言語のための状態空間モデルによるハミルトン力学の時系列予測の試み

An Attempt at Time Series Prediction of Hamiltonian Dynamics Using State Space Models for Language Modeling

徐 百歌 (Baige Xu)[1], 谷口 隆晴 (Takaharu Yaguchi)[1,2]
[1] 神戸大学 (Kobe University), [2] 理化学研究所 (RIKEN)
e-mail : baigexu@stu.kobe-u.ac.jp

1 はじめに

近年，状態空間モデル（State Space Models, SSMs）は，大規模言語モデルにおける効率的な時系列処理の枠組みとして注目されている [1, 2, 3]．特に，SSMs は従来のトランスフォーマーより，長い時系列データの処理における計算コスト削減において優れた特性を示している [4, 5]．

ここで，一部の既存の言語モデルは，その内部構造として減衰項を含むハミルトン方程式を用いる状態遷移が導入されており，従来の物理モデリングにおける数理モデルと類似の構造をもつことに着目すると，SSMs を物理系のモデリングに応用することは，自然な拡張と考えられる．そこで，本研究では，このような言語のための状態空間モデルを用いて，ハミルトン系の時間発展を学習と予測手法を試みる．特に，減衰項つきハミルトン方程式を内在化した状態空間構造を活用することで，軌道データからの物理的時系列予測における精度と計算効率の両立を目指す．この方法を発展させることにより，将来的には，言語モデルのために開発された学習済みモデルをファインチューニングによって物理モデルとして利用できる可能性がある．

2 状態空間モデルとハミルトン力学

本研究では，SSMs として，以下の線形時不変系に代表されるような形式で表現される系を考える：

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t). \tag{1}$$

ここで，$x(t) \in \mathbb{R}^{2n}$ は時刻 t における内部状態，$u(t) \in \mathbb{R}^m$ は外部入力，$y(t) \in \mathbb{R}^d$ は観測可能な出力である．状態遷移行列 $A \in \mathbb{R}^{2n\times 2n}$，入力行列 $B \in \mathbb{R}^{2n\times m}$，出力行列 $C \in \mathbb{R}^{d\times 2n}$ は学習可能なパラメータ行列である．

一方，力学系としてハミルトン系は，状態ベクトル $z(t) = (q(t), p(t)) \in \mathbb{R}^{2n}$ に対して，次のような方程式で記述される：

$$\frac{\mathrm{d}}{\mathrm{d}t}z(t) = J\nabla H, \quad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{2}$$

ここで，$H(z)$ はエネルギー関数，ハミルトニアンであり，J は標準的なシンプレクティック行列である．本研究では，より一般的な減衰項つきハミルトン系を対象とし，次のような形式を考える：

$$\frac{\mathrm{d}}{\mathrm{d}t}z(t) = \begin{bmatrix} 0 & I \\ -I & -\Gamma \end{bmatrix} \nabla H, \quad \Gamma \in \mathbb{R}^{n\times n},\ \Gamma \succeq 0. \tag{3}$$

このように，減衰項 Γ を含む拡張的なハミルトン構造は，SSMs における状態遷移行列 A に類似した役割を果たしうる．したがって，適切に設計した SSMs は，ハミルトン系の時間発展を学習するこ

とが可能であると考えられる．また，逆に，既存の SSMs の中には，行列 A として，実質的に (3) を利用しているものも存在する．したがって SSMs とハミルトン系は強い関係をもつと言える．

3　提案手法：状態空間モデルによるハミルトン系の時系列予測

本研究では，ハミルトン系から得られた軌道データ $\{(q(t), p(t)\}_{t=0}^{T}$ に基づいて，その時間発展を予測可能なニューラル状態空間モデルを構築する．例えば，近年提案された高効率な状態空間モデル Mamba などを導入し，その内部構造を物理系のモデリングに応用する．Mamba は，(1) を効率的にニューラルネットワークとして実装するために，状態カーネルに基づく入力畳み込みと，メモリゲート制御を組み合わせ以下の形式の離散モデルを用いる：

$$x_t = \mathrm{SSMConv}(u_{<t}; \theta) + G_t(u_t), \quad y_t = W x_t. \tag{4}$$

ここで，$\mathrm{SSMConv}(u_{<t}; \theta)$ は過去の入力列 $u_{<t}$ に対して状態カーネルに基づいた畳み込みをする演算であり，θ はカーネル構造に含まれる学習可能パラメータである．また，G_t は，現在の入力に対するゲートによるスケーリングであり，Mamba の特徴的な非線形性と時変構造を与える．出力は W_t により変換され，時間発展の予測を行う．モデルの訓練においては，予測された状態系列 $(\hat{q}_t, \hat{p}_t)$ が軌道データ系列 (q_t, p_t) に近づくよう，以下の損失関数を最小化する：

$$\mathcal{L}(\theta) = \sum_{t=1}^{T} \|\hat{q}_t - q_t\|^2 + \|\hat{p}_t - p_t\|^2. \tag{5}$$

