

令和元年6月10日現在

機関番号：10101

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K05264

研究課題名(和文) マルチカーネル学習の新展開

研究課題名(英文) New framework of multiple kernel learning

研究代表者

田中 章 (Tanaka, Akira)

北海道大学・情報科学研究科・教授

研究者番号：20332471

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：カーネルに基づく機械学習の分野において、主に、以下の二つの成果を得た。

- 1.従来のマルチカーネル学習の理論限界に関する理論解析を行い、通常の2-ノルムに基づく最適化基準によって得られるマルチカーネル学習の最適解は、マルチカーネルモデルの理論限界を一般に達成できないこと、及び、当該最適解は、ある種の単一カーネルを用いた学習結果と一致することを明らかにした。
- 2.単一カーネルを用いた学習において、推定対象である未知関数の自己相関関数を有効に利用する枠組みを構築し、汎化性能を向上させることが可能であること、及び、推定対象である未知関数の自己相関関数が期待値の意味で最適なカーネルとなることを明らかにした。

研究成果の学術的意義や社会的意義

現代社会において、機械学習理論が果たす役割がますます増大するであろうことは最早論を俟たない。一方、それら技術の実用面での爆発的な発展とは裏腹に、理論的な背景や基盤が置き去りにされていることは、当該分野の未来にとって由々しき事態であると言わざるを得ない。本研究課題で得た成果は、主に、機械学習理論における理論整備として位置づけられるものであり、当該分野の科学的な側面を補強する一助になるものと期待できる。

研究成果の概要(英文)：We obtained the following two results in kernel-based machine learning.

1. We analyzed the theoretical limit of conventional multiple kernel learning, and clarified that (a) the optimal solution based on an ordinary 2-norm criterion can not achieve the theoretical limit of the multiple kernel model, and (b) the optimal solution is identical to the solution based on a certain single kernel model.
2. We developed a novel framework of kernel-based learning in which the autocorrelation prior of the unknown true function is taken into account, and also clarified that the autocorrelation function itself is the best kernel in terms of the expectation.

研究分野：機械学習

キーワード：機械学習 カーネル回帰 再生核 再生核ヒルベルト空間 汎化性能

## 様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

### 1. 研究開始当初の背景

計算機性能の劇的な向上を背景に、機械学習理論の成果が実用的なレベルで利用される局面が増大していることは周知の事実である。機械学習理論には、様々なアプローチが存在するが、中でも、再生核ヒルベルト空間論に基礎を置く、カーネルトリックと呼ばれる方法論に基づく学習理論は、非常に有望な分野の一つとして広く認知されている。従来、カーネルトリックに基づく学習理論においては、カーネル関数は、システムの入力を(無限次元を含む)高次元の特徴空間へ写像するとともに、当該特徴空間における内積を規定し算出する道具立てとして認知されており、故に、機械学習問題の定式化に内包されない恣意的な道具立てである、と解釈する立場が支配的であった。それ故、応用面での発展とは裏腹に、その最適性や理論的妥当性は置き去りにされている感が拭えなかった。それに対し、小川らの研究グループは、機械学習の問題が、一般標本化定理が扱う問題と本質的に同一であることを指摘し、この枠組による射影学習という概念を提唱した。この概念に基づく解釈により、カーネル関数が、未知の推定対象が属する関数空間を規定し、かつ、解空間上の最適解、すなわち未知の推定対象の、解空間への直交射影を与える際の計量をも規定するものとして、機械学習問題の定式化に内包されるのみならず、解の最適性に大きく影響を与えるものであることを理論的に明らかにした。これらの成果は、カーネル関数に基づく機械学習問題の、最も重要なマイルストーンの一つと言える。

ところで、近年では再生核ヒルベルト空間論に基づく学習法の中で、複数のカーネル関数を利用する学習法、いわゆる「マルチカーネル学習」が注目されている。これは、複数のカーネル関数を利用できるため、潜在的に自由度の高いモデルを構築可能である一方、主に実験的な立場でのみ、その有用性を支持する研究がほとんどであり、その理論的な背景について顧みられることはほとんどなかった。他方、申請者を含む北海道大学のグループは、小川らの成果に立脚し、その理論的な解析を行い、結果として、通常用いられる経験誤差最小化の枠組みでは最適解は得られず、更に訓練データのみから最適解を構成することも不可能であることを示した。加えて、従来同様、複数のカーネル関数の線形結合で推定対象を表現するため、劇的な性能向上を期待することは難しく、これらがマルチカーネル学習の枠組みを根本から見直す強い動機を与えていた。

### 2. 研究の目的

本研究では、マルチカーネル学習の新たな枠組みの構築を目的として設定した。個々のカーネルに対して、未知推定対象の解空間への直交射影を得ることが可能であることは先に述べたとおりであるが、当該最適解においては、射影により消失する成分が存在する。すなわち、真の推定対象は、当該最適解に(未知である)解空間の直交補空間成分を付加したものとなる。単一のカーネルのみを利用する場合は、当該直交補空間成分を推定することは不可能であるが、複数のカーネルを利用可能な場合、この構造が複数存在するため、必然的に直交補空間成分に対する拘束を生み出す。よって、当該拘束条件を満足する解を選択すれば、推定精度の大幅な向上を期待できるというのが本研究の着想の核心であった。なお、上で述べた機序から、当該枠組みを「射影バイアス推定」と呼称する。以上をまとめると、本研究期間中に明らかにすることは、主に、

