

リザーバーコンピューティングを用いた 流体モデリングの高度化

齊木吉隆 (一橋大学 大学院経営管理研究科)

概要

流体の基礎方程式である Navier-Stokes 方程式からマクロ変数に関する時間発展方程式を解析的に導出することは困難なことが知られている。研究代表者らは、リザーバーコンピューティングと呼ばれる時系列データの機械学習手法を用いて、流体マクロ変数の時系列データの学習に基づいてミクロ変数に立ち戻ることなく数理モデルを構築し、時間発展を予測することに成功している (Nakai and Saiki, Physical Review E, 2018)。本研究では、リザーバーコンピューティングに定数項を追加した形式を採用し、スペクトル半径を大きく設定して学習を行う。出力項は通常 1 次近似を用いることが通例であったが、ここでは 2 次近似を用いて性能向上を図った。これらをもとにして同程度の次元で精度の高い流体の機械学習時間発展モデリングを行った。

1 共同研究に関する情報

1.1 共同研究を実施した拠点名

- 東京大学 情報基盤センター

1.2 課題分野

- 大規模計算科学課題分野

1.3 参加研究者の役割分担

本共同研究は 4 人体制で行われている。

- 研究代表者 齊木吉隆: 背後にある数学的構造の解明、リザーバーコンピューティングによる数理モデル構築
- 副代表者 中井拳吾 (岡山大学学術研究院 環境生命自然科学学域): トレーニングデータの作成、リザーバーコンピューティングによる数理モデル構築
- その他の共同研究者 小林幹 (立正大学 経済学部): 背後にある数学的構造の解明

- その他の共同研究者 武藤誠 (日本大学 経済学部 経済学科): トレーニングデータの作成

研究の各側面でスーパーコンピュータが活用されている。

2 研究の目的と意義

機械学習の利用は分類、画像認識等の静的なものに留まらず、決定論的なダイナミクスに関する時間発展モデリングなど動的な対象へも広がりつつある。流体の基礎方程式である Navier-Stokes 方程式からマクロ変数に関する時間発展方程式を解析的に導出することは困難なことが知られている。研究代表者らは、リザーバーコンピューティングと呼ばれる時系列データの機械学習手法を用いて、流体マクロ変数の時系列データの学習に基づいてミクロ変数に立ち戻ることなく数理モデルを構築し、時

間発展を予測することに成功した。また、流体運動が示すカオス性 (指数的不安定性) によって、長時間の時間発展予測は必然的に破綻するものの、得られた数理モデルは統計性質を再現することも確認した。また、研究代表者らの JHPCN 課題「機械学習に基づく流体変数の未来予測と数学的背景 (jh190070)」によって、学習可能な時系列データの数が流体運動の不安定次元より少ない数にとどまる場合にも、時間遅れ座標を導入することで効率良く数理モデルが構築できることを確認した。実用上は観測可能な時系列の数は少数に限定されることが通例であるが、そのような場合にも本手法を適用することができることを意味する。また、2020-2023 年度の JHPCN 課題を通して、力学系の性質の一つである不動点や周期軌道や正のリアプノフ指数などの再現が機械学習により学習可能であることなど、機械学習モデリングの数理的な構造を明らかにした。

これらの研究をもとに流体に対する高速でかつ高性能な機械学習時間発展モデルの構成を行った。これにより機械学習による流体モデリングの対象を広げることに貢献できる。

3 当拠点公募型研究として実施した意義

リザーバーコンピューティングに基づく物理ダイナミクスの予測に関しては、アメリカのメリーランド大学のグループがこれまでリードしてきたが、リザーバーコンピューティングを用いた流体乱流の数理モデリングならびにそれを用いた予測に、我々、研究代表者と副代表者のグループがはじめて成功した。拠点公募型共同研究として数学・経済学・大規模計算の学際的なチームを編成して多くのスーパーコンピュータ資源を用いた研究を推進することによって、リザーバーコンピューティングモデルの力学系

解析においても我が国が世界をリードすることが期待される。

本研究課題では流体の時系列データの機械学習によるモデリングを行った。高次元力学系である流体の時系列データを精度良く学習するためには、大量の時系列データを取得し大規模な計算資源を使うことが必要不可欠である。JHPCN の枠組みを最大限に活用することによって研究目標の達成が実現できた。

