

論文の要旨

題目 ランダムフォレストを用いた非線形システムのモデリングとその応用
(Modeling Method of Nonlinear Systems Using Random Forest and Their Applications)

氏名 今地 大武

産業界のシステムの多くは多入力であり、さらに非線形かつ動的な特性、および時間変化する内部状態を有しているため、モデリング技術には、これらの特徴を考慮することが求められている。モデリング技術はホワイトボックスモデリングとブラックボックスモデリングに大別される。前者はシステムの物理的特性から、数式によってモデルを構築する手法である。一方、後者は、システムの構造が不明であり、入出力信号のみが与えられている場合に、入出力関係に着目し汎用的なモデルを構築する手法である。一般的に、産業システムは複雑な数学的構造を有しているため、ホワイトボックスモデリングの適用は困難である。したがって、本論文ではブラックボックスモデリングを取り扱う。

ブラックボックスモデリングを実現する手法として、機械学習技術を用いる方法がある。この方法は、システムの入出力関係をヒューリスティックな手法に基づいてモデリングする方法である。機械学習を用いた代表的な手法として、ディープラーニングを用いる手法が挙げられる。ディープラーニングは、ニューラルネットワークの一種であり、複数の事前学習器を重ねた構造をもつ。これにより、高レベルの学習が実現できるため、複雑なシステムのモデリングが可能となる。一方で、設計が難しく、かつモデルの構造解釈が極めて困難であるため、ディープラーニングは、画像処理のように内部構造の解析が必須ではないシステムのモデリングには有効ではあるが、本論文で取り扱う産業システムのように、安全面考慮やインシデント対応次第で生成されたモデルの解析が必要になる場合には不向きであると考えられる。

一方、比較的構造解釈の容易な機械学習手法として、決定木が提案されている。決定木はその構造の解釈の容易さから、モデリングのほかにもシステムの解析に用いられる。しかしながら、過学習を起しやすく、外れ値やノイズの多いデータを用いたモデリングには不向きであると考えられる。ところで、ランダムフォレスト(Random Forest: RF)と呼ばれる複数の決定木を統合することで過学習を抑制する手法が提案されている。この手法は、先述のディープラーニングに比べると、一般的に構造解釈が容易な一方で、モデルの表現能力は劣るが、本研究で対象とする産業システムへの適用としては、十分な性能を発揮できると考えられる。したがって、本手法の特性を利用することにより、非線形性の考慮と、構造解釈の容易さ確保の両立を目指す。

先述のとおり、RFは非線形性と構造解釈の容易さの両立が可能な一方で、システムの時間変化する特性を取り扱うことが難しい。これは、RFが自己回帰項などの過去のシステムの情報を取り入れることが難しいためである。したがって、本論文では、時間変化する特性を取り入れること

が可能な手法に対して、RFにより解析機能を付与することで、時間変化する特性の解析手法の確立を目指す。

本論文では、非線形システムの動的特性を考慮したモデリングと解析を実現するために、2種類の手法を考える。

一つは、まずシステムの状態を複数定義し、システムの時間変化を状態遷移モデルとして表すことで、時間変化する特性をRFによって解析・モデリングする方法である。具体的には、RFの内部構造を利用してシステムの状態を定量化し、その結果から状態遷移時に表れる特徴を解析する手法である。本手法は油圧シャベルによる土砂のダンプ積み込み作業データにより実証実験を行う。

もう一つは、RFとシステムの動的特性を考慮可能な手法を組み合わせる手法である。本手法によってシステムのモデリングに必要なデータを解析し、その結果をモデリングに利用する。本手法は自動車の線形空燃比センサの出力予測により実証実験を行う。

本論文では、動的な特性を考慮でき、かつ解析への応用が容易な非線形システムモデリングの実現を目的に、先述のRFを中心とした2つの手法について考察する。

まず、第2章では、本論文の中心となる技術であるRFについて述べる。RFは機械学習の一種であり、解析が比較的容易な決定木と呼ばれる学習機によって構成されているため、構造解釈の容易さを前提としたモデリングが可能である。前半で決定木の構造とアルゴリズムについて述べたのち、RFについて解説する。さらに、RFを用いた説明変数の重要度算出についても述べる。

第3章では、RFによる状態の定量化と、状態遷移モデルを用いた解析手法について述べる。一般的に、RFは時系列データを取り扱うことが難しいため、RFの構造を状態遷移モデルに落とし込むことによって、RFによる時系列データの解析を試みる。本章の後半は、油圧シャベルの掘削作業の解析を通して、本手法の有効性を検証する。作業員の操縦する油圧シャベルを解析対象のシステムとして、各作業工程をシステムの状態とみなすことで本手法を適用し、作業員のスキルを定量的に評価する。

第4章では、RFによる解析に加えて、時系列データのモデリングについて述べる。動的特性に特徴のあるシステムに対して、システムの動的特性と非線形性を考慮可能なデータベース駆動型モデリング(DDM)に、RFの解析機能を搭載することにより、解析とモデリングを同時に行える手法を提案する。本手法により、モデルの精度向上が期待できる。なお、本手法は自動車の線形空燃比センサ(LAFS)の出力予測に適用することにより、有効性について検証する。LAFSはエンジンの燃焼を制御する上で重要な役割を担っているが、自動車の走行状態や経年劣化によって動的特性が変動する。したがって、本章ではRFによって改良したDDMを用いて、動的特性に影響を与えている要因の特定、動的特性の推定、本来のLAFS値の推定を行う。

第5章では本研究の総括をするとともに、残された課題点について言及し、今後の解析可能なモデリングの展望について述べる。