深層学習の物理モデリング・シミュレーションへの応用における最近の話題

Recent Topics in Applications of Deep Learning to Physical Modeling and Simulations

谷口 隆晴 (Takaharu Yaguchi)[1]
[1] 神戸大学, 理化学研究所 (Kobe University, RIKEN)
e-mail : yaguchi@pearl.kobe-u.ac.jp

1 背景

大規模言語モデルに代表される，近年の機械学習手法の急速な発展は社会に大きな影響を与えている．このような研究は，物理モデリング・シミュレーションに対しても同様に応用されており，Scientific Machine Learning(SciML) などと呼ばれている．科学への機械学習の応用は AI for Science などと呼ばれることもある．SciML は主に科学技術計算と機械学習の融合研究として進められてきたため AI for Science とは，少し目的が異なっている場合もあるが，最近では，これらの研究の差は少なくなってきている．SciML は，主に，物理シミュレーションを加速するための方法，物理モデリングのための方法，これらのアプローチに物理学の知識を活用する方法などの開発を目的とする．これらにより，物理シミュレーションの大幅な加速や，方程式が未知の現象のシミュレーションなどが可能となってきている．本発表では，このような研究における最近の研究のうち，特に Kolmogorov–Arnold ネットワークや，その Physics-Informed Neural Networks への応用などについて説明する．

2 Kolmogorov–Arnold ネットワーク

Kolmogorov–Arnold ネットワーク [1] は，Kolmogorov–Arnold の表現定理に基づく，新しいネットワーク構造である．従来の多層パーセプトロンと同様に，連続関数に対する普遍近似性をもっており，多層パーセプトロンと同様に利用することができる．多層パーセプトロンと大きく異なる点は，多層パーセプトロンでは行列やベクトルの値をデータに対して学習するのに対し，活性化関数を定めるためのパラメータを学習するという点である．また，Kolmogorov–Arnold ネットワークのもう一つの特徴は，多層パーセプトロンに比べて解釈性が高いことである．以下，これについて，より詳細に説明する．

Kolmogorov–Arnold ネットワークは，Kolmogorov–Arnold の表現定理や，それに関連した類似の定理に基づく．最も代表的な Kolmogorov–Arnold の表現定理では，任意の多変数連続関数 $f : [0,1]^n \to \mathbb{R}$ が

$$f(x_1, \ldots, x_n) = \sum_{j=1}^{2n+1} \psi_j \left(\sum_{k=1}^{n} \phi_{j,k}(x_k) \right)$$

という形に表せることが示される．ただし，$\phi_{j,k} : \mathbb{R} \to \mathbb{R}$ と $\psi_j : \mathbb{R} \to \mathbb{R}$ は，すべて単変数の関数である．これは，ヒルベルトの第 13 問題に関連したものであり，与えられた実数値多変数関数を，実数値単変数の合成で表すことができるか，という問いに対する一つの答えを与えたものである．

Kolmogorov–Arnold ネットワークは，このような定理に基づくモデルであり，上の式の $\phi_{j,k}$ や

ψ_j をデータに合うように推定する．これらの関数は，普通のニューラルネットワークにおける活性化関数に相当すると考えることができ，Kolmogorov–Arnold ネットワークでは，これらをパラメタライズされた関数でモデル化し，データに合うように学習する．具体的には，スプライン関数などが利用されることが多い．この方法では，多変数関数を，単変数関数の合成で近似しようとするが，単変数関数は，グラフをかいて可視化したり，多項式などで近似したりすることが簡単である．そのため，従来法に比べて，高い解釈性をもつ．また，本質的に単変数関数を学習すれば良いため，次元の呪いの影響を受けない可能性があると期待されている．一方，実装が複雑になりやすく，GPU を利用した並列化がききにくいこともある．そのような場合は，普通の多層パーセプトロンに比べて計算時間が長くなってしまうことがあるために，注意が必要である．

3　Physics-Informed Neural Networks への応用

Physics-Informed Neural Networks は，偏微分方程式とニューラルネットワークを組み合わせた研究である [2]．様々な用途に応用されているが，代表的な応用は，偏微分方程式の近似解を求めることである．具体的には，偏微分方程式の解をニューラルネットワークで表し，偏微分方程式が満たされるようにニューラルネットワークを学習させることで，近似解を求める．近年，この方法について，Kolmogorov–Arnold ネットワークした方法が提案された [3, 4]．この方法は，普通のニューラルネットワークを利用した従来手法に比較して，少ないパラメータで精度よく近似解が計算できることが報告されている．

当日の発表では，このような方法を含め，主に Kolmogorov–Arnold ネットワークの物理への応用について，理論的な側面も含め，より詳細に説明する．

謝辞　本研究は JST CREST Grant Number JPMJCR1914 および JPMJCR24Q5, JST ASPIRE JPMJAP2329 の補助を受けた．

参考文献

[1] Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y. Hou, Max Tegmark, KAN: Kolmogorov-Arnold Networks, arXiv:2404.19756 (2024).

