jh210027-NAH

機械学習モデルのリアプノフ指数ならびに リアプノフベクトルの解析

齊木 吉隆（一橋大学）

概要

近年、機械学習の一種で、リカーレントネットワークをもったリザーバーコンピューティングが決定論的構造を背後にもつ時系列予測などに有効であることが報告されている。我々はこの機械学習手法を用いて比較的平穏な流体のエネルギー変数のモデルを構成し、予想を成功させている (Nakai and Saiki, Physical Review E, 2018)。時間発展モデリングにおける機械学習の適切な活用のために、どのようなクラスの決定論的ダイナミクスが機械学習を用いてどのような観点、どのような精度でモデル化しうるかといったことを明らかにすることが、期待されている。本研究では、力学系において重要な構造のひとつである安定多様体と不安定多様体の接構造をもつダイナミクスが機械学習でモデリング可能であることを機械学習モデルの詳細な力学系解析を通して明らかにした。

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　京都大学

1.2　共同研究分野

■超大規模数値計算系応用分野

1.3　参加研究者の役割分担

本共同研究は 4 人体制でおこなわれている。

- 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 (東京海洋大学 学術研究院　流通情報工学部門): 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

研究代表者らは、流体乱流の時系列予測を可能とする機械学習を活用した時間発展モデリングに 2018 年に世界ではじめて成功し、短時間予測や長時間統計分布の再現を確認した。一方、力学系理論の視点で機械学習時間発展モデルを詳細に解析した研究や、訓練データがもつ力学系的な情報を機械学習で得た時間発展モデルが保持するかといった観点の研究は不十分な状況にある。どのようなクラスの決定論的ダイナミクスが機械学習を用いてどのような観点、どのような精度でモデル化しうるかといったこ

とが明らかになることによって、時間発展モデリングにおける機械学習の適切な活用が広がると期待される。本研究では、力学系理論の分野で長年研究をしてきた研究代表者らが、力学系の不変集合、その安定性/不安定性、構造安定性、双曲性/非双曲性、間欠性などの概念に着目して時間発展モデリングの可能性、限界を明らかにする。本研究は、力学系理論、数理モデリング、機械学習、大規模計算にまたがる学際的・融合的研究である。

力学系は大きく双曲力学系と非双曲力学系に分類される。双曲力学系は、相空間の各点で伸びる方向と縮む方向に"連続的に"分離可能な系であり、少なくとも幾何学的立場では S. Smale 以降の研究によって構造安定な力学系を特徴づける構造としてかなり理解が進んでいる。これは、記号力学系で表現され、摂動しても周期軌道の個数が変わらない。したがって、ロバストなモデリングが期待できる対象であると考えられる。ここでロバストなモデリングとは、訓練データを生み出す力学系 A は (十分多くの訓練データとニューラルネットワークの次元のもとで) 精度良く定まったどんな機械学習モデル B とも"位相共役"になることであると考える。気象など間欠性ダイナミクスを示す多くの現象は非双曲力学系であると考えられ、その種の力学系では上記のような狭い意味でのロバストなモデリングは期待できない。しかし、なんらかの広い意味でのロバストなモデリング可能性は力学系理論の観点からも期待できる。近年の力学系研究によれば、二つの性質: (i) 安定多様体と不安定多様体が接する構造 (Tangency) (ii) 不安定次元が異なる周期軌道の安定/不安定多様体の交差 (ヘテロ次元サイクル)、が非双曲力学系にあらわれる典型的な構造が非双曲力学系の典型的構造であると考えられている。本研究ではその二つの構造のうち、(i) に着目した研究を推進した。

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、研究代表者と中井拳吾氏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。

本研究課題では時系列データの機械学習によるモデリングを行った。時系列データの力学系構造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。

このためには、JHPCN の枠組みを最大限に活用し、長時間の連続計算が可能な京都大学の system A での長時間データの学習と、非常に高い並列計算を得意とする東京大学の Oakforest-PACS での力学系解析を使い分けて計算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