また，系の物理性質の保存については，必要に応じてエネルギー関数 $H(q, p)$ を用いた正則化損失などを追加することも可能である．本稿では，一般的な SSMs の適用可能性に焦点を当て，構造的制約を課さない損失関数を用いる．

また，実際の予測では，例えば，初期状態 x_0 を入力系列から逐次的に構築する形でモデルに与え，時間発展に沿った (q_t, p_t) の予測列を得る．数値実験結果については，当日，報告する．

謝辞　本研究は JST CREST (JPMJCR1914, JPMJCR24Q5), JST ASPIRE (JPMJAP2329), JSPS KAKENHI (25K15148) の助成を受けた．

参考文献

[1] Gu, Albert, Karan Goel, and Christopher Ré, Efficiently modeling long sequences with structured state spaces, arXiv preprint arXiv:2111.00396 (2021).

[2] Fu, Daniel Y., et al, Hungry hungry hippos: Towards language modeling with state space models, arXiv preprint arXiv:2212.14052 (2022).

[3] Gu, Albert, and Tri Dao, Mamba: Linear-time sequence modeling with selective state spaces, arXiv preprint arXiv:2312.00752 (2023).

[4] Patro, Badri Narayana, and Vijay Srinivas Agneeswaran, Mamba-360: Survey of state space models as transformer alternative for long sequence modelling: Methods, applications, and challenges, arXiv preprint arXiv:2404.16112 (2024).

[5] Wang, Xiao, et al, State space model for new-generation network alternative to transformers: A survey, arXiv preprint arXiv:2404.09516 (2024).

Physics-Informed Neural Network に対する距離関数・動的重みの導入と非圧縮流の逆解析への適用

Smooth distance functions and unbiased dynamic weights enable reliable inverse analysis using physics-informed neural networks

出口 翔大 (Shota DEGUCHI)[1,2], 浅井 光輝 (Mitsuteru ASAI)[3]
[1] 筑波大学 (University of Tsukuba)
[2] 日本学術振興会 特別研究員 PD (JSPS Research Fellow PD)
[3] 九州大学 (Kyushu University)
e-mail : deguchi.shota.gm@u.tsukuba.ac.jp

1 Introduction

Physics-informed neural networks (PINNs) have become a powerful tool for solving forward and inverse problems governed by partial differential equations. However, accurately enforcing boundary conditions remains a key challenge. Conventional penalty methods approximate boundary conditions weakly and are sensitive to the choice of penalty weights.

In this work, we introduce a hard imposition approach that utilizes smooth distance fields to exactly satisfy boundary conditions, making it applicable to a wide range of geometries, including nonconvex domains. Furthermore, we combine this approach with unbiased dynamic weight tuning to enhance convergence when solving inverse problems. Numerical experiments demonstrate that this method is more accurate and reliable than conventional approaches, extending the applicability of PINNs [1].

2 Methods

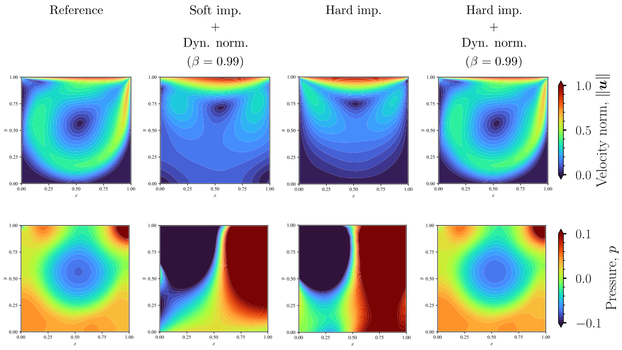
Physics-informed neural networks Consider the following steady-state incompressible Navier-Stokes equations

$$\nabla \cdot \boldsymbol{u} = 0 \quad \text{and} \quad (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = -\nabla p + \frac{1}{\mathrm{Re}} \nabla^2 \boldsymbol{u} \quad \text{in} \quad \Omega, \quad \boldsymbol{u} = \boldsymbol{u}|_\Gamma \quad \text{on} \quad \Gamma, \tag{1}$$

