

機械学習による流体変数の予測

中井 拳吾 (Kengo NAKAI)* , 齊木 吉隆 (Yoshitaka SAIKI)†

概要

カオス的なふるまいを示す流体運動のマクロ変数に関する時間発展方程式を流体の基礎方程式である Navier-Stokes 方程式から解析的に導出することは困難なことが知られている。我々は、リザーバーコンピューティングと呼ばれる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。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、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。

1 導入.

機械学習は様々な分野で注目されている。近年、機械学習の一種であるリザーバーコンピューティング [1, 2] が時系列データやリャプノフ指数などの予測において有効であることが報告されている [2, 3]。リザーバーコンピューティングにおける学習は入力データから得られたリザーバーベクトルと出力するベクトルを線形にフィッティングさせることに重点が置かれている。この学習はニューラルネットワーク構造を学習しないためフィッティングにかかる計算コストを減らしている。そこで我々はリザーバーコンピューティングの手法を用いて流体変数の挙動を予測した。

2 章でリザーバーコンピューティング, 3 章で流体の説明をしたのちに, 4 章で流体マクロ変数としてエネルギー関数に対して行った予測について述べる。さらに, 5 章では利用可能な時系列データが流体運動の不安定次元よりも少ない場合の時系列データの予想について述べる。

2 リザーバーコンピューティング.

$d\phi/dt = f(\phi)$ で表される力学系の変数, $\mathbf{u} = \mathbf{h}_1(\phi) \in \mathbb{R}^M$ と $\mathbf{s} = \mathbf{h}_2(\phi) \in \mathbb{R}^P$ について考える。ただし, ある時刻まで (学習時間と呼ぶ) の入力変数 \mathbf{u} , 出力変数 \mathbf{s} の時系列は既知とする。出力変数 \mathbf{s} の時系列を予測する時刻 (予想時間と呼ぶ) において, 入力変数 \mathbf{u} の時系列データは未知とする。学習時間において \mathbf{u} を要素分解することで得た高次元ベクトル \mathbf{r} の時系列データに対して出力が \mathbf{s} を近似できるように \mathbf{r} と \mathbf{s} の線形関係を決定する。この決定がリザーバーコンピューティングにおける学習に相当する。

このように物理的な知見を用いずに時系列データの学習のみからモデルを構築する。また, ニューラルネットワークの構造自体を学習する機械学習よりも圧倒的に計算量が少なくて済

*東京大学大学院 数理科学研究科 (〒153-8914 東京都目黒区駒場 3-8-1, E-mail: knakai@ms.u-tokyo.ac.jp)

†一橋大学大学院 経営管理研究科 (〒186-8601 東京都国立市中 2-1, E-mail: yoshi.saiki@r.hit-u.ac.jp)

むため、代わりにニューラルネットワークの次元を大きく出来る。そのためたとえダイナミクスが複雑なふるまいを示すとしても、それが決定論的である場合にはこの種の方法が有効である。

予測する時刻に \mathbf{u} の時系列データが既知ではない場合は Pathak らが行った考察 [3, 4] の特殊な場合に相当する。

リザーバーベクトル $\mathbf{r} \in \mathbb{R}^N$ ($N \gg M$) は次の式で時間発展させる。

$$\mathbf{r}(t + \Delta t) = (1 - \alpha)\mathbf{r}(t) + \alpha \tanh(\mathbf{A}\mathbf{r}(t) + \mathbf{W}_{\text{in}}\mathbf{u}(t)), \quad (1)$$

ただし、 Δt は時間ステップ幅を表す。 \tanh はそれぞれのベクトルの要素を引数とする。また、行列 \mathbf{A} , \mathbf{W}_{in} は $N \times N$, $N \times M$ の疎行列である。パラメータ α ($0 < \alpha \leq 1$) は \mathbf{r} の力学系の非線形性を表し、式 (1) の Δt に依存して決める。リザーバーベクトルは $(0, 1]$ の一様乱数から定めた $\mathbf{r}(-\tau)$ を初期値とする。ただし、 $\tau/\Delta t$ ($\gg 1$) はトランジェント時間に相当する。

