

ニューラルネットワークによる非線形 4 階偏微分方程式の係数と解の予測

寺嶋 達磨[†] 村手 翼[‡] 渡辺 崇[‡][†]名古屋大学 情報学部 [‡]名古屋大学院 情報学研究科

1. はじめに

複雑な現象の解析を、限られたデータ量で行う事例は多い。このため、物理データではなく、系を支配する微分方程式から、現象をニューラルネットワーク(NN)によって、学習、同定、予測する試みがなされ、これまでに 3 階までの非線形偏微分方程式が対象にされてきた[1], [2]。本研究では、物理現象を支配する方程式と境界条件を事前情報として組み込んだ NN(Physics-informed NN)[1]の非線形高階偏微分方程式に対する有効性を、非線形 4 階偏微分方程式である蔵本-Sivashinsky(KS)方程式[3]に着目し、「係数の予測」と「解の予測」の観点から検証する。

2. データセット

本研究で扱う KS 方程式は以下の通りである、

$$\frac{\partial u}{\partial t} + \lambda_1 \frac{\partial u^2}{\partial x} + \lambda_2 \frac{\partial^2 u}{\partial x^2} + \lambda_3 \frac{\partial^4 u}{\partial x^4} = 0, \\ x \in [0, \alpha\pi], t \in [0, \beta], \quad (1)$$

$$u(0, x) = \cos(1/\gamma)(1 + \sin(1/\gamma)).$$

使用するデータセット($t, x, u(t, x)$)は、(1)式において、本研究では、 $(\lambda_1, \lambda_2, \lambda_3) = (0.5, 1.0, 1.0)$ として、MATLAB を用いて生成する。「係数の予測」では、 $(\alpha, \beta, \gamma) = (32, 200, 16)$ とし、 $(t, x, u(t, x))$ を与え、(1)式を満たす $\lambda_1, \lambda_2, \lambda_3$ を予測する。「解の予測」では、 $(\alpha, \beta, \gamma) = (4, 15, 2)$ とし、 $\lambda_1, \lambda_2, \lambda_3$ の値と、 u の初期条件と x 方向への周期性を定められた選点で与え、(1)式を満たす u の予測値 u_{pred} を得る。

3. NN モデル

3.1. NN モデルの構成

本研究では、順伝搬全結合型 NN を採用する。入力 t, x の 2 入力、出力は解 u の 1 出力である。中間層の層数(M_L)およびユニット数(M_N) は任意の数に設定する。活性化関数は $\tanh$ を、最適化関数は Adam と L-BFGS を用いる。

3.2. 「係数の予測」の NN モデル

3.1 節で述べた NN は、以下に示す損失(Loss)を最小化するように NN の重みとバイアスおよび微分方程式パラメータ $\lambda_1, \lambda_2, \lambda_3$ を学習する、

Neural network based prediction of parameters and solutions of nonlinear 4th-order partial differential equation

Tatsuma TERASHIMA[†], Tsubasa MURATE[‡], Takashi WATANABE[‡] [†]School of Informatics, Nagoya University,

[‡]Graduate School of Informatics, Nagoya University

$$Loss = MSE_u + MSE_f. \quad (2)$$

ここで、 MSE_u はデータセットから選んだ N_u 個の u と予測した u_{pred} の平均二乗誤差(MSE)である。

MSE_f は予測した u_{pred} を(1)式の微分方程式の左辺に代入した値の MSE である。

3.3. 「解の予測」の NN モデル

ここでは、以下に示す Loss を最小化するように NN の重みとバイアスパラメータを学習する、

$$MSE = MSE_0 + MSE_b + MSE_f. \quad (3)$$

ここで、 MSE_0 は初期値の選点 N_0 個における実際の $u(0, x)$ と予測した $u_{pred}(0, x)$ の MSE である。 MSE_b は周期境界の選点 N_b 個における予測した $u_{pred}(t, 0)$ と $u_{pred}(t, \alpha\pi)$ の 0 階から 3 階の微分同士の MSE の和である。 MSE_f は t と x が張る空間から選んだ N_f 個の選点において予測した u_{pred} を(1)式の微分方程式の左辺に代入した値の MSE である。