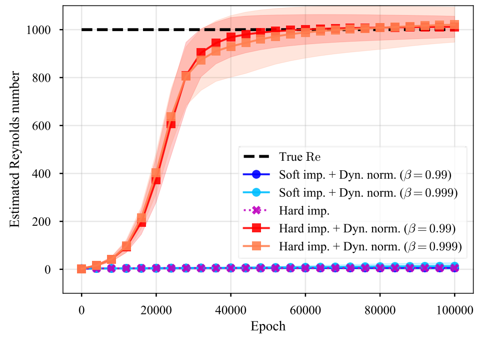
where Ω is the domain, $\Gamma = \partial\Omega$ is the boundary. The solution $(\boldsymbol{u}, p)$ is approximated by a neural network $(\hat{\boldsymbol{u}}_{\boldsymbol{\theta}}, \hat{p}_{\boldsymbol{\theta}})$, which are parameterized by a set of learnable parameters $\boldsymbol{\theta}$. PINN solves the above problem by minimizing $\mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}_{\mathrm{PDE}}(\boldsymbol{\theta}) + \lambda_{\mathrm{BC}}\mathcal{L}_{\mathrm{BC}}(\boldsymbol{\theta}) + \lambda_{\mathrm{Data}}\mathcal{L}_{\mathrm{Data}}(\boldsymbol{\theta})$, where $\lambda_{\mathrm{BC}}, \lambda_{\mathrm{Data}}$ are penalty parameters for balancing the contributions of different loss terms.

Smooth distance functions To enforce boundary conditions exactly, we employ a smooth distance function ϕ, which approximates the exact distance Φ to the boundary. Following techniques presented in [2], we construct an approximate solution that identically satisfies the boundary conditions, thereby ensuring that $\mathcal{L}_{\mathrm{BC}}(\boldsymbol{\theta}) = 0$.

Unbiased dynamic weights While the exact enforcement of boundary conditions provides a key advantage, a main strength of PINNs lies in their flexible applicability to inverse problems. Nevertheless, when solving such problems, choosing an appropriate weight λ_{Data} can often be challenging. To address this, we employ an unbiased dynamic weight tuning method [1]. This ap-



(a) Reference and PINN solutions



(b) Estimated Reynolds number

Figure 1: *Shear-driven cavity flow at Re = 1,000.* Similar results follows at higher Re numbers (see [1]).

proach adaptively adjusts λ_{Data} during training, thereby avoiding expensive parameter searches and offering an efficient and robust way to balance the different components of the loss function.

3 Results

Methods described in Section 2 were applied to the inverse analysis of shear-driven cavity flow at Reynolds number $\text{Re} = 1,000$. The task was to recover velocity and estimate pressure and Reynolds number from sparse velocity observations. The results, summarized in Figure 1, show that employing only the distance function or the dynamic weights does not produce a fundamental improvement. However, when these two techniques are combined, PINN successfully recovers both the flow field and the Reynolds number with high accuracy. We emphasize that similar results were obtained at higher Reynolds numbers, confirming the robustness and applicability of the proposed approach (see [1]).

4 Conclusion

In this study, we demonstrated that the combination of smooth distance functions and unbiased dynamic weights can enable reliable inverse analysis using PINNs. The distance fields enforce boundary conditions, while the adaptive weights balance the different loss components. Together, these techniques produce a significant improvement in accuracy and robustness when solving inverse problems. Numerical experiments show that applying either approach in isolation results in limited improvement; however, their combination leads to a substantial enhancement. We expect that this approach can be broadly applied to a range of problems, including those with complex boundary conditions and scarce observation data.

Reference

[1] Shota Deguchi and Mitsuteru Asai. Reliable and efficient inverse analysis using physics-informed neural networks with distance functions and adaptive weight tuning, 2025.

[2] V.L. Rvachev, T.I. Sheiko, V. Shapiro, and I. Tsukanov. Transfinite interpolation over implicitly defined sets. *Computer Aided Geometric Design*, Vol. 18, No. 3, pp. 195–220, 2001.

転位の形状不変性と物理深層学習を用いた地殻変動観測からの地震断層推定

Earthquake fault estimation from crustal deformation observations using geometry invariance of dislocations and physics-informed deep learning

岡﨑 智久 (Tomohisa Okazaki)[1], 縣 亮一郎 (Ryoichiro Agata)[2],
加納 将行 (Masayuki Kano)[3], 伊藤 武男(Takeo Ito)[4],
平原 和朗 (Kazuro Hirahara)[1], 上田 修功 (Naonori Ueda)[1]
[1]理化学研究所革新知能統合研究センター (RIKEN AIP), [2]海洋研究開発機構 (JAMSTEC),
[3]東北大学 (Tohoku University), [4]名古屋大学 (Nagoya University)
e-mail: tomohisa.okazaki@riken.jp

1 概要

結晶や地殻の変形を記述する転位論において，一様な相対変位による変形が転位面の境界にのみ依存するという形状不変性を示し，対応するポテンシャル場を導入することで転位による変形の効率的な解析方法を提案する．この方法を物理深層学習(physics-informed neural network)により実行し，地震時地殻変動の観測から断層の形状・運動をベイズ推論する際に使用することで，その有効性を示す．