行列 \mathbf{W}_{in} の各列はただ一つの要素のみに $[-\sigma, \sigma]$ 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 $D_1 \times N$ 個の非ゼロの要素を持つ疎行列とする。ただし、 $D_1 \times N$ 個の非ゼロ成分は $[-1, 1]$ の一様乱数から値を与える。さらに、行列 \mathbf{A} については最大固有値が ρ になるように規格化する。

リザーバーベクトル \mathbf{r} は次のようにして得たいベクトル $\hat{\mathbf{s}}(t)$ に変換する。

$$\hat{\mathbf{s}}(t) = \mathbf{W}_{\text{out}}\mathbf{r}(t) + \mathbf{c}. \quad (2)$$

これにより $L = T/\Delta t$ ステップのリザーバーベクトル $\{\mathbf{r}(l\Delta t)\}_{l=1}^L$ を得る。

$0 < t \leq T$ (学習時間と呼ぶ) での時系列 $\hat{\mathbf{s}}$ が既知の時系列 \mathbf{s} の近似になるように \mathbf{W}_{out} と \mathbf{c} を決定する。この決定がリザーバーコンピューティングの学習に当たる。これは次の二次形式が最小化するように \mathbf{W}_{out} と \mathbf{c} を決定することに対応する。

$$\sum_{l=1}^L \|(\mathbf{W}_{\text{out}}\mathbf{r}(l\Delta t) + \mathbf{c}) - \mathbf{s}(l\Delta t)\|^2 + \beta[\text{Tr}(\mathbf{W}_{\text{out}}\mathbf{W}_{\text{out}}^T)], \quad (3)$$

ただし \mathbf{q} に対して $\|\mathbf{q}\|^2 = \mathbf{q}^T\mathbf{q}$ とし、 $\beta(\geq 0)$ を含む項は \mathbf{W}_{out} の過剰適合を避けるために導入している。この決定が妥当であれば、 $t > T$ (予測時間と呼ぶ) での $\hat{\mathbf{s}}(t)$ は $\mathbf{s}(t)$ を再現できるはずである。二次形式 (3) を最小化する解を $\mathbf{W}_{\text{out}}^*$ と \mathbf{c}^* とすると予測時間では

$$\hat{\mathbf{s}}(t) = \mathbf{W}_{\text{out}}^*\mathbf{r}(t) + \mathbf{c}^*, \quad (4)$$

として時系列 $\hat{\mathbf{s}}$ を得る。 $\mathbf{W}_{\text{out}}^*$ と \mathbf{c}^* は次のように書くことができることが知られている ([5] P.140):

$$\begin{aligned} \mathbf{W}_{\text{out}}^* &= \delta\mathbf{S}\delta\mathbf{R}^T(\delta\mathbf{R}\delta\mathbf{R}^T + \beta\mathbf{I})^{-1}, \\ \mathbf{c}^* &= -[\mathbf{W}_{\text{out}}^*\bar{\mathbf{r}} - \bar{\mathbf{s}}], \end{aligned}$$

ただし、 $\bar{\mathbf{r}} = \sum_{l=1}^L \mathbf{r}(l\Delta t)/L$, $\bar{\mathbf{s}} = \sum_{l=1}^L \mathbf{s}(l\Delta t)/L$ とし、 \mathbf{I} は $N \times N$ の単位行列、 $\delta\mathbf{R}$ (rsp, $\delta\mathbf{S}$) は l 行が $\mathbf{r}(l\Delta t) - \bar{\mathbf{r}}$ (rsp, $\mathbf{s}(l\Delta t) - \bar{\mathbf{s}}$) となるような行列とする。

ただし、予測時間 $t > T$ での時系列データ \mathbf{u} は未知であるため、 $\mathbf{u}(t)$ として式 (1) で得た $\hat{\mathbf{s}}(t)$ を用いる:

$$\mathbf{r}(t + \Delta t) = (1 - \alpha)\mathbf{r}(t) + \alpha \tanh(\mathbf{A}\mathbf{r}(t) + \mathbf{W}_{\text{in}}\hat{\mathbf{s}}(t)).$$

学習に用いる変数 $X(t)$ を対等に扱うために正規化した $\tilde{X}(t)$ を使うことにする:

$$\tilde{X}(t) = [X(t) - X_1]/X_2.$$

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、学習時間の正規化に用いた X_1 と X_2 を用いる。

3 流体.