[2] M. Raissi, P. Perdikaris, G.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, Vol. 378 (2019), 686–707.

[3] K. Shukla, J. D. Toscano, Z. Wang, Z. Zou, G. E. Karniadakis, A comprehensive and FAIR comparison between MLP and KAN representations for differential equations and operator networks, Computer Methods in Applied Mechanics and Engineering, Vol. 431 (2024), 117290.

[4] Juan Diego Toscano, Vivek Oommen, Alan John Varghese, Zongren Zou, Nazanin Ahmadi Daryakenari, Chenxi Wu, George Em Karniadakis, From PINNs to PIKANs: Recent Advances in Physics-Informed Machine Learning, arXiv:2410.13228 (2024)

Neural Operator による津波シミュレーションの代理モデル

Surrogate Modeling of Tsunami Simulation using Neural Operators

染矢 真好 (Masayoshi Someya)[1], 古村 孝志 (Takashi Furumura)[1]
[1] 東京大学地震研究所 (Earthquake Research Institute, University of Tokyo)
e-mail : someya@eri.u-tokyo.ac.jp

1　はじめに

津波の数値シミュレーションは、海底の鉛直変位を初期条件として海面波高の時間発展を計算する手続きであり、事前のリスク評価、地震発生直後の津波警報、事後の波源逆解析（観測波形を用いた断層運動の推定）などにおいて必要不可欠な技術である。これらのシミュレーションでは、通常、線形長波方程式などの支配方程式を離散化し、数値的な時間積分を行うが、計算領域の広さや必要な分解能の高さから、一般に計算コストが高くなる。特に、観測波形を用いて地震断層のパラメータを推定する逆解析では、最適な波源を求めるために同様の計算を多数繰り返す場合があり、計算コストの削減が課題となる。

こうした課題を解決するため、本研究では Neural Operator（NO）[1, 2] を用いた津波シミュレーションの軽量な代理モデル（surrogate model）を構築した。NO とは、入力と出力が関数空間に属する演算子（operator）を学習する新しい機械学習の枠組みである。本研究で構築した訓練済みモデルでは、初期条件（海底の鉛直変位）が入力されると、その後の波高の時空間発展を高速に出力することができる。本発表では、NO モデルの構造と訓練方法、逆解析への応用可能性について述べる。

2　代理モデルの構築

本研究では、津波の時空間的な波高分布を予測するために、U-shaped Neural Operator（U-NO）[3] を採用した。U-NO 全体は U-Net に類似したエンコーダ・デコーダ構造を持ち、段階的なダウンサンプリング・アップサンプリングとスキップ接続を通じて、局所的な特徴と大域的な構造の両方を抽出する。また、畳み込み層として用いられている Fourier 層では、離散フーリエ変換（FFT）により入力を（周）波数領域に変換し、各（周）波数成分に複素数重みを乗じた後、逆 FFT により（時）空間領域に戻す。この構成により、波動場の様々なスケールを効果的に学習することができる。

NO モデルの訓練に必要なデータセットは、津波シミュレーションにより生成した。東北沖の海域を計算領域として設定し、（1）ガウス型の滑らかな波源、（2）断層パラメータ（位置、大きさ、すべり量など）から計算された海底変位、の 2 種類の波源をランダムに生成した。これらを初期条件として、差分法ベースの津波シミュレーションライブラリ JAGURS [4] を用いて、波高の時空間発展を計算した。得られた波源・波動場のペアをデータセットとし、訓練に用いた。

訓練済みモデルの精度を検証するために、訓練に用いていない波源を入力し、その出力を津波シミュレーションによる正解波動場と比較した。その結果、NO モデルは空間的な波動構造を良好に再現するものの、波数スペクトルにおいて短波長成分の振幅をやや過小評価する傾向も確認された。

3　自動微分を用いた逆解析の試み

NO モデルは Python の機械学習ライブラリ PyTorch 上に構築されており、モデルの入力に対する出力の勾配を、PyTorch の自動微分機能を用いて簡単に計算できる。このため、観測された津波波

形と NO モデルの予測とのミスフィット関数を最小とする初期波高を、勾配法により求めることが可能である。ただし、初期波高を全ての空間格子点で推定するため未知数は数万個となり、劣決定問題となる。このため、空間平滑化のためのラプラシアン正則化を導入し、安定した推定を実現した。