4. 評価実験と結果

4.1. 「係数の予測」における評価実験と結果

$N_u = 60000$ として、任意の M_L と M_N における、予測係数の実係数に対する誤差(%)を調べる。

$\lambda_1, \lambda_2, \lambda_3$ で同様の傾向が見られたため、 λ_1 の予測結果のみを Table 1 で示す。 M_L および M_N と精度には、相関があるが、一部、例えば、 $M_L = 3$ における、 $M_N = 80$ と $M_N = 100$ のような非単調な傾向も見られる。 λ_1 に関して、最も良い精度を示す $M_L = 6$ 、 $M_N = 100$ の時の予測誤差は 0.28% であり、その時の実際と予測の、 u と式を Fig.1 に示す。

$M_L = 6, M_N = 100$ として、任意の N_u とノイズ率(R_N)における、誤差(%)を調べる。ノイズは、標準正規分布に訓練データ u の標準偏差と既定の R_N を乗じたものである。 λ_1 の予測結果を Table 2 に示す。 N_u の減少に伴い、精度は減少する。また、 R_N の増加に伴い、精度は減少する。しかし、 N_u が増加するにつれ、高い R_N に対して誤差が小さく抑えられる。これは、 N_u の増加により、ノイズの平滑化が起るためと考えられる。

Table 1: Error(%) of λ_1 depending on M_L and M_N .

$M_L \backslash M_N$	40	60	80	100
3	47.72	12.96	6.96	7.65
4	30.89	13.20	4.92	1.03
5	28.27	2.74	1.12	0.73
6	14.22	1.56	2.51	0.28

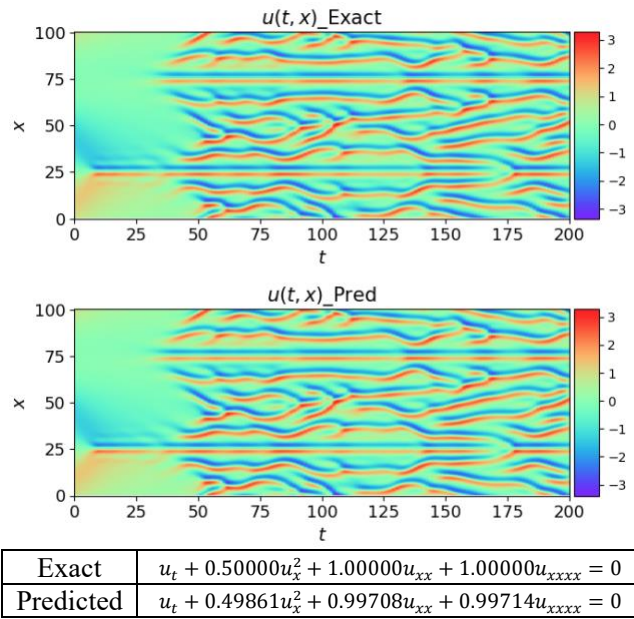


Fig.1: Exact and predicted $u(t, x)$ and equations. $M_L = 6$, $M_N = 100$, $N_u = 60000$, $R_N = 0\%$.

Table 2: Error(%) of λ_1 depending on N_u and R_N .