4 前年度までに得られた研究成果の概要

該当なし

5 今年度の研究成果の詳細

本計算資源を利用し、まずは Navier-Stokes 方程式の直接数値計算によってエネルギー変数の時系列データを生成した。そのエネルギー変数を学習データとしてリザーバーコンピューティングを用いて時間発展モデルを構成した。まずはじめに本研究の要である時系列データの学習手法であるリザーバーコンピューティングと学習に用いる流体データについて言及した後、時系列予測した結果について述べる。

リザーバーコンピューティング、力学系

$$\phi_{n+1} = \mathbf{f}(\phi_n)$$

に関して、変数 $\mathbf{u} = \mathbf{h}_1(\phi_n) \in \mathbb{R}^M$ の時系列データを学習することを考える。まずは、この変数 \mathbf{u} に対して、リザーバー状態ベクトル $\mathbf{r} \in \mathbb{R}^N$ ($N \gg M$) は次の式で支配されているとする:

$$\mathbf{r}(t+1) = \tanh(\mathbf{A}\mathbf{r}(t) + \mathbf{W}_{\text{in}}\mathbf{u}(t) + \boldsymbol{\xi}). \quad (1)$$

行列 \mathbf{A} , \mathbf{W}_{in} は $N \times N$, $N \times M$ 行列である。リザーバー状態ベクトルの初期値はその各成分を $(0, 1]$ の一様乱数から定めた $\mathbf{r}(-\tau)$ とする。た

だし、 $\tau (\gg 1)$ はトランジェント時間のステップ数に相当する。また、 $\boldsymbol{\xi} = (\xi, \dots, \xi)^T \in \mathbb{R}^N$ はバイアス項である。以下ではリザーバー状態ベクトル \mathbf{r} が属する空間 \mathbb{R}^N をリザーバー空間と呼ぶ。

行列 \mathbf{W}_{in} の各列はただ一つのみ $[-\sigma, \sigma]$ の一様乱数から値を与え、それ以外は 0 とする。行列 \mathbf{A} は $D \times N$ 個の非零成分を持つ疎行列とする。 $D \times N$ 個の非零成分は $[-1, 1]$ の一様乱数から値を与える。さらに、行列 \mathbf{A} はスペクトル半径が ρ になるように行列のすべての要素を調整する。

$0 < t \leq L$ (学習時間と呼ぶ) での時系列データ $\{\mathbf{W}_{\text{out}}\mathbf{r}(t) + \mathbf{r}(t)^T \mathbf{W}_{\text{Qout}}^* \mathbf{r}(t)\}$ が既知の時系列データ $\{\mathbf{u}(t)\}$ の近似になるように \mathbf{W}_{out} , $\mathbf{W}_{\text{Qout}}^*$ を決定する [2]。ここでは次の二次形式が最小値をとるように \mathbf{W}_{out} , \mathbf{W}_{Qout} を決定する:

$$\sum_{l=1}^L \|\mathbf{W}_{\text{out}}\mathbf{r}(l) + \mathbf{r}(l)^T \mathbf{W}_{\text{Qout}}^* \mathbf{r}(l) - \mathbf{u}(l)\|_2^2 + \beta(\|\mathbf{W}_{\text{out}} \circ \mathbf{W}_{\text{out}}\|_1 + \|\mathbf{W}_{\text{Qout}} \circ \mathbf{W}_{\text{Qout}}\|_1). \quad (2)$$

ただし \circ はアダマール積を意味する。また、ベクトル \mathbf{a} に対して $\|\mathbf{a}\|_p$ は、 \mathbf{a} の各要素の p 乗し、それらをすべて足して p 乗根とった値である。 $\beta (\geq 0)$ を含む項は \mathbf{W}_{out} , \mathbf{W}_{Qout} の過剰適合を避けるために導入している。

また、予測時間 $t > L$ での時系列データ \mathbf{u} は利用できないため、以下のように時間発展させる。

$$\begin{cases} \mathbf{r}(t+1) = \tanh(\mathbf{A}\mathbf{r}(t) + \mathbf{W}_{\text{in}}\mathbf{u}(t) + \boldsymbol{\xi}), \\ \mathbf{u}(t) = \mathbf{W}_{\text{out}}^* \mathbf{r}(t) + \mathbf{r}(t)^T \mathbf{W}_{\text{Qout}}^* \mathbf{r}(t). \end{cases}$$