また、初期波高（海底変位）は断層すべりに起因するため、断層パラメータの推定も重要な課題である。津波のモデリングでは、断層運動による海底変位を Okada の式 [5, 6] と呼ばれる解析解を用いて計算する手順が広く用いられている。本研究では、Okada の式を PyTorch 上に実装し、NO モデルと結合することで、断層パラメータを入力として津波の波動場を予測する統合モデルを構築した。さらに、統合モデルに勾配法を適用して断層パラメータの推定を行なった。従来、津波波形に非線形に影響する断層パラメータはあらかじめ妥当な値に固定するか、grid search により探索されていたが、それらを勾配法で探索可能になったことは、大きな利点である。一方で、局所解への落ち込みや初期値依存の課題も明らかになった。

4　今後の展望

本研究で構築した代理モデルは東北沖を対象としており、他海域を含む任意の場所に適用できる代理モデルの構築にはさらなる工夫が必要である。また、本モデルは最も単純な線形長波方程式に基づいているため、より複雑で計算コストの高い分散性津波の方程式などに対応した代理モデルを構築することを目指す。さらに、逆解析においては、勾配を用いたパラメータ探索手法の改良に加え、Hamiltonian Monte Carlo（HMC）や No U-Turn Sampler（NUTS）などのベイズ推定手法を導入し、パラメータの事後分布を推定することも検討している。

謝辞　第一著者は、日本学術振興会から特別研究員（DC1）としての支援を受けています。また、本研究ではデータ活用社会創成プラットフォーム mdx の計算資源を利用しました。

参考文献

[1] Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A., Fourier Neural Operator for Parametric Partial Differential Equations, arXiv preprint arXiv:2010.08895 (2020).

[2] Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., and Anandkumar, A., Neural Operator: Learning Maps Between Function Spaces With Applications To PDEs, Journal of Machine Learning Research, 24(89) (2023), 1–97.

[3] Rahman, M. A., Ross, Z. E., and Azizzadenesheli, K., U-NO: U-shaped Neural Operators, arXiv preprint arXiv:2204.11127 (2022).

[4] Baba, T., Takahashi, N., Kaneda, Y., Ando, K., Matsuoka, D., and Kato, T., Parallel Implementation of Dispersive Tsunami Wave Modeling with a Nesting Algorithm for the 2011 Tohoku Tsunami, Pure and Applied Geophysics, 172(12) (2015), 3455–3472.

[5] Okada, Y., Surface deformation due to shear and tensile faults in a half-space, Bulletin of the Seismological Society of America, 75(4) (1985), 1135–1154.

[6] Okada, Y., Internal deformation due to shear and tensile faults in a half-space. Bulletin of the seismological society of America, 82(2) (1992), 1018–1040.

無限次元ハミルトン系を学習するのための近似的なシンプレクティックニューラル作用素

Approximate Symplectic Neural Operator for Learning Infinite-Dimensional Hamiltonian Systems

Yeang Makara[1], 田中 佑典 (Yusuke Tanaka)[2], 松原 崇 (Takashi Matsubara)[3],
谷口 隆晴 (Yaguchi Takaharu)[4]
[1] 神戸大学 (Kobe University),
[2]NTT コミュニケーション科学基礎研究所 (NTT Communication Science Laboratories),
[3] 北海道大学 (Hokkaido University),
[4] 神戸大学, 理化学研究所 (Kobe University, RIKEN)
e-mail : makarayeang99@gmail.com

1 背景と目的

無限次元ハミルトン系は，波動方程式，電磁波方程式など，多くの現象を記述することができるため，物理学や工学において広く現れる．これらの系は，無限次元の関数空間上で定義されるシンプレクティック構造を持つが [1]，その構造を保持しながら高精度なシミュレーションを実現することは，数値解析および機械学習における困難な課題である．

特に既存の Fourier Neural Operator（FNO）や Graph Neural Operator（GNO），および DeepONet といった機械学習ベースの手法 [2] は，無限次元空間に対応しているものの，シンプレクティック構造を保持しないため，長時間の予測においてエネルギーの発散や構造の劣化が生じる．また，SympNet などの構造保存型ネットワークも，その適用範囲は有限次元に限られている [3]．

本研究では，これらの課題を解決するために，無限次元ハミルトン系の構造（特にシンプレクティック性）を近似的に保存するニューラル作用素「近似的なシンプレクティックニューラル作用素（Approximate Symplectic Neural Operator；ASNO)」を提案する．ASNO は関数空間上で構造保存的な時間発展写像を実現するよう設計されており，長時間にわたる安定かつ精度の高いシミュレーションを可能にする新たな枠組みである．

