、ランダム結合をもつニューラルネットワークの解析に基づいて、深層学習に有用な数理的知見を獲得することである。本研究は結合パラメータがランダム変数として生成されたランダム深層ニューラルネットワーク解析し、モデルや学習手法の詳細に依存しない普遍的な現象を数理的に解明する。計算論的神経科学と統計物理の分野では、ランダムな再帰結合を持つリカレントニューラルネットワークの解析が進んできた。ランダムネスはモデルサイズが十分大きい極限で、中心極限定理のような平均操作を可能にしてモデルを粗視化する (e.g. 平均場理論, ランダム行列理論)。すなわち、モデルやパラメータの詳細に依存しない“普遍性”を本質的情報として抽出できる。このランダム結合の数理を深層学習の各種問題に適用する点に本研究の特色がある。この操作によって深層学習の複雑な設定を、層から層への低次元の状態発展方程式に縮約し、深層学習のメカニズムやヒューリスティクスを理解を促進できると期待される。

3. 研究の方法

ランダム結合をもつニューラルネットワークモデルにおいて、統計力学的解析を主軸に行った。より具体的には、古典的に順方向の信号伝播を解析してきた統計神経力学、近年になり逆誤差伝播に拡張された平均場理論、さらにモデルの挙動を支配する行列の固有値解析にランダム行列理論、を用いて解析を実施した。具体的な深層学習の問題を解析するなかで、各解析の適用範囲の拡張や限界の理解を推し進め、数理的基盤の構築を進めた。また、このように解けるモデルでの解析を行うなかで、解析対象の学習手法のメカニズム理解や未知の現象の同定も進めた。同時に、明らかになったメカニズムや現象が実際の深層学習の各種モデルで見られるかを計算機実験によって実証し、理論的にも実験的にも整合するアルゴリズムや学習手法の理解と設計指針を与えた。

4. 研究成果

(1) パラメータ空間の構造

本テーマは研究課題の根幹をなすものであるが、これに対して複数の研究成果を得ることができた。まず、深層学習で広く用いられる batch normalization ありのモデルにおいて、ランダム結合下での Fisher 情報行列の固有値解析を行った。これにより、batch normalization されたモデルの最大固有値が小さくなることを発見した。すなわちパラメータ空間の幾何学的な構造は、より平坦に近い。この結果は、より大きな学習率をモデルの大きさに依存せずに設定できることを示唆しており、経験的な知見ともコンシステントである。ランダム初期化された幅の大きなモデルに着目することで、層数や活性化関数の種類の詳細に依存しない、普遍的な batch normalization の理論的な理解を与えることができた。特に我々の結果は、類似の手法である layer normalization ではこの効果が見えないことを示唆しており、アルゴリズム間の挙動の違いに対して、新しい知見を与えた [Karakida+ NeurIPS 2019]。

さらに、これまでの解析では簡単のために損失関数として平均二乗誤差を仮定していたが、識別問題に対応するため、実問題で広く使用されているクロスエントロピー損失関数において同様に Fisher 情報行列の固有値を解析した。この場合も、平均二乗誤差の場合と同様にモデル幅に依存した外れ固有値が出現し、さらにソフトマックス出力によってばらつきが生まれることを明らかにした。Batch normalization の効果も解析し、外れ固有値を抑える効果が現れることを定量的に明らかにした。また、Fisher 情報行列はパラメータ空間の計量であるが、同様の性質は入力空間の計量においても現れることを明らかにし、入力の摂動に対するモデルの頑健性を議論するための基盤を提供した [Karakida+ Neural Computation (2021)]。

以上は平均場理論の範囲で理解できる固有値の統計的性質の結果である。より高次の情報、具体的にはスペクトル形状に関しては一般にランダム行列理論による取扱いが必要となる。我々は自由確率論を利用し、ある種の条件付き Fisher 情報行列のスペクトル形状について成果を得た。また超深層の状況では最大固有値が層数に依存するため、適切な初期値スケールが必要となる。理論的に得られた最大固有値に基づいて、勾配法が収束に向かうために必要な学習率の大きさを推測し、実験的にその妥当性を明らかにした [Hayase & Karakida, AISTATS

(2) 深層学習手法の解析

ランダム神経回路に基づく解析はランダム初期値におけるのぞましいアーキテクチャや初期化の設定に示唆を与えてきた。近年は、Neural Tangent Kernel (NTK) レジームが同定されたことで、ランダム初期値における NTK や Fisher 情報行列の性質がある程度、訓練ダイナミクスを左右することがわかっている。より詳細には、NTK レジームは Lazy レジームの一種で、パラメータが微小変化しモデルが一次近似される範囲で大域収束する状態である。これは過剰パラメータ系の最も単純な可解モデルとみることができる。恣意性やヒューリスティクスの多い深層学習法が、このような解けるモデルでいかに理解されるかは、学習法の本質を理解するうえでも、さらに洗練された学習手法を開発するうえでも重要な知見である。

