2023 International Symposium on Nonlinear Theory and Its Applications
NOLTA2023, September 26-29, 2023, Catania and Online

Topological Inference of State Space as Effective Goal of Dynamics Learning

Taiki Yamada†, Kantaro Fujiwara†

†Graduate School of Information Science and Technology, University of Tokyo
7–3–1 Hongo, Bunkyo–ku, Tokyo 113-8656, Japan

Email: yamada-taiki@g.ecc.u-tokyo.ac.jp, kantaro@g.ecc.u-tokyo.ac.jp

Abstract— Our brain functions rely on the ability to detect the current states of the outside world and infer their evolution. Thus, studying the requirements to achieve dynamics learning is necessary to understand the brain.

Reservoir computing (RC) is a recent representative method for deterministic dynamics learning. Although both applicational and theoretical works support the potential of RC, few studies about the achievability of learning in realistic situations exist. This lack of understanding is fatal for further applications, especially when one wants to use the RC framework as a model of the brain.

Inspired by previous works, this paper will consider a relaxed but necessary goal of dynamics learning regarding topological inference of state space. This relaxation allows us to consider the achievability of learning with the finite number of learning iterations and the presence of noise in observations.

Considered learning goal enables us to treat both deterministic and stochastic observations in a unified way. Therefore, for future works, we expect the development of theories motivated by deterministic and stochastic notions, which facilitate further understanding of our brain.

1. Introduction

Recent developments in machine learning methods allow us to shift our attention from static information learning to dynamics learning [1, 2]. The mechanism of dynamics learning has been studied with focusing on the reservoir network (RC) as a learning method [1] by using the theory of dynamical systems [3].

However, there are few studies about the mechanism of dynamics learning considering the finiteness of samplings size and the effects of noise in computations potentially leading to learning failure. This is because results of dynamics learning were always reported through the achievements of necessary conditions [1, 4], although dynamics learning failure is comparably as common as success (See Supplementary Material). Therefore, investigations of the cases where dynamic learnings fail is crucial for future theoretical development and deeper understanding of the mechanism of dynamics learning.

ORCID iDs Taiki Yamada: 0009-0000-4216-3623, Kantaro Fujiwara: 0000-0001-8114-7837

This study focused on the inference of topological invariant called homology as a necessary condition of dynamics learning inspired by previous work [4]. We suggested a conjecture on the dependencies of success and failure of homology inference on the number of samplings and sampling noise. This study aims to investigate the validity of our conjecture for typical dynamical systems: a rotation map, an expanding map, the Lorenz equation, and the Rössler equation.

To state our conjecture, we first reviewed the relationships between homology inference and dynamics learning in 1.1, then briefly introduced homology in 1.2. Then, we proposed our conjecture and explained simulation settings for testing our conjecture in 2. In section 3, We illustrated simulation results and their indications.

1.1. Topological inference for invariant set as an necessary goal of Dynamics Learning

We denoted a sequence of samplings with length $N \in \mathbb{N} \cup \{\infty\}$ as $\mathcal{S}^{(N)} = \langle x_0, x_1, \dots x_N \rangle$, and the set of all points in $\mathcal{S}^{(N)}$ as $\{\mathcal{S}^{(N)}\} = \{x_i\}_{i=0}^{N}$.

In typical studies of dynamics learning, a sequence of samplings is an orbit $\mathcal{S}^{(N)}_{f-\mathrm{orbit}} = \langle x, f(x), \cdots, f^{N-1}(x) \rangle$ obtained by a target dynamical system (X, f), where X is a state space and f is a map from X to itself. The common goal of dynamics learning is to obtain an estimated dynamical system (Y, g), which is topologically conjugate to (X, f). Topological conjugacy between (X, Y) and (Y, g) implies X and Y has the same topological property. Thus, a sampling sequence $\mathcal{S}^{(N)}$ for dynamics learning with a target (X, f) should contain knowledge on the topological property of X. This fact motivates us to investigate whether the behavior of topological inference based on $\mathcal{S}^{(N)}_{f-\mathrm{orbit}}$ tells us about the significance of a map f in terms of dynamics learning. In this study, we focused on one of the topological properties of a space called homology. We considered dynamical systems which satisfy $\mathrm{cl}(\{\mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}\}) = X$, so that homology inference will surely success with $N \gg 1$.

