

計測・制御セレクションシリーズ企画書:

深層常微分方程式

京都大学医学研究科 岡本有司

2023 年 7 月 24 日

1 タイトル案

深層常微分方程式

2 筆者プロフィール

氏名: 岡本有司

現職: 京都大学医学研究科特定助教

ORCID: 0000-0003-0024-549X

研究領域:

制御理論と深層学習理論の研究者。東京工業大学の博士後期課程及び東大のポスドク時代、システムバイオロジーなどの生体システムに対して制御理論を取り入れる研究を行う。現職では、より複雑なシステムに適用するために深層学習理論を制御理論に取り入れる研究を行っている。制御理論を背景に持ち、深層学習理論のトップジャーナルである NeurIPS に出しているのは珍しい。(NeurIPS は Google Scholar の h5-index ランキングで現在 9 位)

主要論文:

Yuji Okamoto and Kojima, Ryosuke. "Learning deep input-output stable dynamics." Advances in Neural Information Processing Systems 35 (2022): 8187-8198.

3 出版趣旨

本書は、深層学習と常微分方程式 (ODE) の理論を融合した新たな時系列予測手法「ニューラル ODE」を中心に、深層学習を用いた時系列予測手法について解説する。ニューラル ODE は、深層学習を用いた時系列予測で一般的であった離散時間の時系列予測を拡張し、連続時間に対応した時系列予測モデルの学習を可能にした。これにより、安定性などの常微分方程式の特性を利用した時系列予測が可能になる。本書「深層常微分方程式」は、ニューラル ODE を中心とした深層学習および時系列予測理論を一本化して扱い、これらを扱う研究者に対して新たな視点を提供する。

本書は、深層学習を用いた時系列予測手法の基礎理論とそれを用いた応用手法の二部構成となっている。第一部は、さらに離散時間と連続時間の時系列予測手法の 2 つに分けられている。離散時間の時系列予測手法では、大規模言語モデル (LLM) で一般的に用いられている Transformer と深層学習理論の重要課題である初期値依存性に一石を投じた HiPPO について解説する。連続時間の時系列予測手法では、連続時間の予測機に対する学習の問題点である逆伝播の蓄積を解説し、ニューラル ODE による解決策を示す。第二部の応用手法では、確率モデルおよび状態空間への拡張について解説する。確率モデルへの拡張では、差分方程式の視点から現在の生成モデルの主流となりつつある diffusion model の解説を行う。また、状態空間への拡張されたニューラル ODE の学習は、最適制御問題を利用して学習することができることを示す。

4 本書のターゲット

本書の対象読者は、深層学習を用いた時系列システムの研究に従事する学生と研究者すべてになる。現時点、時系列予測手法の学術書は線形モデルを用いた入門書がほとんどであり、最先端の深層学習及び微分方程式の研究に関しては未だ調査論文に留まっている。線形モデルを始めとし、非線形モデル、深層学習モデルへと遷移する時系列予測手法を統合的に扱っている書は未だに存在しない。そのため、本書は幅広い読者層を対象とした時系列予測の入門書であり、最先端の深層学習の研究をまとめた専門書にもなるだろう。

5 内容詳細

以下に、章の構成予定を示す。2-4 章は第一部の基礎理論、5-6 章はその応用手法である。

1 章：はじめに

時系列予測の問題提起と定式化を行います。

1. **時系列データと時系列予測**：時系列データ及び、時系列予測の定義を行います。説明において、近年話題になっている、大規模言語モデル (LLM) を始め、ロボティクス、電力システムにおける実施例を紹介する。
2. **時系列予測と常微分方程式**：時系列予測を行うためのモデル設計方法を導入します。直接的な解モデルの推定、差分方程式の推定、常微分方程式の推定の 3 つの推定方法、および、それぞれの問題点を示す。
3. **本書の構成**：本書の構成の説明します。

2 章：時系予測のための深層学習

以降の章の説明する準備として、深層学習の解説を行います。

1. **機械学習と連続最適化**：既存の機械学習理論とその学習方法を解説する。
2. **深層モデルと誤差逆伝播**：現在の深層学習モデルとその学習方法を解説する。
3. **深層学習の汎化性能**：深層二重降下を中心とした機械学習と深層学習の差異について示す。

3 章：深層差分方程式

現在でも主流に研究されている、離散時間の時系列予測モデルについて紹介する。

1. **線形差分方程式**：時系列予測の基礎となる AR モデルを示す。

2. **スパースな差分方程式の推定**：Dynamic Mode Decomposition(DMD) や非線形化した Sparse identification of non-linear dynamics(SINDy) についての説明する。
3. **深層モデルを用いた差分方程式**：NLP 界隈で用いられていた RNN と、近年の LLM でスタンダードとなった Transformer を紹介する。
4. **深層差分方程式の初期値問題**：深層学習を用いた時系列予測手法の初期値問題について一石を投じた HiPPO について解説する。

4 章：深層常微分方程式

ニューラル ODE とその利用法について紹介する。

1. **ODEsolver と常微分方程式の学習**：常微分方程式とその解法 (ODEsolver) を導入する。さらに、ODEsolver を用いた常微分方程式の学習の問題点である逆伝播の蓄積について示す。
2. **深層学習による解の推定**：解を直接推定する Physics-Informed Neural Network(PINN) を紹介する。
3. **ニューラル常微分方程式**：ニューラル ODE を用いた、深層学習の学習手法では逆伝播の蓄積が行われないことを示す。
4. **機能バイアスを用いた常微分方程式の学習**：ニューラル ODE が直接ベクトル場を推定できることを利用した、安定性などの機能バイアスを保証する学習手法を解説する。

5 章: 確率モデルへの拡張

近年、よく話題に登る生成モデルを時系列予測という観点から解説する。

1. **ベイズ統計と変分推定**：以降の説明をするために、ベイズ統計と変分推定についての簡単に説明する。
2. **離散深層確率モデル**：Stable diffusion など、画像生成で近年標準的に用いられている diffusion model の説明を行う。
3. **連続深層確率モデル**：確率モデルに拡張された NeuralODE とその (因果推定などへの) 利用を示す。

6 章: 深層状態空間モデル

制御理論を導入するための状態空間モデルに対する拡張を示す。

1. **状態空間モデル**：状態空間モデルの簡単な解説を行う。

2. **Deep Kalman Filters** : 確率差分方程式を用いた状態空間モデルである Deep Kalman Filters を解説する。
3. **最適制御と状態空間モデル** : ニューラル ODE の状態空間モデルへの拡張は、最適制御問題となる。ここでは、ニューラル ODE の最適制御問題への変換を行う。

連絡先

電子メール: okamoto.yuji.2c@kyoto-u.ac.jp

住所: 〒606-8303 京都府京都市左京区吉田橘町 5 4

京都大学附属病院 先端医療機器開発・臨床研究センター 408 号室