- 1) 「射影バイアス推定」に基づく、マルチカーネル学習の理論的枠組みの構築。
- 2) 当該枠組みの、実用へ向けた展開。

の二点である。これらの課題の遂行により、再生核ヒルベルト空間論に基づく機械学習の大幅な性能向上や適用範囲の拡大が期待できる。

### 3. 研究の方法

本研究の目標を完遂することとは、すなわち「研究目的」の項で述べた項目の各々に対し、主に数学的な定理を証明するという形で決着させることである。これを実現するために実行可能なことは、これまで蓄積した数学的知見を駆使することの他は、効果的に情報を収集する以外にない。研究期間中は、学術的な会議への出席や、他機関の研究者との研究討論を通じて効果的に情報を収集し、目標の完遂を目指して研究活動を行った。

他方、本研究課題の主要な目的である、「射影バイアス推定」に基づく、マルチカーネル学習の理論的枠組みの構築については、二年目の段階で実施した理論検討の結果、複数のカーネルを用いた場合でも、単一のカーネルを用いた場合以上の拘束条件が得られないことが理論的に明らかとなったため、以降は、軌道修正し、単一カーネルを用いた場合の学習法の構築や、最適カーネルの性質についての理論解析を行うこととした。

### 4. 研究成果

#### (1) 従来のマルチカーネル学習に関する理論解析

研究課題の遂行に先立ち、従来のマルチカーネル学習の理論限界について明らかにすべく、理論解析を行った。その結果、通常用いられる2-ノルムに基づく最適化基準の最小化によって得られるマルチカーネル学習の最適解は、マルチカーネルによる関数モデルの潜在表現力の理論限界を一般には達成できないことが明らかとなった。加えて、当該最適解は、マルチカーネル学習で用いられる複数カーネルの線形結合によって得られる単一カーネルを用いた学習結果と一致することを明らかにした。(主な発表論文等〔雑誌論文〕の )

#### (2) 未知関数の自己相関を用いた学習

単一カーネルを用いた学習において、推定対象である未知関数の自己相関関数が利用可能である場合に、当該情報を有効に利用して学習を行う枠組みを構築し、汎化性能を向上させることが可能であることを示した。(主な発表論文等〔学会発表〕の )

その後、推定対象である未知関数の自己相関関数そのものをカーネルとして利用する学習についても検討を行い、当該自己相関関数が期待値の意味で最適なカーネルとなることを明らかにした。(主な発表論文等〔学会発表〕の )

#### (3) その他

再生核ヒルベルト空間論、及び、再生核ヒルベルト空間論と密接に関連する、標本化理論について以下の成果を得た。

- a. 複数の再生核ヒルベルト空間の積集合として得られる再生核ヒルベルト空間に対応するカーネルの近似計算法を開発した。(主な発表論文等〔学会発表〕の )
- b. 移動不変な再生核ヒルベルト空間において、等間隔標本点による標本化定理が成立するための、必要十分条件として新たな判別式を発見した。(主な発表論文等〔学会発表〕の )
- c. 再生核ヒルベルト空間に基づく標本化理論に関する解説論文を執筆した。(主な発表論文等〔雑誌論文〕の )

カーネルを用いた関数推定の応用事例として、鉄鋳炉中の投入材料表面形状の推定法を開発した。(主な発表論文等〔雑誌論文〕の )

## 5 . 主な発表論文等

### 〔雑誌論文〕(計3件)

J. Tian, A. Tanaka, Q. Hou, and X. Chen, "Radar Detection-based Modeling in a Blast Furnace: a Prediction Model of Burden Surface Shape after Charging", ISIJ International, Vol.58, No.11, pp.1999-2008, 2018. (査読有)

田中章, 「再生核ヒルベルト空間から眺める標本化定理」, 日本音響学会誌, Vol.73, No.9, pp.577-584, Sep. 2017. (査読無)

A. Tanaka, H. Imai, "Theoretical Analyses on 2-Norm-Based Multiple Kernel Regressors", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E100-A, No.3, pp.877-887, Mar. 2017. (査読有)

### 〔学会発表〕(計4件)

Akira Tanaka: "Kernel-Induced Sampling Theorem for Translation-Invariant Reproducing Kernel Hilbert Spaces with Uniform Sampling", 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2018), Calgary, Canada (2018,4), pp.4554-4558

Yuuki Saito, Akira Tanaka: "Optimal Kernel in Kernel Regression Problems with Autocorrelation Prior", Proceedings of APSIPA Annual Summit & Conference 2017, Kuala Lumpur, Malaysia, (2017,12), Paper ID:78

Akira Tanaka: "Kernel Regression with Autocorrelation Prior", 12th International Conference on Sampling Theory and Applications, Tallinn, Estonia (2017.7), pp.318-321

田中章, 「再生核ヒルベルト空間の積空間におけるグラム行列の近似計算」, 第31回信号処理シンポジウム, 吹田 (2016.11), 講演論文集, pp.405-408

### 〔図書〕(計0件)

### 〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

### 〔その他〕

ホームページ等

<http://ims2.main.ist.hokudai.ac.jp>

## 6 . 研究組織

(1)研究分担者 : なし

(2)研究協力者 : なし

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。