

令和 5 年 4 月 17 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2017～2022

課題番号：17K12653

研究課題名（和文）転移学習の理論的解析とその応用

研究課題名（英文）Theoretical Analysis of Transfer Learning and Its Applications

研究代表者

熊谷 亘 (Kumagai, Wataru)

東京大学・大学院工学系研究科（工学部）・特任助教

研究者番号：20747167

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：第一の主結果として、パラメトリックなモデルをドメイン間で転移する状況における理論バウンドを導出した。特に特徴抽出器として複雑なパラメトリックモデルを取ることでもでき、深層ニューラルネットワークやスパース符号化を理論的に取り扱うことができるものである。第二の主結果として、同変性とよばれる代数的な性質を加味した場合のメタ学習器の普遍性を示した。同変性はデータ処理や自然界の過程で自然に現れる性質であり、同変的なニューラルアーキテクチャによる処理により学習を効率化させることができるという利点がある。第三の主結果として、同時分布に関する期待リスクの差の分解定理を導出した。

研究成果の学術的意義や社会的意義

初めに学術的意義について述べる。転移学習は現在の機械学習や人工知能の研究において欠かせない技術である。本研究により、転移学習の理論的側面の一端が明らかになり、効率的なモデルの構築や転移学習手法の構築に資することが期待できる。特にメタ学習において同変性を用いた新規のモデルを提案したが、これはデータ内の対称性という代数的性質を学習の効率化に結びつけるために重要な結果と言える。次に社会的な意義について述べる。転移学習技術は多数のドメインでの学習をサポートするもので、幅広い応用で成功を収めている。本研究結果は今後の転移学習の応用においてその理論的基盤の構築に貢献するものである。

研究成果の概要（英文）：As the first main result, we derived theoretical bounds for transferring parametric models across domains. Notably, this can involve employing complex parametric models as feature extractors, enabling the theoretical treatment of deep neural networks and sparse coding. As the second main result, we demonstrated the universality of meta-learners when considering an algebraic property called equivariance. Equivariance naturally emerges in data processing and natural processes, and it offers the advantage of enhancing learning efficiency through processing with covariant neural architectures. As the third main result, we derived a decomposition theorem for the difference in expected risks with respect to joint distributions.

研究分野：人工知能

キーワード：転移学習 メタ学習 パラメータ転移

1. 研究開始当初の背景

本研究は転移学習の理論解析を行うことを目的としてきた。まずは転移学習について説明する。転移学習とは機械学習理論の分野の一つである。転移学習の手法を用いると、注目しているタスクに関連したデータが少ない場合でも、直接関連のないデータの助けを借りて学習を行うことができる。はじめに、転移学習と従来の機械学習の差について具体例を通して説明する。ここでは例として、画像のカテゴリ分類を考える。データベース DB1 に“犬”と“猫”の画像が保存されており、すでに正しくカテゴリ分けなされているとしよう。ここでのタスクは、新たに“犬”または“猫”の画像がデータベース DB1 に保存されたとき、その画像を正しいカテゴリに分類することである。DB1 に保存されている画像が少ない場合、従来の機械学習では学習が十分に行えず、カテゴリ分類の成功率が低くなってしまふ。そこで転移学習では、別のデータベース DB2 に保存されている画像を用いて学習を行い、それによってデータの少なさを補うことを考える。ここでは、DB2 には“犬”や“猫”とは限らず、カテゴリ分けもされていない多種多様な画像が保存されていてもよい。従来の機械学習では、与えられるデータは同一分布に従って発生すると仮定しており、このような状況は扱うことはできない。一方、転移学習では DB1 と DB2 のデータの分布の同一性は仮定されない。そのような場合でも、DB2 で何らかの構造を学習し、その知識を転移することができる。そして転移された知識を用いて、DB1 における“犬”と“猫”の特徴を適切に抽出し、高い正解率でカテゴリ分類を行うことができるのである。

次に研究の背景を述べる。転移学習の文脈では、転移する知識の送り手側は元ドメイン、受け手側は目標ドメインと呼ばれる。上記の例では、DB2 が元ドメイン、DB1 が目標ドメインに当たる。