新規課題のため該当しない。

5　今年度の研究成果の詳細

本計算資源を利用した研究では、(i) の安定多様体と不安定多様体が接する非双曲構造

(Tangency) を持つ力学系について研究を実施し、当初の目標であった研究項目 1(機械学習によるモデリング)、研究項目 2(時系列データから空間微分の計算)、研究項目 3(リャプノフ解析) のいずれも研究が完了し論文 [a] として出版されている。

まずはじめに本研究の要である機械学習モデリングの学習手法であるリザーバーコンピューティングについて言及した後、研究項目 1∼3 の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

$$\frac{d\phi}{dt} = \mathbf{f}(\phi)$$

に関し、その変数 $\mathbf{u} = \mathbf{h}_1(\phi) \in \mathbb{R}^M$ と $\mathbf{s} = \mathbf{h}_2(\phi) \in \mathbb{R}^P$ について考える。ただし、時系列データ $\mathbf{u}$ は $\mathbf{s}$ の時系列を予測する時刻にデータはないとする。リザーバーベクトル $\mathbf{r} \in \mathbb{R}^N$ $(N \gg M)$ は次の式で支配されている。

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

ただし、Δt は時間幅を表す。また、行列 $\mathbf{A}$, $\mathbf{W}_{\mathrm{in}}$ は $N \times N$, $N \times M$ 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 $\mathbf{W}_{\mathrm{in}}$ の各列はただ一つのみに $[-\sigma, \sigma]$ 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 $D \times N$ 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 $\mathbf{A}$ については最大固有値が ρ になるように行列のすべての要素を調整する。

リザーバーベクトル $\mathbf{r}$ は

$$\hat{\mathbf{s}}(t) = \mathbf{W}_{\mathrm{out}}\mathbf{r}(t)$$

により、得たいベクトルに変換する。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 $\mathbf{W}_{\mathrm{out}}$ 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相当する。これは次の二次形式が最小化するように $\mathbf{W}_{\mathrm{out}}$ を決定することに対応する。

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

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 $\mathbf{W}_{\mathrm{out}}$ 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 (2) を最小化する解を $\mathbf{W}_{\mathrm{out}}^*$ とすると予測時間では

$$\hat{\mathbf{s}}(t) = \mathbf{W}_{\mathrm{out}}^*\mathbf{r}(t), \quad (3)$$

として時系列 $\hat{\mathbf{s}}$ を得る。$\mathbf{W}_{\mathrm{out}}^*$ は次のように書くことができることが知られている:

$$\mathbf{W}_{\mathrm{out}}^* = \delta\mathbf{S}\delta\mathbf{R}^T(\delta\mathbf{R}\delta\mathbf{R}^T + \beta\mathbf{I})^{-1}, \quad (4)$$

ただし、

$$\overline{r} = \sum_{l=1}^{L} \mathbf{r}(l\Delta t)/L, \quad \overline{s} = \sum_{l=1}^{L} \mathbf{s}(l\Delta t)/L$$

とし、$\mathbf{I}$ は $N \times N$ の単位行列、$\delta\mathbf{R}$, $\delta\mathbf{S}$ はそれぞれ l 行が $\mathbf{r}(l\Delta t) - \overline{r}$, $\mathbf{s}(l\Delta t) - \overline{s}$ 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 $\mathbf{u}$ は利用できないため、$\mathbf{u}(t)$ として式 (1) で得た $\hat{\mathbf{s}}(t)$ を用いる:

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

学習に用いる変数 $X(t)$ を対等に扱うために正規化した $\tilde{X}(t)$ を用いる: $\tilde{X}(t) = [X(t) - X_1]/X_2$. 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。この正規化より σ の調整を避けることができる。