1.2. Homology estimation by samplings

Homology is a topological invariant between topologically equivalent spaces. Roughly speaking, homology tells us the information of generalized holes in a topological space. For instance, a 0-dimensional hole is a connected

component, and a 1-dimensional hole is a cycle in topological space. For rigorous definitions of homology, see [5].

We now consider homology inference by using a finite set of samplings $\{\mathcal{S}^{(N)}\} \subset X \subset \mathbb{R}^d$. Since a finite set of samplings $\{\mathcal{S}^{(N)}\}$ itself has nothing other than the trivial homology: at most N separated components, one needs to consider the thicken set: $B_R(\{\mathcal{S}^{(N)}\}) = \bigcup_{i=1}^{N} B_R(x_i)$, where $B_R(x)$ is the closed ball of $\mathbb{R}^d$ with radius $R > 0$ centerd at x (Figure 1).

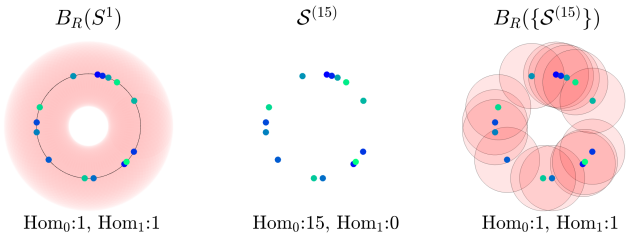

Figure 1: Homology estimation by using a finite number of samplings. Hom_0 and Hom_1 indicate the number of connected components and the number of cycles, respectively.

We used the algorithm [5] based on matrix reductions to compute the homology of the thicken set $B_R(\{\mathcal{S}^{(N)}\})$, which approximates the homology of $B_R(X)$, and thus so of X.

Notice that there is the interval $[\underline{R}, \overline{R}]$ of thickening radius R which results in the correct estimation of the homology of the target space. Hence, we needed to choose an appropriate value of radius $R \in [\underline{R}, \overline{R}]$ to infer the homology of the target space. The topological data analysis called the persistent homology solves this issue [6]. Thus, by using the persistent homology, we could reliably obtain an interval of appropriate values of radius $[\underline{R}, \overline{R}]$ and a sufficient number of samplings N for the success of homology estimation at each simulation.

2. Methods

2.1. Conjecture statement

For a given sequence of samplings $\mathcal{S}^{(\infty)}$ from X, we defined the inference result $I_{N,R}(\mathcal{S}^{(\infty)})$ by:

$$I_{N,R}(\mathcal{S}^{(\infty)}) = \begin{cases} 1 & \text{if } \mathrm{Hom}[B_R(\{\mathcal{S}^{(N)}\})] = \mathrm{Hom}[B_R(X)], \\ 0 & \text{otherwise.} \end{cases}$$

, where $\mathrm{Hom}[A]$ denotes the homology of a space A.

In order to investigate the effects of noise in samplings, we considered the psuedo-orbits of $\mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}$ restricted on $\{\mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}\}$ denoted as $\mathcal{S}^{(\infty)}_{f-\delta}$ with a parameter $\delta \geq 0$ (Figure 2). Specifically, we defined the i-th sampling of $\mathcal{S}^{(\infty)}_{f-\delta}$ as a randomly chosen element of $B_\delta(f^{i-1}(x)) \cap \{\mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}\}$. Note that $\mathcal{S}^{(\infty)}_{f-0} = \mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}$, and $\mathcal{S}^{(\infty)}_{f-\infty}$ is equivalent to a sequence of random samplings of points of $\mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}$. We denoted $\mathcal{S}^{(\infty)}_{f-\infty}$ as $\mathcal{S}^{(\infty)}_{f-\mathrm{random}}$ for clarity.

