

氏 名 沈 迅

学位(専攻分野) 博士(統計科学)

学位記番号 総研大甲第 2321 号

学位授与の日付 2022 年 3 月 24 日

学位授与の要件 複合科学研究科 統計科学専攻
学位規則第6条第1項該当

学位論文題目 Residual Analysis for Machine Learning

論文審査委員 主 査 中野 慎也
統計科学専攻 准教授
庄 建倉
統計科学専攻 准教授
矢野 恵佑
統計科学専攻 准教授
野村 俊一
早稲田大学 商学学院 准教授

(様式3)

博士論文の要旨

氏名 沈 迅

論文題目 Residual Analysis for Machine Learning

In this thesis, we consider the problem of applying residual analysis to improve machine learning algorithms, including model improvement for state-space models with nonlinear response (Chapter 3), approximate approach for solving chance constrained optimization (Chapter 4), and computing the probabilistic bounds on state trajectories for uncertain nonlinear systems (Chapter 5).

In Chapter 2, we summarize the original concept of using residual analysis to check the goodness of fit and improve model. The example of residual analysis in regression problems is used to help understanding the concept. Besides, some recent applications of residual analysis have been reviewed, including applications in point processes, time series models, and hidden Markov models. At the end of Chapter 2, we give an intuitive introduction for the idea of applying residual analysis in model improvement of state space models and approximation of chance constrained optimizations, which are the main contributions of this thesis.

In Chapter 3, we address the problem of model improvement for state space models based on residual analysis. The residual of a state-space model is defined. For an SSM with unknown nonlinear response, we propose a novel algorithm for model learning and hidden state inference. A neural network model is used to approximate the unknown nonlinear part in the observation equation, and an Expectation-Maximization (EM) algorithm is proposed to infer the hidden state and learn the parameters in both the linear part and the neural network model, from the given sequences of input data and observation data. In the E-Step, the posterior mean and covariance for the system hidden state given the sequences of the system input and observations is inferred via a Kalman filter-based forward recursion and Rauch-Tung-Streifel smoother backward recursion. In the M-Step, the model parameters are optimized according to the inferred hidden state, input data, and observation data. The M-Step consists of two components: a reconstruction procedure, in which uses the residuals of the linear model to fit the neural network model, and a parametrization procedure, which identifies the parameters in the linear part of the state space model. We apply this newly proposed method to a numerical example and in a case study of battery capacity estimation. The results show that the proposed method can achieve better performance on the model learning and hidden state inference than previously developed tools.

In Chapter 4, we present the residual analysis-based algorithm design for approximate chance constrained program. After reformulating the probabilistic

constraints as the quantile function, a sample-based neural network model is used to approximate the quantile function which can dramatically improve the efficiency of the algorithm. The statistical guarantees of the neural approximation are discussed by showing the convergence of the approximate residual and feasibility analysis. Interval Predictor Model (IPM) of wind power is investigated to validate the proposed method.

In Chapter 5, we present the chance constrained optimization-based algorithms to compute predictive probabilistic bounds on state trajectories for uncertain nonlinear systems. A probabilistic constrained problem is formulated for calculating the probabilistic ellipsoidal bounds of the future trajectory of system states. Scenario approach and sample average approach are used to approximate the probabilistic constrained problem by formulating a deterministic problem with samples of the uncertain parameters in the system. For a given probabilistic level and upper bound of the violation probability, the least number of samples required for calculating the bound can be determined for both scenario approach and sample average approach. The optimality of the solutions obtained by scenario approach and sample average approach is discussed theoretically. The results of numerical example show that more samples will improve the feasibility by sacrificing optimality slightly for sample average approach. However the scenario approach needs more sacrifice on optimality.