5.1 研究項目 1: 機械学習によるモデリング

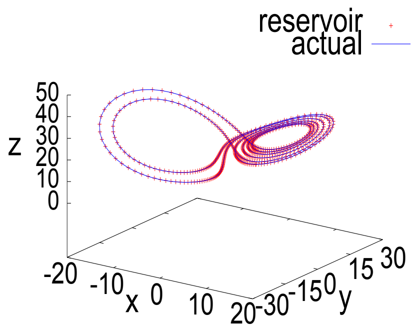
決定論的な時系列データとしてローレンツ方程式の学習を行った。ローレンツ方程式はあるパラメータでのもとでカオスな振る舞いをすることが知られている。

ローレンツ方程式

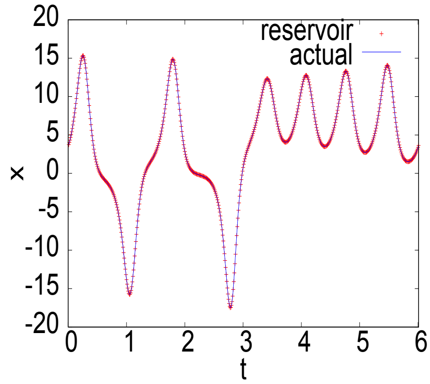
$$\begin{cases} \dot{x} = -10x + 10y \\ \dot{y} = -xz + rx - y \\ \dot{z} = xy - \frac{8}{3}z. \end{cases}$$

に関して、本研究ではパラメタ r として古典パラメタである 28 に加えて、安定多様体と不安定多様体の接する構造が存在すると考えられる 60 について、ローレンツ方程式の時間により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により学習させて得られた機械学習時間発展モデルが、元の力学系構造をどの程度再現するかを解析する。 $r = 60$ のパラメタにおいては力学系が構造不安定的であり、どの程度の精度でモデリングが可能となるか明らかではなかった。

しかし、本計算資源を使うことでいずれのパラメタにおいても、時間発展予測が可能な高精度な機械学習モデルの構成に成功した。実際、さまざまな初期値から時間発展予測が可能であることを確認している。特にカオス力学系における骨格をなす基本的な軌道である不安定周期軌道を機械学習モデルが近似することを確認した (図 1)。



(a)



(b)

図 1 **A periodic orbit-like trajectory.** (a) A periodic orbit-like trajectory obtained from the data-driven model is plotted together with the corresponding unstable periodic orbit (period T_p =5.9973192969) of the actual Lorenz system with $r = 28$, and (b) their time developments of the x variable. (論文 a Fig.4)

5.2 研究項目 2: 時系列データから空間微分の計算

適当なカオス的時系列データから背後にある微分方程式を推定する手法を用いて、研究項目 1 で構成した機械学習モデルを時間発展させることで得た時系列データからその支配方程式

$$\frac{d\mathbf{X}}{dt} = F(\mathbf{X}), \quad \mathbf{X} \in \mathbf{R}^D, \tag{5}$$

を推定した。そして、この式を X で偏微分することにより研究項目 3 のリャプノフ解析の計算に用いるヤコビ行列を求めた。

微分方程式モデル (式 (5) に相当) を以下の手続きを経て推定する。

1. リザーバーモデルで生成した変数 x の時系列を用いて微分方程式を推定するために、変数 x を成分毎に標準化して得られた変数を X とする。
2. テイラー展開を用いて観測された各離散時刻における X の時間微分 (式 (5) の左辺) を推定する。
3. 式 (5) の右辺を X に関する多項式として表現するモデルを定め、その係数をリッジ回帰を用いて求める。項目 2 で推定された X の時間微分をあらわす (式 (5) が定まる)。
4. 項目 3 で求まった式 (5) を X で偏微分することによってヤコビ行列を得る。

離散時間で観測可能な $\mathbf{X}(t)$ $(t=\tilde{t}+k\delta)$ を用いて、テイラー近似によって各観測時刻における時間微分値を推定する．本研究では以下の 6 次テイラー展開による．

$$\begin{aligned}\frac{d\mathbf{X}(\tilde{t})}{dt} \approx \frac{1}{60\delta}\{&\mathbf{X}(\tilde{t}+3\delta)-9\mathbf{X}(\tilde{t}+2\delta)\\ &+45\mathbf{X}(\tilde{t}+\delta)-45\mathbf{X}(\tilde{t}-\delta)\\ &+9\mathbf{X}(\tilde{t}-2\delta)-\mathbf{X}(\tilde{t}-3\delta)\}.\end{aligned}$$