これを機械学習モデルと呼ぶ。

学習の際には、学習に用いる変数 $X(t)$ を $\tilde{X}(t) = [X(t) - X_1]/X_2$ のように標準化して用いる。ただし、 X_1 は平均値、 X_2 は標準偏

差を表す。予測時間の時系列データ $X(t)$ を $\tilde{X}(t)$ から再現するときは、学習時間の標準化に用いた X_1 と X_2 を用いる。この標準化により σ の調整を避けることができる。

このリザーバーコンピューティングの計算に計算資源の多くを利用している。特に (2) の最小化解を求めるところがメインの計算であり、最小化解を求める際に必要となる行列-行列積は適切なサイズになるようにプログラミングし効率よく計算している。

学習時系列流体データの生成。周期境界条件の下で 3 次元非圧縮 Navier-Stokes 方程式の直接数値計算によって得られた時系列データをリザーバーコンピューティングの学習データとして用いる:

$$\begin{cases} \partial_t v - \nu \Delta v + (v \cdot \nabla)v + \nabla \pi = f, \quad \nabla \cdot v = 0, \\ v|_{t=0} = v_0 \quad \text{with } \nabla \cdot v_0 = 0, \end{cases}$$

ただし、 $\mathbb{T} = [0, 2\pi)$, $\nu > 0$ は動粘性係数、 $\pi(x, t)$ は圧力、 $v(x, t) = (v_1(x, t), v_2(x, t), v_3(x, t))$ は速度とする。空間離散化のためにフーリエスペクトル法 [1] を用いる。時間積分は 4 次のルンゲクッタ法を用いる。また、低周波成分のエネルギーが一定になるように外力 $f(t)$ を加える。

本報告書では $\nu = 0.058$ のもとで計算した流体に対して、流体のマクロ変数としてエネルギー関数に関する時間発展モデリングを行う。波数 $k \in \mathbb{N}$ におけるエネルギー関数 $E_0(k, t)$ を次で定義する:

$$E_0(k, t) := \frac{1}{2} \int_{D_k} \sum_{\zeta=1}^3 |\mathcal{F}_{[v_\zeta]}(\kappa, t)|^2 d\kappa,$$

ただし、 $D_k := \{\kappa \in \mathbb{Z}^3 | k-0.5 \leq |\kappa| < k+0.5\}$ とする。 $\mathcal{F}_{[v_\zeta]}(\kappa, t)$ は v のフーリエ変換を表す。これ以降、 $k = 3$ のエネルギー関数 $E_0(k, t)$ を扱う。低周波成分や高周波成分は機械学習に

よるモデリングが困難になるため、時間方向のフィルターをかけて取り除く。これ以降、このフィルターをかけて得たものをエネルギー関数 $E(t)$ と呼ぶこととする。エネルギー関数 $E(t)$ の時系列データのみから、 $E(t)$ の時間発展モデルを構成する。

入出力変数. リザーバーコンピューティングによる時間発展モデリングをする際に、学習する時系列データの背後にある力学系の次元に比べて観測可能な変数の数が少ない場合は時間遅れ座標系が有効であることがわかっている [3]。そこで、エネルギー関数 $E(t)$ の時間発展モデルを構成するために、入出力変数を E を用いて M 次元の $\Delta\tau$ 時間遅れ座標系 $\mathbf{u}(t) = (E(t), E(t - \Delta\tau), \dots, E(t - (M - 1)\Delta\tau))^T$ を導入する。

モデルの構成. この時間遅れ座標系を入力とし、リザーバーコンピューティングのパラメータを表 1 となるように設定し、学習を行った。これ以降、あるひとつの行列の組 \mathbf{A} , \mathbf{W}_{in} , $\mathbf{W}_{\text{out}}^*$, $\mathbf{W}_{\text{Qout}}^*$ からなる機械学習モデルを考える。

パラメタ		値
M	入出力の次元	23
N	リザーバー状態ベクトルの次元	80
L	学習データ量	1×10^5
ρ	\mathbf{A} のスペクトル半径	1.5
σ	\mathbf{W}_{in} の生成に用いる入力強度	0.4
ξ	リザーバーモデルのバイアス項	0.1
α	モデル (1) の非線形性	0.7
β	正則化パラメタ	$5 * 10^{-5}$

表 1 リザーバーモデルの構成に用いたパラメタとその値.