2 理論：転位の形状不変性と Dislocation Potential

転位論(dislocation model)は，連続体（主に弾性体）の内部に変位不連続（転位）が生じた場合の変形場を記述する物理モデルである．転位論は材料科学における結晶の格子欠陥の記述を起源とし，地球科学においては地震やプレート運動による地震動や地殻変動の記述に用いられている．

先行研究[1]において，2 次元面外歪み（変形が面の直交成分のみ）を仮定した線形弾性体における一様な変位不連続による変形場が，転位端の位置にのみ依存し転位面の形状に依存しないという「形状不変性」が示された．一般の形状・弾性定数分布において成立する点が肝要である．さらに古典力学における保存力とエネルギーに着想を得て，断層端位置の関数としての変位場を Dislocation Potential (DP)と定義し，任意の転位面形状・転位量に対する変位場が DP の転位面に沿った線積分により計算できることが導かれた．これは，転位面（2 次元空間の曲線）に依存する無限次元の問題が，転位端（2 次元空間の点）という 2 次元問題に帰着することを意味する．

本研究では，形状不変性が 2 次元面内歪み（変形が面内成分のみ）および一般の 3 次元構造において成立することを示す．その際「一様な変位不連続」として，転位面に平行な剪断運動ではなく，転位面の向きに依らない定ベクトルの変位を対象とすることが要点である．転位論における境界条件の性質と線形性（重ね合わせの原理）を用いることで，形状不変性を簡明に導出できる．さらに，面外歪みの場合と同様に面内歪みにおいても DP を定義できる．転位・変位が面内 2 成分を持つため DP が 2 次元テンソルとなるほかは，線積分公式などを面外歪みの場合と同様に定式化できる．

3　応用：物理深層学習とベイズ推論による地震断層推定

形状不変性と DP の応用として，2 次元面内歪みを対象に，地震時地殻変動の観測データから地下の断層形状とすべり分布を同時推定する問題を考える．本研究では数値実験を扱う．

まず，深層学習により微分方程式を解く physics-informed neural network (PINN)[2]を地殻変動に適用[3]することで DP を解析する．PINN は一般の地下構造を扱えることに加え，解を連続関数として表現し，転位端位置のようなパラメタを入力変数とできるため，DP の解析に適している．面内歪みにおける PINN[4]を拡張することで DP 解析を実装できる．

次に，ベイズ推論により断層形状・すべり分布を推定する．具体的には，マルコフ連鎖モンテカルロ(MCMC)法の一種であるハミルトニアン・モンテカルロ(HMC)法を用いる．その尤度計算に表れる DP の線積分において上記の訓練済み PINN を使用する．Neural network に備わる自動微分により HMC の実行に必要な勾配を計算できる．仮定した断層運動に対する数値解から地表観測の模擬データに本手法を適用し，その有効性を検証する．

これまで MCMC を用いた地震断層推定では，多数の尤度計算が必要のため解析解の得られる単純な地下構造が仮定されることが一般的であった．本研究で示す通り，転位の形状不変性と PINN を組み合わせることで，複雑な地下構造における解析を実現できる．

4　議論

転位の形状不変性を一般の 3 次元構造において導出したが，DP は 2 次元構造において定義された．本節では DP の 3 次元解析への応用可能性を議論する．

まず，2 次元構造における 3 次元変動への拡張は容易である．面外・面内歪みにおける DP をブロック対角成分とする 3 次元 DP テンソルを定義することで 3 次元変位場を解析できる．このとき面外・面内成分の変形は転位面を共有するほかは独立である．

それに対し，3 次元構造における DP の定義は困難である．転位面の境界は閉曲線であり，それ自体が無限次元の自由度を持つため，効率的に解析できる DP を単純には定義できない．数学的な洞察により有用な物理量が導かれるかもしれない．

参考文献

[1] Okazaki, Tomohisa, Kazuro Hirahara and Naonori Ueda, Fault geometry invariance and dislocation potential in antiplane crustal deformation: physics-informed simultaneous solutions. Progress in Earth and Planetary Science, 11 (2024), 52.

[2] Raissi, Maziar, Paris Perdikaris and George E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational physics, 378 (2019), 686-707.

[3] Okazaki, Tomohisa, Takeo Ito, Kazuro Hirahara and Naonori Ueda, Physics-informed deep learning approach for modeling crustal deformation, Nature Communications, 13 (2022), 7092.

[4] Okazaki, Tomohisa, Kazuro Hirahara, Takeo Ito, Masayuki Kano and Naoroni Ueda, Physics-informed deep learning for forward and inverse modeling of inplane crustal deformation, Journal of Geophysical Research: Machine Learning and Computation, 2(1) (2025), e2024JH000474.