(A)本研究では、パラメータ転移のアプローチに着目してきた。これを一般的に述べると、元ドメインで何らかのパラメータを学習し、それを目標ドメインに転移する方法とすることができる。パラメータ転移アプローチは、様々な応用がなされている一方で、理論的解析がほとんどなされていなかった。従って、提案されている様々な手法が、どのような仮定のもとでどの程度有効なのかは明らかになっていなかった。申請者は論文 [1] において、パラメータ転移アプローチに特化した理論解析を行っていた。

(B) パラメータ転移アプローチを用いるためには、元ドメインで有効なパラメータを十分正確に学習しておく必要がある。従って、元ドメインではしばしば大量のデータ要する。この問題を解決する手段として、自己教示学習と呼ばれる、元ドメインのデータとしてラベルなしデータを許す問題設定が提案されていた [2]。自己教示学習は、深層学習などの強力な教師なし学習を援用できるため、特に注目を集めていた。しかし、自己教示学習アルゴリズムは、実験的な性能が優れていることが確認されている一方で、理論的な性能評価が全くなされていなかった。申請者は論文 [1] において、データがスパースモデルに従うという仮定のもとで、スパース符号化について考察した。さらに、自己教示学習の文脈で、初めて汎化誤差を導出した。

(C) 転移学習の設定はドメイン間のデータ空間の違いにより二つに分類される。元ドメインと目標ドメインでのデータ空間が等しいときの知識の転移を同質的、異なっているときを変異的という。例として元ドメイン DB2 と目標ドメイン DB1 で共に画像データを扱う場合、両ドメインの画像データの構造(ピクセルサイズ、カラーか白黒か、など)が、完全に一致している場合は同質的であり、何か一つでもデータ構造が異なっている場合は変異的である。同質的である状況は限定的であり、応用的な多くの状況は変異的な設定に属する。変異的な状況で、ドメイン間の構造が大きく異なる場合は、知識の転移によって逆にアルゴリズムの性能が低下する、負転移と呼ばれる現象がしばしば起こる。負転移に関しては、ヒューリスティックな回避方法が多く提案されているものの、理論的に負転移を回避する方法は提案されていない。

2. 研究の目的

本研究の目的は、転移学習の理論的基礎を深化させ、性能の良い転移学習アルゴリズムを提案することである。具体的には、(A)パラメータ転移と呼ばれるアプローチに着目し、汎化誤差などの理論評価を行う。また、パラメータ転移と親和性があり、かつ深層学習などと相性の良い(B)自己教示学習の理論を構築する。さらに、応用上重要である(C)変異的転移と呼ばれる、同質でないタスク間で知識を転移させる問題設定を扱う。これらのテーマにおいて、新たなアルゴリズムを作成し、実データにおいてその有効性を検証する。

3. 研究の方法

(A)申請者が論文[1]においてパラメータ転移アルゴリズムを提案した。そのアルゴリズムは、ドメイン間で共通のパラメータ θ を、元ドメインで学習し、それを目標ドメインへ転移するものであった。論文 [1] を含む多くの研究では、目標ドメイン内の データに限れば、独立同一分布から発生していることを仮定している。従って、元ドメインで学習したパラメータ θ を固定してしまえば、残りのパラメータは目標ドメインの独立同一分布からのデータのみで学習を行っていることになる。論文 [1] では、この点に着目し、独立同一性に基づく従来の汎化誤差の評価に帰着している。しかし実際の応用では、転移したパラメータ θ は初期値として用いられ、目標ドメインのデータでさらに再学習するウォームスタートに利用されている。本研究では、このウォームスタートの状況でも汎化誤差の解析を行う。ウォームスタートを用いる場合、転移されたパラメータ θ は、目標ドメインのデータを用いてさらに更新される。従って、2 つのドメインで二段階でパラメータを学習する過程で、データの同一性は成り立たない。これを解決するために、Azuma の不等式や McDiarmid の不等式などの、同一性を仮定しない集中不等式を用いて評価を行うことを試みた。

既存の転移学習の文脈では、元ドメインのデータが従う分布は同一とは限らないが、独立性は仮定していた。本研究では、元ドメインのデータの独立性の仮定を排除することも試みる。これが可能になれば、時系列データなど、独立とは限らない幅広いデータセットからの知識の転移を行うことができる。具体的には、オンライン学習に基づいたクラス予測に関する結果を用いて、元ドメインで適切な仮説のパラメータを学習する。そして、そのパラメータを目標ドメインへ転移する。データの独立性すら失われているので、そのようなアルゴリズムに関しては汎化誤差ではなく、オンライン学習に対するリグレット解析を援用する。