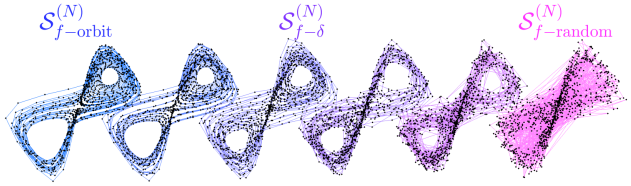

Figure 2: Examples of $\mathcal{S}^{(N)}_{f-\delta}$. Colored lines indicate temporal adjacency of samples.

The sequence of samplings $\mathcal{S}^{(\infty)}_{f-\delta}$ ($\delta \geq 0$) depends on a random choice of initial state. Additionally, for $\delta > 0$, $\mathcal{S}^{(\infty)}_{f-\delta}$ also depends on a sequence of random choices of a next state. Thus inference result $I_{N,R}(\mathcal{S}^{(\infty)}_{f-\delta})$, which is a function of the sequence of samplings $\mathcal{S}^{(\infty)}_{f-\delta}$, is a random variable. Then, we defined the parameter $\theta_{f-\delta}(N,R)$ by:

$$\theta_{f-\delta}(N,R) := \mathrm{Prob}\left[I_{N,R}(\mathcal{S}^{(N)}_{f-\delta}) = 1\right].$$

Here, we are ready to state our following conjecture for a given dynamical system (X, f):

Conjecture: *One can learn dynamics f by using $\mathcal{S}^{(\infty)}_{f-\delta}$ iff* $\exists N \in \mathbb{N}, \exists R > 0, \theta_{f-\delta}(N,R) \neq \theta_{f-random}(N,R)$.

This conjecture is motivated by the fact that no one can learn dynamics f by using $\mathcal{S}^{(\infty)}_{f-\mathrm{random}}$ since the stochasticity hides knowledge of dynamics f. In other words, we hypothesized that we could detect the gap between $\mathcal{S}^{(\infty)}_{f-\delta}$ and $\mathcal{S}^{(\infty)}_{f-\mathrm{random}}$ in terms of dynamics learnings through the difference of parameters $\theta_{f-\delta}(N,R)$ and $\theta_{f-\mathrm{random}}(N,R)$ which reflect the properties of homology inference. Note that the superiority of orbits than random sampling: $\theta_{f-\delta}(N,R) > \theta_{f-\mathrm{random}}(N,R)$ doesn't necessarily hold, since an psuedo-orbit $\mathcal{S}^{(N)}_{f-\delta}$ reflecting property of f could bias and prevent homology inference with intermidiate values of (N,R).

In this study, we conducted an approximation of the hypothesis testing with the null hypothesis $H_0 : \forall N \in \mathbb{N}, \forall R > 0, \theta_{f-\delta}(N,R) = \theta_{f-\mathrm{random}}(N,R)$. As we mentioned in 1.2, by using the persistent homology, we reliably estimated an adequate number of samplings and the thickening radius denoted as $N_{\max}$ and $\overline{R} < R_{\max}$ respectively. To test our null hypothesis H_0, we considered 10×10 gird choice of paris $\{(N_i, R_j)\}_{1 \leq i,j \leq 10}$, where $N_i = 1 + i\lfloor N_{\max}/10 \rfloor, R_i = i\lfloor R_{\max}/10 \rfloor$. For each pair (N_i, R_j), we conducted two-sided tests by using Fisher's exact test, which is implemented by the `fishertest` function in MATLAB, with the following statistical hypotheses:

$$\text{(Null) } H_0^{(i,j)} : \theta_{f-\delta}(N_i,R_j) = \theta_{f-\mathrm{random}}(N_i,R_j),$$
$$\text{(Alternative) } H_1^{(i,j)} : \theta_{f-\delta}(N_i,R_j) \neq \theta_{f-\mathrm{random}}(N_i,R_j).$$

Then, we supposed H_0 was rejected iff $H_0^{(i,j)}$ are rejected for some $1 \leq i, j \leq 10$. We conducted all tests by 10 samples of $\mathcal{S}^{(\infty)}_{f-\delta}$ with a significance level $\alpha = 0.01$ for each $\delta \in \{0, \texttt{logspace}(10^{-3}, \mathrm{diam}(X), 10)\}$, where `logspace` is the logspace function in MATLAB and $\mathrm{diam}(X)$ is the diameter of X.