我々はまず、Fisher 情報行列とも本質的にかかわる勾配法として、自然勾配法の解析を NTK レジームにおいて実施した。自然勾配法は学習の収束を高速化するために開発されている手法であるが、深層学習では計算量を下げするために多数の近似が用いられている状況である。特に、K-FAC はその簡便さから広く使われている手法であるが、厳密な自然勾配とは異なる大胆なヒューリスティクスに基づいており、その挙動の理解は限定的であった。我々は NTK レジーム内でこうした近似自然勾配法を解析し、適切なハイパーパラメータの設定のもと、経験的に開発されてきた近似自然勾配法のいくつかを、近似なしの厳密な自然勾配法と同じ訓練収束のダイナミクスを達成できることを明らかにした。訓練誤差の収束レートは K-FAC を含む近似自然勾配法と厳密な自然勾配法で一致させることができる。さらに論文では、近似自然勾配法が高速な収束を実現するために満たすべき一般的な条件にも洞察を与えており、勾配法の改良につながる有用な知見を与えられたといえる [Karakida & Osawa, NeurIPS 2020]。

次に、近年の深層学習の発展によって注目を集めている知識転移、特に継続学習において、汎化誤差解析を実施した。転移学習や継続学習は近年の深層学習の発展を支えている枠組みであるが、その運用には多くの恣意性があり、どのような状況化でモデルの性能向上がみられるかは単純な可解モデルであってもその理解が限定的であった。本研究では NTK レジームにおける線形モデルを考え、継続学習の汎化性能をレプリカ解析によって遂行した。レプリカ解析は汎化性能をバウンド評価ではなく、その期待値(典型値)を直接評価することを可能にするため、タスク間での汎化誤差の増減を正確に追うことができる点に利点がある。解析の結果、知識転移が性能の劣化なく機能するためには、学習するタスクの類似度だけでなく、タスク間の訓練サンプル数の均衡が重要であることが明らかとなった。この未知の現象は、実際の深層モデルの継続学習でも観測されることも明らかにし、知識転移において理論と実験の両面から理解を進めることができた [Karakida & Akaho, ICLR 2022]。今後、サンプル数の均衡が崩れた場合に忘却を回避する学習の補正方法の開発が必要だろう。

(3) 構造のあるモデル

ここまで簡単のために、MLP に代表される均一的な構造を持つモデルで解析を実施してきたが、畳み込みや注意機構といったモデルの構造が本質的に寄与を持つ現象や性能向上があるだろう。これに関してはまず、所受容野の構造を埋め込んだランダム結合を持つモデルで NTK レジームにおける学習を考え、局所受容野の構造に応じて、入力摂動に対する頑健性が得られることを明らかにした [Watanabe+, Neural Networks (2023)]。

古典的なランダム神経回路では、完全ランダム結合とランダム記憶埋め込み結合の 2 つのモデルが研究されてきた背景がある。前者は深層学習においてランダム初期化として位置づけることができる。後者は訓練後のモデル挙動を理解するための抽象化された数理モデルであるが、この数理モデルの知見が近年の深層学習にどのような示唆を与えるかは興味深い問題である。まず、我々は訓練された VAE モデルにおいて、連想記憶モデルと類似した記憶の引き込みダイナミクスが見られることを実験的に明らかにした。すなわち、モデルにおける発火パターンのサンプリングを繰り返すとき、このパターンは訓練パターンの平均(コンセプトベクトル)に対応したものに近づく。この知見は、実際の訓練済みモデルで連想記憶モデルの挙動が見られることを示しており、より広範で現代的なモデルの連想記憶的側面の理解につながっていくと期待できる。

また、連想記憶モデルに関連しては近年、深層学習で広く使われる自己注意機構と連続値 Hopfield モデルとの対応が報告されている。このモデルの理解を広げるため、この Modern Hopfield モデルの学習、すなわち対応する Boltzmann machine を提案し、その学習の典型的挙動を理論解析および数値実験から明らかにした。注意機構は対応する Boltzmann machine においてある種の基底選択(混合ガウス分布の各ガウス分布の選択)と解釈でき、また既存の Boltzmann machine に比べて尤度計算が容易であることも判明した。記憶埋め込み型モデルの視点から利便性の高い深層モデルの拡張・提案を行う研究群に対し、基盤となる知見を与えることが期待される [Ota & Karakida, Neural Computation (2023)]。

