

令和 6 年 6 月 13 日現在

機関番号：32641

研究種目：若手研究

研究期間：2022～2023

課題番号：22K17860

研究課題名（和文）サポートベクター回帰のハイパーパラメータの新たな推定

研究課題名（英文）New estimation of hyperparameters for support vector regression

研究代表者

牧草 夏実 (Makigusa, Natsumi)

中央大学・理工学部・助教

研究者番号：60908341

交付決定額（研究期間全体）：（直接経費） 1,800,000円

研究成果の概要（和文）：本研究では、特にカーネル法を用いた手法の一つであるサポートベクター回帰に絞って、パラメータの推定の問題に取り組んだ。特に非線形サポートベクター回帰のパラメータを、罰則付き尤度を用いることによって推定するための方法を構築し、その理論を確立するために、いくつかの補題の導出を行った。また、実データへの適用を通して、実際のデータへの応用として十分利用可能な手法であることが確認された。理論研究のために、サポートベクター回帰にとどまらない再生核ヒルベルト空間に関する研究も行った。本研究に関連し期間中に7件の学会発表および1本の論文が出版された。科研費の支援のもと、成果を得ることができた。

研究成果の学術的意義や社会的意義

非線形サポートベクター回帰のパラメータを決定するにあたってよく知られている方法としては、クロスバリデーション法と呼ばれる方法である。この方法では、候補となるあらゆるパラメータでの回帰を行い、その回帰についての評価を行う。あらゆる組み合わせにおける回帰を計算する必要があるため、非線形サポートベクター回帰のような選択すべきパラメータが多い手法において計算量が非常に多くなるという問題点が挙げられていた。このため、本研究の成果はこの問題点を解決する手法の一つとなっている。さらに、サポートベクター回帰のような機械学習手法はAI技術としても用いられており、昨今の科学技術発展につながる。

研究成果の概要（英文）：In this study, we focused on the problem of parameter estimation in one of the kernel methods, specifically Support Vector Regression (SVR). We developed a method to estimate the parameters of nonlinear Support Vector Regression using penalized likelihood and derived several lemmas to establish the theoretical foundation. Additionally, through the application to real datasets, we confirmed that the method is sufficiently practical for actual data applications. For the sake of theoretical research, we also conducted studies related to Reproducing Kernel Hilbert Spaces (RKHS) beyond just Support Vector Regression. During the research period, we made seven conference presentations and had one paper accepted and published. With the support of the Grants-in-Aid for Scientific Research (KAKENHI), we were able to achieve these results.

研究分野：数理統計学

キーワード：カーネル法 サポートベクター回帰 漸近理論 再生核ヒルベルト空間 機械学習

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

カーネル法は、高次元データの取り扱いや非線形化の可能な、非常に有用なデータ解析手法である。パラメトリックモデルのような分布や回帰関数の指定は不要だが、カーネルの選択や、事前に指定するパラメータ(ハイパーパラメータ)の推定が必要となる。回帰や分類を行う手法では、クロスバリデーションという方法があり、カーネルとハイパーパラメータに対する評価値を計算することで、候補の中から最適なものを選択することが可能である。その中でハイパーパラメータはモデルの性能に大きく影響を与えるため、最適なハイパーパラメータを推定できるように候補点の幅を広げなければならない。しかし、カーネルやハイパーパラメータの候補が多くなれば、その分組み合わせが増え、計算コストが増大することが問題点となっていた。このハイパーパラメータの推定について以下のテーマに取り組んだ。

テーマ「ハイパーパラメータを含めた推定法の構築」では、種々のカーネル法を用いた手法の中で、既存研究での理論的結果を用いることが容易であるため、特にサポートベクター回帰に焦点を当ててこの研究に取り組んだ。従来の手法では最適化問題により回帰関数のみ推定していたサポートベクター回帰であるが、回帰関数とハイパーパラメータ(γ , C)を同時に最適化することができないだろうかという学問的問いに取り組んだ。また、この手法により推定されたパラメータの理論的評価のため、関連する研究調査に取り組んだ。

2. 研究の目的

本研究の目的はサポートベクター回帰におけるハイパーパラメータを含めた推定法の構築として定めていた。特に、従来の最適化問題では事前に与えていた γ や C といったパラメータも最適化問題の最適化したい変数として考えることで、回帰関数とハイパーパラメータを一挙に推定することであり、この最適化問題を解くことを具体的な目的と定めた。

3. 研究の方法

サポートベクター回帰のハイパーパラメータを回帰関数と同時に最適化問題から推定することが目的であるが、線形の場合は、ある特定の条件付き密度関数を持つと仮定し、その仮定の下での最尤推定量が考えられている。ただし、負の対数尤度を最小にする問題では過適合が起こる、解が求まらない等の問題を考慮し、罰則付き負の対数尤度を最小にする方法が考えられている。この最小化問題の式が、一般的なサポートベクター回帰の最小化問題と非常に似通っているため、罰則付き対数尤度による最適化問題を解くことにより、サポートベクター回帰のハイパーパラメータも最適化する手法になっている。本研究は、この研究の非線形化に対応しており、同様に罰則付き対数尤度を用いることにより、サポートベクター回帰のハイパーパラメータの推定を行った。したがって、回帰関数とハイパーパラメータの推定量というのは、この罰則付き対数尤度の最小解である。この推定量が、真のパラメータ、すなわち条件付き密度関数に現れているパラメータに確率収束すること、さらに真のパラメータへの収束の様子が正規分布に従っている、すなわち漸近正規性を持つことについて調べた。