2.2. Simulations

To illustrate the principle of our conjecture, we first conducted hypothesis testing at $\delta \in \{0, \infty\}$ for the rotation map f_{rot} and the expanding map f_{exp} on S^1 defined by:

$$x_{i+1} = f_{\mathrm{rot}}(x_i) = x_i + 1 + \epsilon_i \mod 2\pi, \quad (1)$$
$$x_{i+1} = f_{\mathrm{exp}}(x_i) = 2x_i + \epsilon_i \mod 2\pi. \quad (2)$$

ϵ_i in (1, 2) is a random variables which takes a 0 or ϵ with the same probability at each index i, where $\epsilon := 2^{-52}$ is the machine epsilon in MATLAB. Infinitesimal perturbation ϵ_i is mainly for the expanding map f_{exp} to avoid artificial convergence of trajectories due to the digital representation of an initial value, which can also be seen in numerical simulations of the Bernoulli shift map [7]. For consistency, we implemented perturbation ϵ_i to f_{rot}, as well as f_{exp}. We embedded S^1 to Euclidean space by the transformation $x \mapsto [\cos x, \sin x]$, and used $\{[\cos x_i, \sin x_i]\}_{i=0}^{\infty}$ as $\mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}$ for homology inference. For f_{rot} and f_{exp}, we got initial state x_0 by the uniform distribution on S^1, which is equivalent to chose a point randomly from $\{\mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}\}$.

In addition to f_{rot}, f_{exp}, we conducted hypothesis testing with the null hypothesis H_0 at $\delta \in \{0, \texttt{logspace}(10^{-3}, \mathrm{diam}(X), 10)\}$ for flows of the Lorenz equation and the Rössler equation defined by:

$$\text{Lorenz:} \begin{cases} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z, \end{cases} \quad (3)$$

$$\text{Rössler:} \begin{cases} \dot{x} &= -y - z, \\ \dot{y} &= x + ay, \\ \dot{z} &= b + xz - cz, \end{cases} \quad (4)$$

We used fixed parameters $(\sigma, \rho, \beta) = (10, 28, 8/3)$ and $(a, b, c) = (0.36, 0.4, 4.5)$ for all simulations. In this study, we only cared about 1-dimensional holes (cycles) in an attractor of systems (3, 4) to reduce computational costs. In this case, it is sufficient to consider 2-dimensional state space. For this purpose, we took the 2-dimensional time-delay coordinate of the first variable $[x(t), x(t-T)]$ of each system. We fixed time delay T at different values for each system, respectively. Then we defined a map of dynamical system by $f : [x(t_i), x(t_i - T)] \mapsto [x(t_{i+1}), x(t_{i+1} - T)]$, and used temporary discretized orbits $\{[x(t_i), x(t_i - T)]\}_{i=0}^{\infty}$ as $\mathcal{S}^{(N)}_{f-\mathrm{orbit}}$ for homology inference. We got an initial state $[x(t_0), x(t_0 - T)]$ by a random choice from $\mathcal{S}^{(50000)}_{f-\mathrm{orbit}}$, which is numerical approximation of $\mathcal{S}^{(\infty)}_{f-\mathrm{orbit}}$. We used $\mathrm{Hom}[B_R(\{\mathcal{S}^{(N)}_{f-\mathrm{orbit}}\})]$ with sufficiently large $N = N_{\max}$ as an approximation of true homology (Figure 3).

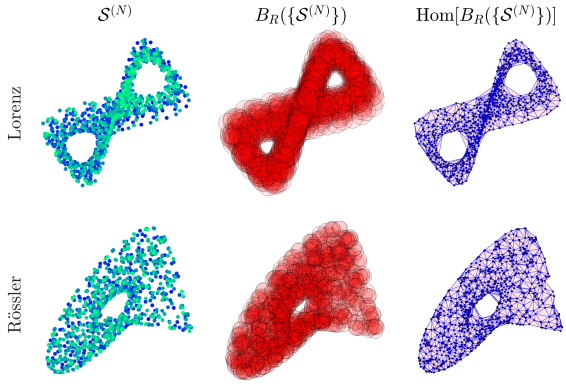

Figure 3: Homology of attractors in the Lorenz equation and the Rössler equation on time-delay coordinate. The right columns show connections between points (blue lines). red-colored sheds indicate there is no blank.