結果. 構成した機械学習モデルを用いてエネルギー関数 $E(t)$ の時系列予測を行った。図 1

は様々な時刻から予測開始した様子を示したものである。開始する時刻によって、差はあるもののどの予測も時刻 50 程度までは変数予測がうまく行っていることがわかる。

次に、長時間発展のエネルギー関数 $E(t)$ の振る舞いを確認するため、機械学習モデルを長時間発展させて得た時系列データを図 2 に示した。流体では時間間欠的にエネルギーが大きい値を取るところが見られるが、機械学習モデルでもその様子が再現されていることがわかる。さらに、値が 10 を超えるようなまれにしか起きない振る舞いも再現していることがわかる。

また、エネルギー関数 $E(t)$ とその時間遅れ変数 $E(t - \Delta\tau)$ の振る舞いを確認するため、横軸を $E(t)$ 縦軸を $E(t - \Delta\tau)$ として 2 次元プロットを図 3 に示した。機械学習モデルは 2 次元プロットの概形をおよそ再現していることがわかった。これは、時間遅れ関係をデータから学習できているということを表している。

参考文献

- [1] Keiichi Ishioka, GFD Dennou Club, (1999), <http://www.gfd-dennou.org/arch/ispack/>.
- [2] Akane Ohkubo, Masanobu Inubushi, “Reservoir computing with generalized readout based on generalized synchronization”, Scientific Reports 14, 1 (2024).
- [3] Kengo Nakai and Yoshitaka Saiki, “Machine-learning construction of a model for a macroscopic fluid variable using the delay-coordinate of a scalar observable”, Discrete and Continuous Dynamical Systems Series S 14, 3 (2021).