5. 主な発表論文等

〔雑誌論文〕 計13件（うち査読付論文 12件 / うち国際共著 0件 / うちオープンアクセス 10件）

1. 著者名 Ryo Karakida, Shotaro Akaho	4. 巻 -
2. 論文標題 Learning Curves for Continual Learning in Neural Networks: Self-Knowledge Transfer and Forgetting	5. 発行年 2022年
3. 雑誌名 International Conference on Learning Representations	6. 最初と最後の頁 1-27
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Watanabe Kaito, Sakamoto Kotaro, Karakida Ryo, Sonoda Sho, Amari Shun-ichi	4. 巻 160
2. 論文標題 Deep learning in random neural fields: Numerical experiments via neural tangent kernel	5. 発行年 2023年
3. 雑誌名 Neural Networks	6. 最初と最後の頁 148 ~ 163
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.neunet.2022.12.020	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Toshihiro Ota, Ryo Karakida	4. 巻 -
2. 論文標題 Attention in a family of Boltzmann machines emerging from modern Hopfield networks	5. 発行年 2023年
3. 雑誌名 Neural Computation	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Tomohiro Hayase, Ryo Karakida	4. 巻 -
2. 論文標題 The Spectrum of Fisher Information of Deep Networks Achieving Dynamical Isometry	5. 発行年 2021年
3. 雑誌名 Proceedings of AISTATS	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Karakida Ryo, Akaho Shotaro, Amari Shun-ichi	4. 巻 33
2. 論文標題 Pathological Spectra of the Fisher Information Metric and Its Variants in Deep Neural Networks	5. 発行年 2021年
3. 雑誌名 Neural Computation	6. 最初と最後の頁 2274 ~ 2307
掲載論文のDOI (デジタルオブジェクト識別子) 10.1162/neco_a_01411	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Karakida Ryo, Osawa Kazuki	4. 巻 2021
2. 論文標題 Understanding approximate Fisher information for fast convergence of natural gradient descent in wide neural networks*	5. 発行年 2021年
3. 雑誌名 Journal of Statistical Mechanics: Theory and Experiment	6. 最初と最後の頁 124010 ~ 124010
掲載論文のDOI (デジタルオブジェクト識別子) 10.1088/1742-5468/ac3ae3	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Takase Tomoumi, Karakida Ryo, Asoh Hideki	4. 巻 442
2. 論文標題 Self-paced data augmentation for training neural networks	5. 発行年 2021年
3. 雑誌名 Neurocomputing	6. 最初と最後の頁 296 ~ 306
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neucom.2021.02.080	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Ryo Karakida, Kazuki Osawa	4. 巻 33
2. 論文標題 Understanding Approximate Fisher Information for Fast Convergence of Natural Gradient Descent in Wide Neural Networks	5. 発行年 2020年
3. 雑誌名 Proceedings of Conference on Neural Information Processing Systems (NeurIPS)	6. 最初と最後の頁 10891--10901
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Nagano Yoshihiro, Karakida Ryo, Okada Masato	4. 巻 10
2. 論文標題 Collective dynamics of repeated inference in variational autoencoder rapidly find cluster structure	5. 発行年 2020年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-020-72593-4	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Ryo Karakida, Shotaro Akaho, Shun-ichi Amari	4. 巻 -
2. 論文標題 Universal Statistics of Fisher Information in Deep Neural Networks: Mean Field Approach	5. 発行年 2019年
3. 雑誌名 Proceedings of Conference on Artificial Intelligence and Statistics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shun-ichi Amari, Ryo Karakida, Masafumi Oizumi	4. 巻 -
2. 論文標題 Fisher Information and Natural Gradient Learning in Random Deep Networks	5. 発行年 2019年
3. 雑誌名 Proceedings of Conference on Artificial Intelligence and Statistics	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shun-ichi Amari, Ryo Karakida, Masafumi Oizumi	4. 巻 10
2. 論文標題 Statistical neurodynamics of deep networks: geometry of signal spaces	5. 発行年 2019年
3. 雑誌名 Proceedings of Nonlinear Theory and Its Applications	6. 最初と最後の頁 322-336
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/nolta.10.322	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ryo Karakida, Shotaro Akaho, Shun-ichi Amari	4. 巻 -
2. 論文標題 The Normalization Method for Alleviating Pathological Sharpness in Wide Neural Networks	5. 発行年 2019年
3. 雑誌名 Proceedings of Conference on Neural Information Processing Systems	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計18件 (うち招待講演 8件 / うち国際学会 4件)

1. 発表者名 唐木田亮
2. 発表標題 カーネル法の統計力学的解析とそれによる継続学習の評価
3. 学会等名 統計物理と統計科学のセミナー (招待講演)
4. 発表年 2022年