確率収束すること，漸近正規性を持つことについては，線形の場合に帰着させることが困難であり，線形の場合だけでなく，サポートベクターマシンの推定量に関しての漸近正規性を示している先行研究を参考にした．ここでは，関数デルタ法を用いることにより，回帰関数の推定量に関する漸近正規性を示しており，この手法を同様に用いるために，罰則付き対数尤度関数に関してのフレッシュ微分(またはアダマール微分)可能性を調べ，推定量の漸近正規性の導出に取り組んだ．

4．研究成果

(1) サポートベクター回帰におけるハイパーパラメータを含めた推定法の構築

本研究では，特に非線形をサポートベクター回帰におけるハイパーパラメータを含めた推定法として，ハイパーパラメータ C と c を回帰関数 f , b と同時に推定することに取り組んだ．ある特定の条件付き密度関数を持つと仮定し，その仮定の下での最尤推定量が考えられ，罰則付き負の対数尤度を最小にするパラメータを考えると，既存のサポートベクター回帰の最小化問題に帰着され，ハイパーパラメータ C の最小化と，回帰関数 f と b についての最小化を繰り返すことにより，最適化問題を解くことができることが得られた．線形の場合のサポートベクター回帰において，同様の結果が既に報告されており，本研究の成果は線形の場合のサポートベクター回帰の拡張に対応している．

(2) 理論的結果

(1) で述べた推定量についての理論的結果として，推定量の漸近正規性について取り組んだ．ハイパーパラメータを含めない場合のサポートベクターマシンの推定量に関しての漸近正規性についてはすでに報告されており，その導出には関数デルタ法が用いられていた．我々の考えている推定量は，ハイパーパラメータを同時に最適化するためユークリッド空間と再生核ヒルベルト空間の直和空間の元として考える必要はあるが，直和空間もまた再生核ヒルベルト空間であるため，先行研究と同様に関数デルタ法を用いることで導出する方針を立てた．そのために必要な補題として，最適化関数がフレッシュ微分可能であることを導出した．

本研究と関連して，再生核ヒルベルト空間に関する論文 1 編が出版された．

(3) シミュレーションおよび実データへの適用結果

シミュレーションを通じて，推定量の漸近挙動について調べた結果，サンプル数の増加とともに回帰関数 f が真の関数に収束していく様子がみられた．さらに，各点に対しての挙動についても確認したところ，回帰関数 f は各点において漸近正規性を持っている様子が見られた．

さらに，実データを用いて，ハイパーパラメータの決定手法として広く知られているクロスバリデーション法との比較を行った．その結果，提案手法でのハイパーパラメータの決定はクロスバリデーションでのパラメータの決定と同等の精度を出すことができ，より少ない計算コストで実行することが可能であることがわかった．

本研究に関連し期間中に 7 件の学会発表および 1 本の論文が出版された．科研費の支援のもと，成果を得ることができた．

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1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 Natsumi Makigusa	4. 巻 220851
2. 論文標題 Two-sample test based on maximum variance discrepancy	5. 発行年 2023年
3. 雑誌名 Communications in Statistics - Theory and Methods	6. 最初と最後の頁 1-18
掲載論文のDOI（デジタルオブジェクト識別子） 10.1080/03610926.2023.2220851	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 -

〔学会発表〕 計7件（うち招待講演 1件/うち国際学会 2件）

1. 発表者名 Natsumi Makigusa, Kanta Naito
2. 発表標題 Nonlinear support vector regression with penalized likelihood
3. 学会等名 IMS Asia Pacific Rim Meeting 2024（国際学会）
4. 発表年 2024年

1. 発表者名 牧草 夏実
2. 発表標題 罰則付き尤度に基づく非線形サポートベクター回帰
3. 学会等名 九州大学統計科学セミナー
4. 発表年 2023年

1. 発表者名 牧草夏実, 内藤貴太
2. 発表標題 非線形サポートベクター回帰のハイパーパラメータの決定手法
3. 学会等名 日本計算機統計学会 第37回シンポジウム
4. 発表年 2023年

1. 発表者名 牧草夏実, 内藤貫太
2. 発表標題 非線形サポートベクター回帰のハイパーパラメータの推定
3. 学会等名 2023年度統計関連学会連合大会
4. 発表年 2023年

1. 発表者名 Natsumi Makigusa
2. 発表標題 Two-sample test based on the variance of a positive definite kernel
3. 学会等名 25th International Conference on Computational Statistics (国際学会)
4. 発表年 2023年

1. 発表者名 牧草夏実
2. 発表標題 正定値カーネルの分散を用いた二標本検定
3. 学会等名 日本計算機統計学会 第37回大会
4. 発表年 2023年

1. 発表者名 Natsumi Makigusa
2. 発表標題 Kernel Method and its Application
3. 学会等名 RIMS研究集会：再生核ヒルベルト空間を中心とした実解析・複素解析・函数解析の総合的研究（招待講演）
4. 発表年 2022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
---------	---------