(B)申請者の論文 [1] の結果を深化させる方針をとった。論文 [1] では、データがスパースモデルと呼ばれる性質の良い確率モデルに従うと仮定している。ここでスパースモデルとは、データはサンプル空間のある過完備基底のスパースな一次結合で表されるとするものである。このとき、過完備基底は辞書と呼ばれ、この辞書をパラメータとみなすことができる。論文 [1] では、この事実を用いてパラメータ転移の文脈に帰着させ、スパース符号化に対する汎化誤差を与えている。本研究では、目標ドメインではスパース符号化だけではなく、構造化スパース符号化や多重カーネルなど、他の手法も対象とする。特に、深層学習も主要な対象とする。深層学習に基づいた転移学習アルゴリズムでは、元ドメインでネットワーク重みを学習し、それを目標ドメインへ転移する。ここではネットワーク重みをパラメータとみることで、パラメータ転移アプローチとみなすことができる。従って(A)の テーマと関連して、深層学習の理論解析を進めることができる。

自己教示学習において、元ドメインのデータはラベルなしなので、データ数はある程度多いと想定できる。しかし、データ数が十分足りているかどうかは、許容誤差に応じて異なり、それは自己教示学習の文脈では全く明らかになっていない。それを明らかにするために、自己教示学習の文脈ではサンプル複雑性についての評価を行う。サンプル複雑性とは、汎化誤差を許容値以下にするために十分なデータ数のことである。汎化誤差からサンプル複雑性を把握することは理論的にだけでなく、実用上重要である。申請者の論文 [1] では、元ドメインで辞書を学習し、目標ドメインでスパース符号化を行うアルゴリズムに対し、汎化誤差の上界を導出した。その結果からは、元ドメインのデータサイズが非常に大きく、辞書学習が十分に正確である場合には、汎化誤差の上界のオーダーが $1/\sqrt{n}$ になることが分かる。ここで n は目標ドメインのデータ数である。しかし、元ドメインのデータ数が具体的にどの程度大きければ十分かは明らかになっていない。本研究では、まずスパース符号化に関する論文 [1] の汎化誤差をより深く解析し、サンプル複雑性を明らかにする。さらにその結果を、構造化スパース符号化や深層学習などの場合へ拡張する。

(C) 初めに、元ドメインと目標ドメインにおいて、画像データが利用できる状況を考える。ここでは、ピクセルサイズや色設定 (カラーか白黒か) が異なる程度の、ドメイン間の差が比較的少ない場合を想定している。ただし、画像内の内容はドメイン間で異なる状況を作る。例としては、元ドメインでは風景の画像が、目標ドメインでは動物の画像が保存されているような状況である。そのような データセットに対して、既存の変異的転移アルゴリズムを適用し、負転移を起こしやすいデータ構造を考察する。ここで用いる具体的なアルゴリズムとして、両ドメインのデータを同一の空間に射影し、同質的な状況に帰着させる転移アルゴリズムを想定している。これは典型的なタイプの変異的転移アルゴリズムであり、既存の手法が負転移に陥る状況を把握するのに適していると考えられる。

さらに、自然言語に対して同様の考察を行う。すなわち、元ドメインと目標ドメインに、内容の異なる文章データが保存されている状況を想定する。そして、変異的転移アルゴリズムを適用し、どのようなデータセットのとき、負転移が生じるかを考察する。さらに、元ドメインと目標ドメインでデータ構造が全く異なる状況で、知識の転移を可能にするアルゴリズムを作成する。ここでは元ドメインが複数ある状況も許容し、元ドメインにおけるデータは画像、動画、自然言語などデータタイプが異なっているとする。各々の元ドメインから抽出された特徴をまとめる

共通空間を構築し,同質的な状況に帰着させるアルゴリズムを作成する.ここで,共通空間としてはいくつかの候補が考えられる.自然言語は word2vec のような手法でベクトル化できるので,ベクトル空間が共通空間の一つの候補である.一方で,人間が画像や動画を自然言語で表現し,知識を統合できることを鑑みて,脳構造を模倣したグラフィカルモデルも共通空間の候補として想定できる.どのような共通空間の構造を仮定するかは,研究の進展と実装した場合の性能を比較し決定する.