6 進捗状況の自己評価と今後の展望

本制度の計算資源を利用することにより当初の予定通り流体の時間発展モデリングを行った。本研究結果はまとめて論文として出版する予定である。

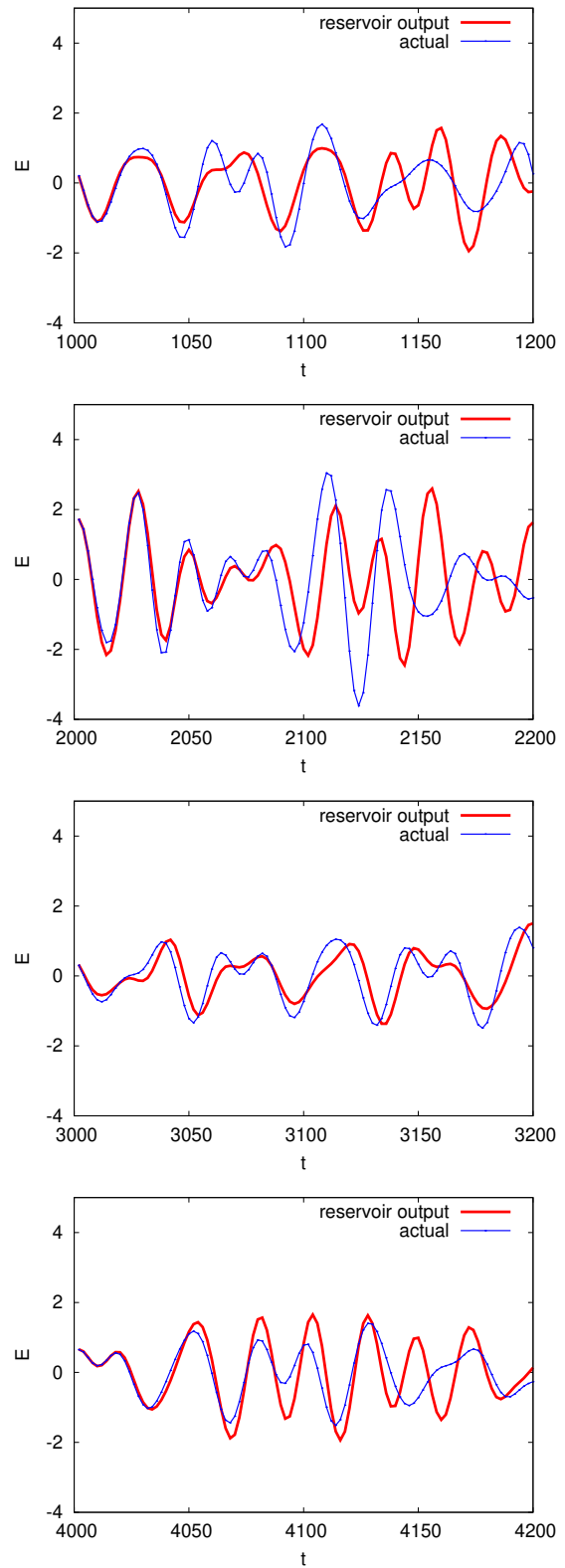


図 1 時系列予測 機械学習モデルを用いてエネルギー関数 $E(t)$ を時刻 1000, 2000, 3000, 4000 から予測した様子を赤線で描いた。比較のため Navier-Stokes 方程式の時間発展により得た $E(t)$ を青線で描いた。

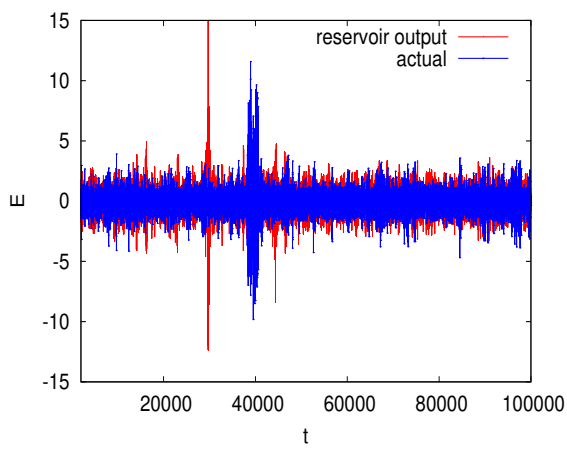


図 2 長時間発展のエネルギー関数 $E(t)$ 機械学習モデルを時刻 10^5 まで時間発展させて得た時系列データを赤線で描いた。比較のため Navier-Stokes 方程式の時間発展により得た $E(t)$ を青線で描いた。Navier-Stokes 方程式のエネルギー関数は時間間欠的に大きな値を取るが、機械学習モデルも同様に時間間欠的に大きい値を取っている様子が見取れる。さらに、値が 10 を超えるようなまれにしか起きない振る舞いも再現していることがわかる。

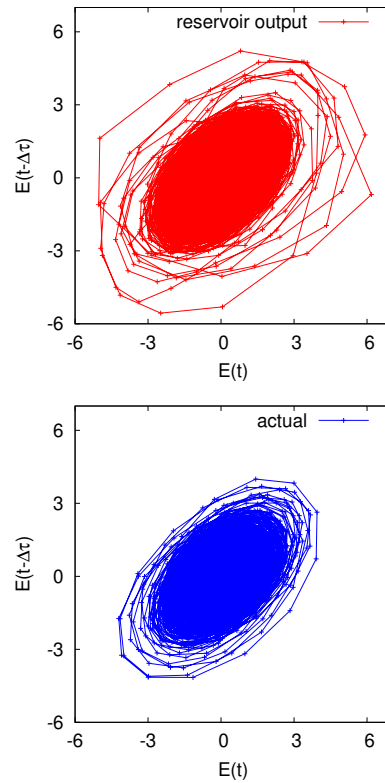


図 3 時間遅れ座標系上での時系列データの振る舞い横軸を $E(t)$ 、縦軸を $E(t-\Delta\tau)$ として 2次元平面上に、機械学習モデル (赤色) と Navier-Stokes 方程式 (青色) から得た 20000 時間分のデータをそれぞれプロットした。振幅が小さいときの時間遅れ平面上的軌道の概形はおおよそ再現している。一方で、振幅が大きいところについては頻度が少ないため、若干の相違がある。