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:

$$\begin{cases} \partial_t v - \nu \Delta v + (v \cdot \nabla)v + \nabla \pi = f, & \nabla \cdot v = 0, & \mathbb{T}^3 \times (0, \infty), \\ v|_{t=0} = v_0 & \text{with } \nabla \cdot v_0 = 0, & \mathbb{T}^3, \end{cases}$$

ただし, $\mathbb{T} = [0, 2\pi)$, $\nu > 0$ は動粘性, $\pi(x, t)$ は圧力, $v(x, t) = (v_1(x, t), v_2(x, t), v_3(x, t))$ は速度とする. 切断波数を $(2 \times 9 + 1)^3$ としてフーリエスペクトル法 [6] を用いる. また, 4次のルンゲクッタ法により時間積分をする. また, 低周波成分のエネルギーが一定になるように毎時刻外力を加える. さらに低周波成分にのみエネルギーを持つように初期値を定める.

流体のマクロ変数としてエネルギー関数を考察する. 粘性は $\nu = 0.058$ とする. 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:

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

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_ζ のフーリエ変換を表す. 微細な振動を除くために短時間の時間平均をとった $E(k, t) = \sum_{s=t-99\Delta t}^t E_0(k, s)/100$ を考察する. これによりエネルギー関数の本質的な挙動を見ることができる. これ以降時間平均を取った E をエネルギー関数と呼ぶことにする.

4 エネルギー関数の挙動の予測.

以下のように時系列データ $\mathbf{u}(t), \mathbf{s}(t)$ を設定し, リザーバーコンピューティングにより学習を行う.

$$\begin{aligned} \mathbf{u}(t) &= (\tilde{E}(1, t), \tilde{E}(2, t), \dots, \tilde{E}(9, t))^t, \\ \mathbf{s}(t) &= (\tilde{E}(1, t), \tilde{E}(2, t), \dots, \tilde{E}(9, t))^t, \end{aligned}$$

ただし, パラメータは [7] の表 I (b) を用いる.

$\tilde{E}(1, \cdot)$ から $\tilde{E}(9, \cdot)$ の9次元の学習によってエネルギー関数の予測に成功した. 図1 (左上) (右上) に $\tilde{E}(4, t)$ と $\tilde{E}(9, t)$ の時系列データを書き出した. 流体のカオス性に由来する初期値鋭敏性により長時間時系列の予測が失敗することが予測される. 実際, 図1 (左下) に書き出したエネルギー関数のエラーの成長は $t - T \lesssim 100$ において指数関数的な増加が確認できた. 一方で, エネルギー関数 $E(k, t)$ 各波数 k の時間平均によって得られるエネルギースペクトル $\bar{E}(k) = \langle E(k, t) \rangle$ について, 時刻 $1000 < t - T < 2000$ の予測した時系列データから再構成することに成功した (図1) (右下). このことから速度のデータなどの微視的な時系列データを使わずに構成したりザーバーによる力学系は統計量もうまく再現できていることがわかる.

以上のことから, このリザーバーの力学系はエネルギー関数の挙動を表現できているカオス力学系として見るることができることが示唆されている.