多項式の線形結合によるモデル

$$\begin{aligned}F_k(\boldsymbol{X}) \approx \tilde{\beta}_0^k + &\sum_{d=1,\cdots,D} \tilde{\beta}_d^k X_d \qquad (6)\\ &+ \sum_{1\le i\le j\le D} \tilde{\beta}_{iD+j}^k X_i X_j + \cdots,\end{aligned}$$

$F_k(\boldsymbol{X})$ $(k=1,\ldots,D)$ は k 番目の変数の成分であり、$\tilde{\beta}_l^k$ $(l=0,1,\ldots,D^2+D)$ をリッジ回帰で推定する．

5.3 研究項目 3: リャプノフ解析

研究項目 2 で得られた空間微分の情報をもとに力学系の基本的な量であるリャプノフ指数やリャプノフベクトルを機械学習モデルに対して計算した。図 2 はリャプノフベクトルから計算される安定多様体と不安定多様体のなす角の分布を書き出したものである。比較のため元のローレンツ方程式から計算されるなす角の分布も合わせてかきだしている。$r=28,60$ いずれの場合についても機械学習モデルのリャプノフ指数の分布は元のローレンツ方程式のなす角の分布をよく再現していることが見て取れる。

また、$r=28,60$ いずれの場合のリャプノフ指数についても元のローレンツ方程式のリャプノフ指数をよく再現していることが見て取れる。研究項目 1 から 3 によって安定多様体と不安定多様体が接する構造をもつ非双曲型力学系についても機械学習によって時系列モデリングが可能であることがわかった。これらの結果は論文 [a] として出版されている。

6 今年度の進捗状況と今後の展望

実施計画において次の3 つの項目

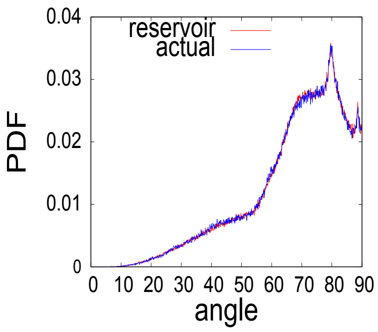
1. 機械学習によるモデリング
2. 時系列データから空間微分の計算
3. リャプノフ解析

を予定していた。

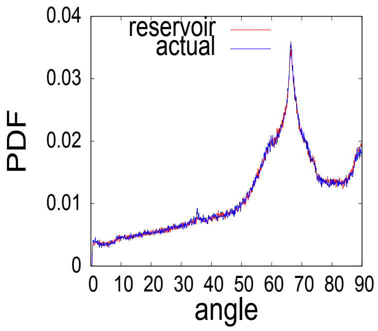
このいずれも予定通り研究が進捗し、非双曲的な力学系構造において典型的な二つの構造のうち、安定多様体と不安定多様体の接構造をもつダイナミクスに関して得た、研究項目 1,3 のモデリングの結果とリャプノフ解析の計算結果については既に論文 [a] として出版されている。また研究項目 2 の時系列データのみからの空間微分の推定については、より一般化して、適当なカオス的時系列データから背後にある微

r	$\lambda^{(1)}_{\rm res}$	$\lambda^{(2)}_{\rm res}$	$\lambda^{(3)}_{\rm res}$	$D^{KY}_{\rm res}$	$\lambda^{(1)}_{\rm actual}$	$\lambda^{(2)}_{\rm actual}$	$\lambda^{(3)}_{\rm actual}$	$D^{KY}_{\rm actual}$
28	0.901	0.000	-14.570	2.06	0.902	0.000	-14.570	2.06
60	1.402	0.000	-15.070	2.09	1.404	0.000	-15.071	2.09