Although temporal discretization with different time width $\Delta t = t_{i+1} - t_i$ result in quantitative difference of $\theta_{f-\delta}(N, R)$, we emphasize that, as long as $\mathcal{S}^{(N)}_{f-\delta}$ and $\mathcal{S}^{(N)}_{f-\mathrm{random}}$ are sharing the same time scale, test result for H_0 is preserved. Thus, as well as delay T, we used different time widths Δt for different dynamical systems as needed.

We summarized the parameters settings for each dynamical system related to homology inference on Table 1.

Table 1: Parameters related to homology inference

	$N_{\max}$	$R_{\max}$	T [s]	Δt [s]
$f_{\mathrm{rot}}, f_{\mathrm{exp}}$	100	1.5	-	-
Lorenz	1000	5	0.15	0.05
Rössler	1000	2	0.7	0.5

3. Results and Discussion

3.1. H_0 was rejected for f_{rot}, but not for f_{exp}

Figure 4 shows the results of hypothesis testings with $H_0^{(i,j)}$ $(1 \le i, j \le 10)$ as null hypotheses for f_{rot} and f_{exp}. The quantitative property of $\theta_*(N, R)$, colored by green to yellow, was intuitively reasonable. Namely, we got low values of inference accuracy $\theta_*(N, R)$ at extremely large value of $R_{\max}$ since, with such value of $R_{\max}$, any pairs of points in $\mathcal{S}^{(N)}$ were connected, and thus there was no chance to detected a cycle in a state space. Also, we got low values of inference accuracy $\theta_*(N, R)$ by insufficiency of the number of samplings N with respect to R, which resulted in the detections of separated components rather than a cycle.

Importantly, we found significant differences between $\theta_{f_{\mathrm{rot}}-\mathrm{orbit}}(N, R)$ and $\theta_{f_{\mathrm{rot}}-\mathrm{random}}(N, R)$ for some paris of (N, R), but not so between $\theta_{f_{\mathrm{exp}}-\mathrm{orbit}}(N, R)$ and $\theta_{f_{\mathrm{exp}}-\mathrm{orbit}}(N, R)$. Thus, we concluded that H_0 was rejected for f_{rot}, but not for f_{exp}. Considering the observation that learning of f_{exp} is harder compared with so of f_{exp} (Figure 6 in Supplementary Material), this results suggests that our conjecture is valid for both f_{rot} and f_{exp}.

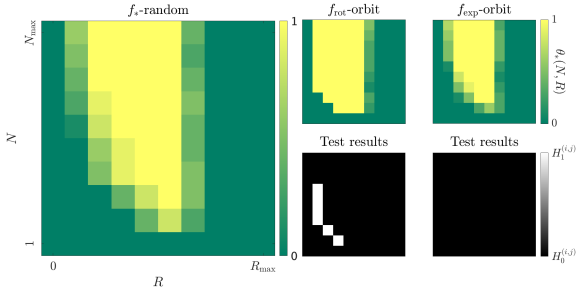

Figure 4: $\theta_*(N,R)$ and testing results for $f_{\rm rot}$ and $f_{\rm exp}$

3.2. H_0 was rejected for the Lorenz equation, but not for the Rössler equation

Figure 5 shows the results of hypothesis testings for the Lorenz equation and the Rössler equation, in the same manner as Figure 4. Similar to the indications of Figure 4, we concluded that H_0 was rejected for the Lorenz equation, but not for Rössler equation. Again, results in Figure 5 supported our conjecture for both $f_{\rm rot}$ and $f_{\rm exp}$.

Moreover, we found that the counts of $H_0^{(i,j)}$ rejections, this presumably corresponds the level of H_0 rejection, increase with the increasing δ from 0 to 0.003 (Figure 5). This was surprising because since tests evaluate the significant difference from the case of $\delta = \infty$, and hence the difference was expected to decrease monotonically with the increasing of δ. Therefore, it is interesting to investigate whether the specific value of δ, which determine the amplitude of sampling noise, has any means for the learning of the Lorenz equation.