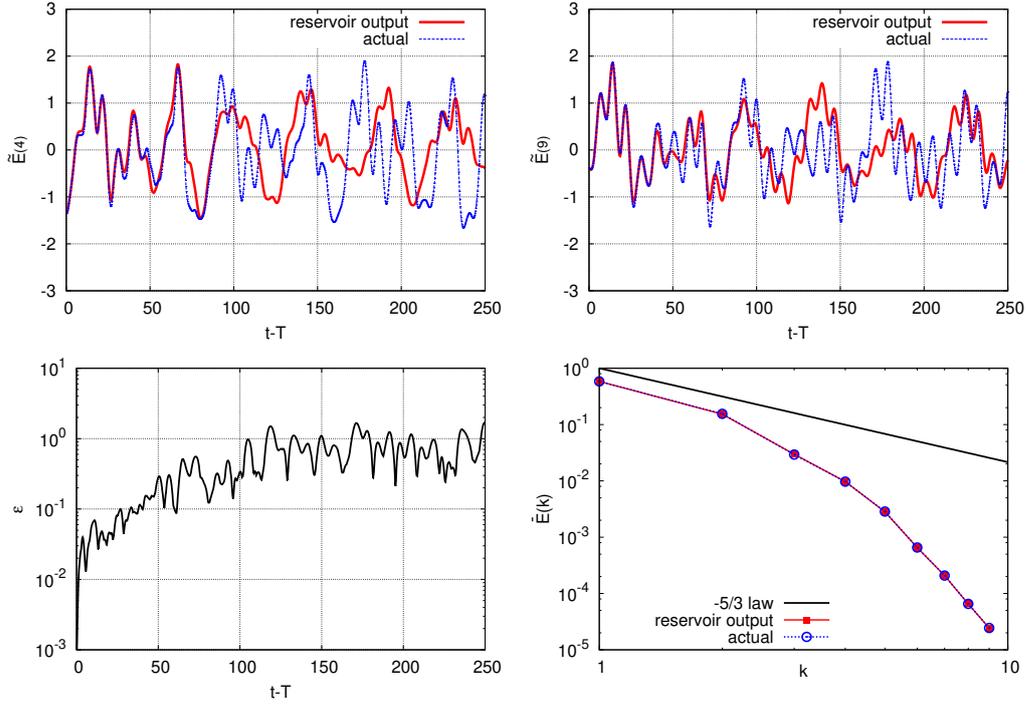


図 1: 流体マクロ変数の時間発展と統計の時系列データの予測。リザーバーコンピューティングにより $k = 4$ (左上) と 9 (右上) のエネルギー関数 $\tilde{E}(k, t)$ の時系列データの挙動を予測した (赤色)。また、比較のため Navier-Stokes 方程式の直接数値計算により得たエネルギー関数の時系列データも書きだした (青色)。これらの結果から時刻 100 程度までは赤色 (リザーバー) 青色 (DNS) は同期している。 $\varepsilon(t) = \sum_{k=1}^{N_0} |\tilde{E}(k, t) - \hat{\tilde{E}}(k, t)| / N_0$ ($N_0 = 9$) で定義される予測の誤差はカオスに見られるように時刻 100 に向かって指数的に大きくなっている (左下)。このことから時系列の予測のずれ方がある程度妥当であることがわかる。予想が直接数値計算の結果からずれた時刻の時系列データを元にエネルギースペクトル $\bar{E}(k)$ を計算した (右下)。また、比較のため Navier-Stokes 方程式の直接数値計算により得たエネルギー関数の時系列データも書きだした。各波数 k の予測したエネルギースペクトル $\bar{\tilde{E}}(k)$ の相対的なエラーは最大で 1.3% である。このことから時系列の予測がずれているところも統計量の意味で予測がうまく行っていることがわかる。

5 1 変数時系列のみの時間遅れ座標系を用いた予測。

高次元力学系の実験や観測において、観測できる変数とそのアトラクターのリアプノフ指数よりも少ないことはしばしばある。このような場合に、時間遅れ座標系 [8, 9] によって得た高次元データ \mathbf{u} を利用することでより効率よく予測することができる。リザーバーベクトル \mathbf{r} は行列 \mathbf{A} によっては過去の入力データ \mathbf{u} の影響を記憶できるが、計算量が莫大になるのでこの種の手続きは非常に有効である。

ここで、4 章で考察した同じエネルギー関数 $E(4, t)$ の予測を考察する。ただし、4 章で学習に用いた $E(k, t)$ ($k = 1, \dots, 9$) の中で $E(4, t)$ のみが既知だとする。遅れ時間 $\Delta\tau = 2.5$ として 36 次元の時間遅れ系を用いる。つまり、次のようにする。

$$\begin{aligned} \mathbf{u}(t) &= (\tilde{E}(4, t), \tilde{E}(4, t - \Delta\tau), \dots, \tilde{E}(4, t - 35\Delta\tau))^t, \\ \mathbf{s}(t) &= (\hat{\tilde{E}}(4, t), \hat{\tilde{E}}(4, t - \Delta\tau), \dots, \hat{\tilde{E}}(4, t - 35\Delta\tau))^t. \end{aligned}$$