表 1 **Lyapunov exponents and Lyapunov dimensions.** Lyapunov exponents of the data-driven model using reservoir computing ($\lambda^{(1)}_{\rm res}, \lambda^{(2)}_{\rm res}, \lambda^{(3)}_{\rm res}$) and those of the actual Lorenz system ($\lambda^{(1)}_{\rm actual}, \lambda^{(2)}_{\rm actual}, \lambda^{(3)}_{\rm actual}$) are listed. The values are computed using the four-stage and fourth-order Runge–Kutta method with time step $2\Delta t$ from the points along an orbit trajectory and the estimated Jacobian matrices. The Lyapunov dimensions $D^{KY}_{\rm res}$ for the data-driven model and $D^{KY}_{\rm actual}$ for the actual Lorenz system are estimated from the Kaplan–Yorke formula. (論文 a Table.III)



(a)



(b)

図 2 **Distribution of the angle between stable and unstable manifolds along a trajectory** ((a) $r = 28$ and (b) $r = 60$). The density distribution of the manifold angles (degree) at points along a trajectory is shown for a data-driven model using reservoir computing together with that of the actual Lorenz system. (論文 a Fig.5)

分方程式を推定する推定手法を開発し、ローレンツ方程式や流体変数のアトラクタ上でふるまう軌道の一変数のみの時系列データから微分方程式を推定した。推定した微分方程式を評価するため、短時間の時間発展予測に加えて、長時間発展によって得られた変数の出現頻度分布の再現性を確認した。この結果については論文として まとめて現在査読中である。

今後、非双曲的な力学系構造の中で典型的なもののうち、本研究の機械学習モデリングで取り扱っていないもう一つの構造、すなわち不安定次元が異なる不変集合とその間のサイクルが存在するダイナミクスの研究を計画している。その予備的研究としてその種の非双曲構造をもつ力学系を研究して論文 [b] を執筆した。

7 研究業績一覧（発表予定も含む）

学術論文（査読あり）

a. M. Kobayashi, K. Nakai, Y. Saiki, and N. Tsutsumi, “Dynamical system analysis of a data-driven model constructed by reservoir computing,” Physical Review E 104, 044215:1-7, 28th October 2021

b. Y. Saiki, H. Takahasi and J. A. Yorke, “Piecewise-linear maps with heterogeneous chaos,” Nonlinearity 34 (8) 5744-

5761, 13th July 2021

国際会議プロシーディングス （ 査読あり）

国際会議発表 （ 査読なし）

1. Kengo Nakai, Differential Equations for Data Science 2022 , “Constructing differential equations using only a chaotic time-series,” Online, 23rd March, 2022.
2. Yoshitaka Saiki, Jim A. Yorke’s 80th Birthday Celebration Workshop, “Low-dimensional paradigms for high-dimensional chaos,” online, 17th September, 2021.

国内会議発表 （ 査読なし）

3. 中井拳吾, 応用数学に関する研究発表会, “力学系不変集合の機械学習モデルによる再現性,” 芝浦工業大学, 2022 年 3 月 30 日
4. 中井拳吾, 2021 年度年会日本数学会, “力学系不変集合の機械学習モデルによる再現性,” 埼玉大学, 2022 年 3 月 30 日
5. 中井拳吾, 微分方程式とデータサイエンス研究会, “機械学習モデルの力学系構造の再現性と流体統計量の予測,” 金沢大学, 2022 年 1 月 17 日
6. 齊木吉隆, 冬の力学系研究集会, “機械学習モデルによる各種力学系的性質の再現性,” 九州大学, 2022 年 1 月 8 日
7. 中井拳吾, 応用数学合同研究集会, “機械学習モデルの力学系構造の再現性と流体統計量の予測,” online, 2021 年 12 月 18 日
8. 齊木吉隆, 応用数学合同研究集会, “カオス的時系列データに基づく微分方程式推定,” online, 2021 年 12 月 18 日

公開したライブラリ等

その他（ 特許，プレス発表，著書等）