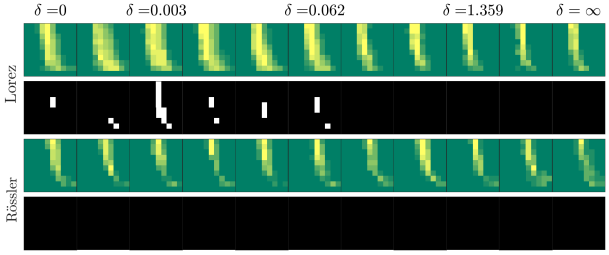

Figure 5: $\theta_*(N,R)$ and testing results for the Lorenz equation and the Rössler equation at different δ

4. Conclusion

In this study, we suggested a conjecture for homology inference properties of samplings from state space of a dynamical system, and investigated its validity. Simulation results suggested that our conjecture holds for typical dynamical systems. For future work, considering to expand the scope of our conjecture and provide concrete interpretations related to dynamics learning is important.

Acknowledgments

We thank Prof. Yuichi Katori (Future University Hakodate, Hakodate, Japan) for discussion. The research was supported by JSPS KAKENHI Grants No. JP20H00596, JP21K12105, JP22K18419, JST CREST Grant No. JPMJCR19K2 and JST Moonshot R&D Grant No. JPMJMS2021 for K.F.

5. Supplementary Material

Here, we illustrated typical examples of dynamics learning by using a reservoir computing (RC) defined by $r_{t+1} = g(x_t, r_t) = \overline{\tanh}\,[Ar_t + Bx_t]$, where $x_t \in \mathbb{R}^m$, $r_t \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $\overline{\tanh} : [r_t^{(1)} \cdots r_t^{(n)}]^\top \mapsto [\tanh(r_t^{(1)}) \cdots \tanh(r_t^{(n)})]^\top$. x_t is a input signal from a dynamical system f to be learned. In this case, obtaining $W \in \mathbb{R}^{m\times n}$ which minimize $|Wr_t - x_t|$ $(t \geq 0)$ completes a dynamics learning, and the learned dynamics is $g_W : r_t \mapsto g(Wr_t, r_t)$. Figure 6 shows an orbit of g_W and its targets, for each dynamical system with the same parameters settings of RC. Figure 6 indicates that dynamcics learning of the expanding map and the Rössler equation is harder than that of the rotaion map and the Lorenz equation.

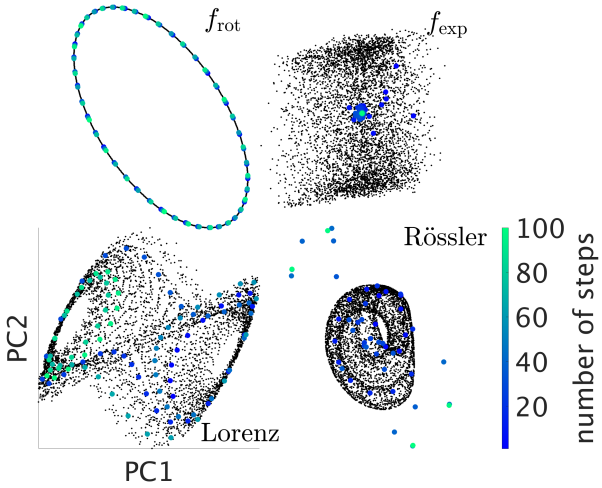

Figure 6: Success and failure of dynamics learning. Plots contain targets (black) and predictions (blue and green).

References

[1] Zhixin Lu et al. *Chaos*, 27(4):041102, 2017.

[2] Ling-Wei Kong et al. *Chaos*, 33(3), 2023.

[3] Allen Hart et al. *Neural Netw.*, 128:234–247, 2020.

[4] Eugene Tan et al. *Chaos*, 31(12):123109, 2021.

[5] Herbert Edelsbrunner and John L Harer. American Mathematical Society, 2022.

[6] Herbert Edelsbrunner et al. *Discrete Comput. Geom.*, 28:511–533, 2002.

[7] Shujun Li et al. *Int. J. Bifurc. Chaos*, 15(10):3119–3151, 2005.