博士論文審査結果

Name in Full
氏名 沈 迅

Title
論文題目 Residual Analysis for Machine Learning

[論文の概要]

提出された論文は、残差分析 (Residual analysis) 手法を適用し、状態空間モデルおよび確率制約付き最適化の方法論を改良する方法を提案したものである。英文で書かれており、全 6 章と附録、参考文献を合わせて 113 頁からなる。

第 1 章は、研究の背景と動機、および本論文全体の構成についての説明である。

第 2 章では、残差分析に関する最近の応用事例を概観し、さらに残差分析を応用した状態空間モデルの改良、および確率制約付き最適化問題の近似方法についてのアイデアが提示されている。

第 3 章では、観測がシステムの状態に対して未知の非線形応答を持つ場合を考慮して状態空間モデルを拡張し、モデルパラメータ、隠れ状態と非線形応答関数を同時に推定する新しいアルゴリズムを提案している。アルゴリズムの具体的な構造は、カルマンフィルタを使用した隠れ状態の推定、最尤法に基づく線形部分のモデルパラメータの推定、線形モデルの予測と観測との間の残差に基づく非線形成分の推定を、期待値最大化 (EM) アルゴリズムに類似した繰り返しアルゴリズムで行うというものである。非線形応答成分は、ニューラルネットワークモデルで表現して推定がなされる。ニューラルネットワークの学習は、全結合係数を推定する方法でも実現可能だが、過剰適合が避けられ、かつ効率的な方法として、出力層のみを学習する extreme learning machine モデルの採用を提案しており、数値実験で有効性を示している。また、応用例としてバッテリー容量の推定に適用し、他手法 (unscented Kalman filter、unscented Kalman filter with neural network approximation) と比較して提案手法が優位であることも示している。

第 4 章では、残差分析に基づく新たな確率制約付き最適化手法を提案している。既存の手法として、シナリオアプローチ、サンプル平均アプローチなどがあるが、前者は保守的な近似に基づいているため最適解が得られる保証がなく、後者には計算上の困難さがある。提案手法では、確率的制約を分位関数で再定式化した後、extreme learning machine モデルを用いたサンプルベースの手法で分位関数を近似して解くことにより、サンプル平均アプローチなどの既存手法と比較してアルゴリズムの効率を大幅に向上させることに成功した。また、最適解への収束性も確認している。簡単な数値実験の他、風力発電の電力予測問題に適用し、提案手法の有効性も確認している。

第 5 章では、確率制約付き最適化手法が、非線形システムにおける解軌道の確率的 outer bound を求める問題に適用できることを議論し、シナリオアプローチ、サンプル平均アプローチのそれぞれに基づいたアルゴリズムを提示している。数値実験でそれぞれのアルゴリズムの結果とモンテカルロ計算との結果を比較し、その有効性も確認しており、特にサンプル平均アプローチではサンプルサイズを十分に取ることで適切な解が得られることを示している。

第 6 章では、今後の展望を議論しており、附録では extreme learning machine の説明が補足されている。

[論文の評価]

残差分析の理論に基づいて、非線形応答を含む状態空間モデルの推定、および確率的制約付き最適化という実用上も重要な問題に取り組んでいる。いずれに研究においても、**extreme learning machine** モデルを用いて効率的な計算を実現し、実用的な問題に適用することで手法の有用性も具体的に示されており、高く評価できる。特に、確率制約付きの最適化問題においては、最適解への収束性を保証しながら先行研究よりも高い計算効率を実現し、さらに非線形システムの予測への適用可能性も提示しており、重要な貢献として高く評価できる。以上の理由により、審査委員会は、本論文が学位の授与に値すると判断した。

なお、本論文の第 4 章前半部分は IEEE Access に、後半部分は IEEE Transactions on Neural Networks and Learning Systems 誌にそれぞれ掲載が決まっており、すでにオンラインで公開されている。第 5 章前半部分の内容は IEEE Transactions on Emerging Topics in Computational Intelligence 誌に掲載決定、公開済みである。また、関連する研究が IEEE/ASME Transactions on Mechatronics 誌に掲載決定、公開済みである。第 3 章の内容は Journal of Computational and Graphical Statistics 誌に投稿中である。