4. 研究成果

第一の主結果として,パラメトリックなモデルをドメイン間で転移する状況における理論バウンズを導出した.特に特徴抽出器として複雑なパラメトリックモデルを取ることでもでき,深層ニューラルネットやスパース符号化を理論的に取り扱うことができるものである.第二の主結果として,同変性とよばれる代数的な性質を加味した場合のメタ学習器の普遍性を示した.同変性はデータ処理や自然界の過程で自然に現れる性質であり,同変的なニューラルアーキテクチャによる処理により学習を効率化させることができるという利点がある.第三の主結果として,同時分布に関する期待リスクの差の分解定理を導出した.

[1] Wataru Kumagai, "Learning Bound for Parameter Transfer Learning," The Thirtieth Annual Conference on Neural Information Processing Systems (NIPS), 2016.

[2] R. Raina, A. Battle, H. Lee, B. Packer, and A. Y. Ng, "Self-taught learning: transfer learning from unlabeled data," ICML, 2007.

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 0件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Wada Yuichiro, Miyamoto Shugo, Nakagama Takumi, Andeol Leo, Kumagai Wataru, Kanamori Takafumi	4. 巻 21
2. 論文標題 Spectral Embedded Deep Clustering	5. 発行年 2019年
3. 雑誌名 Entropy	6. 最初と最後の頁 795 ~ 795
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/e21080795	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 WADA Yuichiro, SU Siqiang, KUMAGAI Wataru, KANAMORI Takafumi	4. 巻 E102.D
2. 論文標題 Robust Label Prediction via Label Propagation and Geodesic k-Nearest Neighbor in Online Semi-Supervised Learning	5. 発行年 2019年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 1537 ~ 1545
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2018EDP7424	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 Matsui Kota, Kumagai Wataru, Kanamori Kenta, Nishikimi Mitsuaki, Kanamori Takafumi	4. 巻 31
2. 論文標題 Variable Selection for Nonparametric Learning with Power Series Kernels	5. 発行年 2019年
3. 雑誌名 Neural Computation	6. 最初と最後の頁 1718 ~ 1750
掲載論文のDOI（デジタルオブジェクト識別子） 10.1162/neco_a_01212	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 Kumagai Wataru, Kanamori Takafumi	4. 巻 108
2. 論文標題 Risk bound of transfer learning using parametric feature mapping and its application to sparse coding	5. 発行年 2019年
3. 雑誌名 Machine Learning	6. 最初と最後の頁 1975 ~ 2008
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s10994-019-05805-2	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 -

〔学会発表〕 計7件（うち招待講演 2件 / うち国際学会 0件）

1. 発表者名 熊谷 亘
2. 発表標題 転移学習から継続学習へ
3. 学会等名 科学研究費 基盤研究(A)「機械学習システムの社会実装に向けた次世代最適化技法の研究」による2019年度ワークショップ（招待講演）
4. 発表年 2019年

1. 発表者名 熊谷 亘
2. 発表標題 転移学習から継続学習へ
3. 学会等名 情報系 WINTER FESTA Episode 5
4. 発表年 2019年

1. 発表者名 熊谷 亘
2. 発表標題 Wasserstein距離を用いた転移学習の理論解析
3. 学会等名 第21回情報論的学習理論ワークショップ
4. 発表年 2018年

1. 発表者名 熊谷 亘
2. 発表標題 転移学習の数理
3. 学会等名 第二回理研AIP数理系合同合宿（招待講演）
4. 発表年 2018年

1. 発表者名 熊谷 亘
2. 発表標題 パラメータ転移学習における汎化誤差の評価
3. 学会等名 科研費シンポジウム「統計学，機械学習の数理とその応用」
4. 発表年 2017年

1. 発表者名 熊谷 亘
2. 発表標題 パラメータ転移学習におけるリスク上界
3. 学会等名 統計学・機械学習若手シンポジウム「大規模複雑データに対する統計・機械学習のアプローチ」
4. 発表年 2017年

1. 発表者名 熊谷 亘
2. 発表標題 パラメータ転移学習におけるリスク上界
3. 学会等名 2017年度 統計関連学会連合大会
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
---------	---------