$N_u \setminus R_N$	0%	1%	5%	10%
10000	4.59	3.08	7.52	17.12
20000	0.52	1.36	1.73	4.78
40000	0.81	1.07	1.40	1.92
60000	0.28	0.55	1.83	1.50

以上より, Physics-informed NN は「係数の予測」において, 高い有効性を示す。

4.2. 「解の予測」における評価実験と結果

$N_f = 60000$, $N_0 = 200$, $N_b = 200$ として, 任意の M_L と M_N における, 予測した u_{pred} の実際の u に対する相対 L2 ノルム誤差を Table 3 に示す. M_L および M_N には, 精度との相関が見られず, 全体的に非単調な傾向がある。

$M_L = 6$, $M_N = 100$ として, 任意の N_0 と N_b と N_f における, 相対 L2 ノルム誤差を Table 4 に示す. この場合も, N_0 , N_b および N_f と, 精度との相関が見られず, 全体的に非単調な傾向がある。

これらの実験で, 相関が見られない要因としては, NN の過学習が考えられる. しかし, 実際の u と予測した u_{pred} を比較した図 Fig.2 より, 解の予測は妥当に行えている。

以上より, Physics-informed NN は「解の予測」に対しても, 有効性を持つという結果が得られる. しかし, 「係数の予測」で用いた比較的大きな (α, β, γ) では精度の減少が見られた. この点については, さらなる NN のチューニングが必要であると考えられる。

Table 3: Relative L2 Error between exact and predicted $u(t, x)$ depending on M_L and M_N .

$M_L \setminus M_N$	40	60	80	100
3	7.72e-3	2.37e-3	2.10e-2	1.12e-2
4	6.72e-3	5.36e-3	3.18e-3	1.04e-2
5	1.28e-3	4.57e-3	8.53e-3	5.20e-3
6	6.67e-3	3.98e-3	5.16e-3	2.40e-3

Table 4: Relative L2 error between exact and predicted $u(t, x)$ depending on numbers of N_0 , N_b and N_f .

$N_0, N_b \setminus N_f$	10000	20000	40000	60000
50,50	2.30e-3	1.25e-2	3.48e-3	4.02e-3
100,100	5.78e-3	7.09e-3	4.67e-3	6.14e-3
200,200	3.84e-3	6.20e-3	5.16e-3	2.40e-3

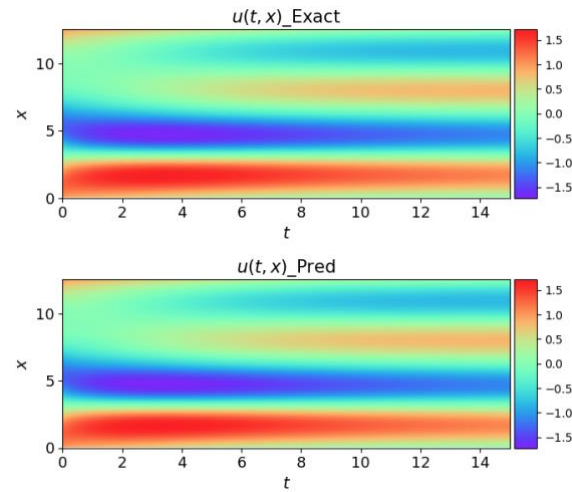


Fig.2: Exact and predicted $u(t, x)$. $M_L = 6$, $M_N = 100$, $N_f = 60000$, $N_0 = 200$, $N_b = 200$.

5. 終わりに

本研究では, Physics-informed NN によって非線形 4 階偏微分方程式である KS 方程式の「係数の予測」と「解の予測」を行った. 本 NN は KS 方程式の「係数の予測」において, 高い有効性を示した. 「解の予測」においては, 初期段階で解が比較的滑らかな時間までの予測は良好であったが, 時間が経過し, 解の空間変動が大きくなると予測に問題が生じた. 今後は, 「係数の予測」および「解の予測」の精度の向上を目指すとともに, さらに高階の偏微分方程式についても研究を進める。

6. 参考文献

- [1] Raissi M., Perdikaris P., Karniadakis G.E., J.Comput.Phys 378(2019), 686-707.
- [2] Qin T., Wu K., Xiu D., J.Comput.Phys 395(2019), 620-635.
- [3] 森肇・蔵本由紀(2000)「散逸構造とカオス」岩波書店