2 問題設定

本研究では，ヒルベルト空間 $\mathbb{H} \times \mathbb{H} = L^2(\Omega) \times L^2(\Omega)$ 上で定義された無限次元ハミルトン系を考える．領域 $\Omega \subseteq \mathbb{R}^d$ において、初期条件 $u_0(x) \in \mathbb{H} \times \mathbb{H}$ と適当な境界条件が与えられた際に，以下のハミルトン偏微分方程式に従って系が時間発展するものとする：

$$\frac{\partial u}{\partial t}(x,t) = J\frac{\delta H}{\delta u}(u(x,t)), \quad u(x,0) = u_0(x), \quad x \in \Omega, \quad t \in [0,T]. \tag{1}$$

ここで，$u(\cdot,t)$ は相空間上の状態を表し，$H : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ は系のハミルトニアンである．$\delta H/\delta u$ は変分導関数である．また，J は歪対称作用素（$J^* = -J$）である．

3 提案手法

シンプレクティック性を持つニューラル作用素を作るためには以下のいくつかの基本的なアイディアに基づく．

補題 1 シンプレクティック作用素の合成は，再びシンプレクティック作用素となる．

補題 2　$\mathcal{K}_i, \mathcal{K}_j : \mathbb{H} \to \mathbb{H}$ が自己随伴作用素ならば，以下のように定義される剪断演算

$$M_{\mathrm{up}}(\mathcal{K}_i) = \begin{pmatrix} I & \mathcal{K}_i \\ 0 & I \end{pmatrix}, \quad M_{\mathrm{low}}(\mathcal{K}_j) = \begin{pmatrix} I & 0 \\ \mathcal{K}_j & I \end{pmatrix}$$

は，$\mathbb{H} \times \mathbb{H}$ から $\mathbb{H} \times \mathbb{H}$ へのシンプレクティック作用素となる．

補題 3　$\mathcal{K}_i, \mathcal{K}_j : \mathbb{H} \to \mathbb{H}$ がある滑らかな汎関数 $\Phi_i, \Phi_j : \mathbb{H} \to \mathbb{R}$ に対して $\mathcal{K}_i = \nabla\Phi_i, \mathcal{K}_j = \nabla\Phi_j$ と表されるとき，以下のように定義される剪断演算

$$M_{\mathrm{up}}(\mathcal{K}_i) = \begin{pmatrix} I & \mathcal{K}_i \\ 0 & I \end{pmatrix}, \quad M_{\mathrm{low}}(\mathcal{K}_j) = \begin{pmatrix} I & 0 \\ \mathcal{K}_j & I \end{pmatrix}$$

は，$\mathbb{H} \times \mathbb{H}$ から $\mathbb{H} \times \mathbb{H}$ へのシンプレクティック作用素となる．

今回は，研究の最初のステップとして，補題 2 と補題 3 で述べた $\mathcal{K}_i, \mathcal{K}_j$ を普通の FNO を用いて学習する実験を行った．このモデルは厳密にいうとシンプレクティック性を持たないが，シンプレクティック作用素に近い形を持つので，シンプレクティック構造はある程度保存できる．このモデルを ASNO（Approximate Symplectic Neural Operator）と呼ぶ．このような構成により，提案する ASNO は各ステップでの時間発展を近似的にシンプレクティックに保ちつつ，関数空間全体での高精度なモデル化を可能にする．特に，従来のモデルと比較して，長期的なエネルギー保存性と数値安定性において優れた性能を発揮することが期待される．

4　数値実験と評価

提案手法の性能評価として，波動方程式，電磁波方程式，シュレディンガー方程式，クライン・ゴルドン方程式の 4 種類の代表的な系を用いて数値実験を行った．数値実験の結果，ASNO は FNO，GNO，DeepONet や CNO など既存の代表的ニューラル作用素手法と比較して，数値精度，長期安定性，ハミルトニアンエネルギーおよびシンプレクティック形式の保存性において大幅に優れた性能を示した．特に，長時間シミュレーションにおいてもエネルギードリフトが抑制され，系の本質的な構造を良く捉えることができた．詳細については，当日，報告する．

謝 辞　本 研 究 は JST ASPIRE JPMJAP2329, JST PRESTO (JPMJPR24TB), CREST (JPMJCR1914, JPMJCR24Q5), ASPIRE (JPMJAP2329), JSPS KAKENHI (24K15105, 25K15148), SGH 奨学金の補助を受けた．

参考文献

[1] Chernoff, P. R., & Marsden, J. E. (1974). Properties of Infinite Dimensional Hamiltonian Systems. *Lecture Notes in Mathematics*, Vol. **425**, Springer-Verlag.

[2] Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A. M., & Anandkumar, A. (2022). Neural Operator: Learning Maps Between Function Spaces With Applications to PDEs. *Journal of Machine Learning Research*, **23**, 1–97.

[3] Jin, P., Zhang, Z., Zhu, A., Tang, Y., & Karniadakis, G. E. (2020). SympNets: Intrinsic structure-preserving symplectic networks for identifying Hamiltonian systems. *arXiv preprint* arXiv:2001.03750.