図 2 に予測した $\tilde{E}(4, t)$ の時系列データを書きだした。 $\tilde{E}(1, \cdot)$ から $\tilde{E}(9, \cdot)$ の 9 次元の学習によって予測した (図 1) (左上) と同じように成功したのを見て取れる。ただし、パラメータは [7] の表 I (c) を用いる。

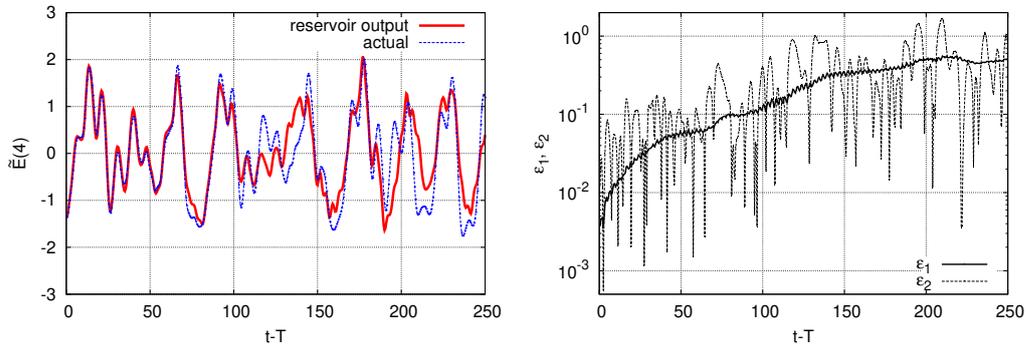


図 2: 時間遅れ座標系を用いたマクロ変数の予測。エネルギー関数 $\tilde{E}(4, t)$ の時系列のみの学習により図 1 (左上) と同じ時刻の $\tilde{E}(4, t)$ の挙動を予測した (左)。比較のため Navier-Stokes 方程式の直接数値計算により得たエネルギー関数の時系列データも書きだした。 $\varepsilon_1(t) = \sum_{n=0}^{35} |\tilde{E}(4, t - n\Delta\tau) - \hat{\tilde{E}}(4, t - n\Delta\tau)|^2 / 36$ と $\varepsilon_2(t) = |\tilde{E}(4, t) - \hat{\tilde{E}}(4, t)|$ で定義される予測の誤差を書きだした (右)。

謝辞 中井は、東京大学数物フロンティア・リーディング大学院 (FMSP) の助成を受けている。齊木は、JSPS 科研費 JP17K05360, JST さきがけ JPMJPR16E5 の助成を受けている。また、本研究でおこなった計算の一部は京都大学のスーパーコンピュータ共同研究制度 (若手・女性奨励枠) に基づく。

参考文献

- [1] Herbert Jaeger and Harald Haas. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, 304:78–80, 2004.
- [2] Zhixin Lu, Jaideep Pathak, Brian Hunt, Michelle Girvan, Roger Brockett, and Edward Ott. Reservoir observers: Model-free inference of unmeasured variables in chaotic systems. *Chaos*, 27:041102, 2017.
- [3] Jaideep Pathak, Zhixin Lu, Brian Hunt, Michelle Girvan, and Edward Ott. Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data. *Chaos*, 27:121102, 2017.
- [4] Jaideep Pathak, Brian Hunt, Michelle Girvan, Zhixin Lu, and Edward Ott. Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Phys. Rev. Lett.*, 120:024102, 2018.
- [5] M. Lukosevicius and Herbert Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3:127–149, 2009.
- [6] Keiichi Ishioka. ispack-0.4.1. <http://www.gfd-dennou.org/arch/ispack/>, 1999. GFD Dennou Club.

- [7] Kengo Nakai and Yoshitaka Saiki. Machine-learning inference of fluid variables from data using reservoir computing. *Phys. Rev. E*, 98:023111, 2018.
- [8] F Takens. Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980 (Coventry, 1979/1980)*, volume 898 of *Lecture Notes in Math.*, pages 366–381. Springer, Berlin-New York, 1981.
- [9] T Sauer, J A Yorke, and M Casdagli. Embedology. *J. Stat. Phys.*, 65:579–616, 1991.