1. 発表者名 唐木田亮
2. 発表標題 Neural tangent kernel regimeにおける継続学習の学習曲線
3. 学会等名 日本応用数理学会2022年度年会 (招待講演)
4. 発表年 2022年

1. 発表者名 唐木田亮
2. 発表標題 継続学習における自己知識転移と忘却
3. 学会等名 第51回統計的機械学習セミナー (招待講演)
4. 発表年 2022年

1. 発表者名 唐木田亮, 高瀬朝海, 早瀬友裕, 大沢和樹
2. 発表標題 効率的な勾配正則化アルゴリズムとその陰的バイアスの解析
3. 学会等名 IBIS2022
4. 発表年 2022年

1. 発表者名 唐木田亮, 高瀬朝海, 早瀬友裕, 大沢和樹
2. 発表標題 対角線形ネットにおける勾配正則化の陰的バイアス
3. 学会等名 日本物理学会2023年春季大会
4. 発表年 2023年

1. 発表者名 Tomohiro Hayase, Ryo Karakida
2. 発表標題 The Spectrum of Fisher Information of Deep Networks Achieving Dynamical Isometry
3. 学会等名 AISTATS (国際学会)
4. 発表年 2021年

1. 発表者名 Ryo Karakida
2. 発表標題 Improving the trainability of deep neural networks: A perspective from the infinite width limit
3. 学会等名 4th international conference on econometrics and statistics (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 唐木田 亮, 赤穂 昭太郎
2. 発表標題 継続学習における転移と忘却: NTK regimeのレプリカ解析
3. 学会等名 IBIS2021
4. 発表年 2021年

1. 発表者名 Ryo Karakida
2. 発表標題 Understanding Approximate Fisher Information for Fast Convergence of Natural Gradient Descent in Wide Neural Networks
3. 学会等名 Math Machine Learning Seminar MPI MIS + UCLA (招待講演)
4. 発表年 2021年

1. 発表者名 唐木田 亮
2. 発表標題 幅無限大深層モデルにおける近似自然勾配法の収束解析
3. 学会等名 日本物理学会 第76回年次大会
4. 発表年 2021年

1. 発表者名 唐木田 亮
2. 発表標題 深層学習の数理: ランダム行列と統計力学的視点
3. 学会等名 Random Matrices, Free Probability, and Machine Learning ワークショップ (招待講演)
4. 発表年 2020年

1. 発表者名 唐木田 亮
2. 発表標題 深層学習の数理: 統計力学的アプローチ
3. 学会等名 ディープラーニングと物理学2020 オンライン (招待講演)
4. 発表年 2020年

1. 発表者名 唐木田 亮, 大沢 和樹
2. 発表標題 深層モデルにおいて高速に収束する近似自然勾配法の理論解析
3. 学会等名 IBIS2020
4. 発表年 2020年

1. 発表者名 Ryo Karakida
2. 発表標題 Understanding Approximate Fisher Information for Fast Convergence of Natural Gradient Descent in Wide Neural Networks
3. 学会等名 Conference on Neural Information Processing Systems (NeurIPS) (国際学会)
4. 発表年 2020年

1. 発表者名 Ryo Karakida
2. 発表標題 Fisher Information of Deep Neural Networks With Random Weights
3. 学会等名 The 11th ICISA international conference (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 唐木田亮
2. 発表標題 ランダムなBackpropagation学習における巨視的ダイナミクスの生成汎関数法的解析
3. 学会等名 日本物理学会 第75回年次大会
4. 発表年 2020年

1. 発表者名 Ryo Karakida, Shotaro Akaho, Shun-ichi Amari
2. 発表標題 Universal Statistics of Fisher Information in Deep Neural Networks: Mean Field Approach
3. 学会等名 Conference on Artificial Intelligence and Statistics
4. 発表年 2019年

1. 発表者名 Ryo Karakida, Shotaro Akaho, Shun-ichi Amari
2. 発表標題 The Normalization Method for Alleviating Pathological Sharpness in Wide Neural Networks
3. 学会等名 Conference on Neural Information Processing Systems
4. 発表年 2019年

〔図書〕 計1件

1. 著者名 唐木田 亮	4. 発行年 2020年
2. 出版社 サイエンス社	5. 総ページ数 7
3. 書名 数理科学 深層神経回路網の幾何～統計神経力学とのつながり～	

〔産業財産権〕

〔その他〕

研究者が作成したwebページ
<https://sites.google.com/view/ryokarakida/english>
唐木田亮 website
<https://sites.google.com/view/ryokarakida/>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------