卒論　ノート

氏名：久野証

所属：東大工学部計数工学科数理情報工学コース

学籍番号: 03-210599

2023 年 9 月 25 日

目次

概要

2023 年 A セメスターの卒論執筆に際して、勉強したことや考えたことのメモをここにまとめた。

1 Papers

1.1 Emergence of a resonance in machine learning

Zheng-Meng Zhai , 1 Ling-Wei Kong , 1 and Ying-Cheng Lai 1,2,* 1School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona 85287, USA 2Department of Physics, Arizona State University, Tempe, Arizona 85287, USA.

(Received 9 June 2022; revised 1 March 2023; accepted 26 July 2023; published 24 August 2023)

キーワード：Resonance in nonlinear dynamical systems,

1.1.1 要旨

1. 入力信号にノイズを挿れた場合の Reservoir Computing を考える。
   (a) hyperparameters が最適化されていない時でも、ノイズを挿れることで予測の精度をあげることができる。
   (b) もっとも良い精度を達成するには、hyperparameters が最適化されていなければならない。
      i. Hyperparameters に対する Baysian optimization で最適化可能。
      ii. 確率共鳴があると決定づけるために、ノイズの振幅を hypermparameter に数える。
2. Macky-Glass (MG) system と Kuramoto-Sivashinsky (KS) system に対して、シミュレーションを行う。
3. 物理側から、確率共鳴が生まれる原理を考える。

1.1.2 状況設定

Appendix A 参照。

1. hyperparameters:
   (a) ρ: the specral radius of the reservoir network.
   (b) γ: the scaling factor of the input weights.
   (c) α: the leakage parameter
   (d) β: the regularization coefficient
   (e) p: the link connection probability of the random network in the hidden layer.
   (f) σ: the noise amplitude
2. hyperparameters の最適化
   (a) MATLAB: SURROGATEOPT を用いる。[*1]
   (b) σ ごとに hyperparameters の最適化を行うので、σ に対する他の hyperparameters の組みは異なる。
3. シミュレーションを行う。
   (a) MG system:

$$\dot{s}(t) = as(t-\tau)/\left(1+[s(t-\tau)]^c\right) - bs(t),$$

   τ is the time delay, a, b, and c are parameters. [*2]

   i. $a = 0.2, b = 0.1,$ and $c = 10$ を固定。
   ii. $\tau = 17$, Lyapunov exponents: $\lambda_+ \approx 0.006$ と $\tau = 30$, Lyapunov exponents: $\lambda_+ \approx 0.011$ and 0.003 の 2 つの場合を比べる.
   iii. $\Delta t = 100h = 1.0$.
   iv. Warmup: $10000\Delta t$.
   v. 時系列データには事前に z-score normalization: $z(t) = [s(t) - \bar{s}]/\sigma_s$ を施す。

[*1] "The Bayesian optimization method can be implemented using PYTHON or other languages. Different packages for Bayesian optimization are now available, such as BAYESIAN-OPTIMIZATION and BOTORCH in PYTHON."としている。

[*2] The state of the system at time t is determined by the entire prior state history within the time delay, making the phase space of the system infinitely dimensional.

1.1.3 面白いと思ったところ

1. Introduction でも述べられているが、ノイズを入れることによって、初期値鋭敏性を持つカオスシステムに対して。短期的にも長期的にも予測の精度を挙げられるということ。

1.1.4 論文を受けての今後の研究方向

1.1.5 疑問点

1. Fig.4 と Fig.5 について，MG system における τ の値が 30 から 17 に変えると，σ にどのような影響があるか．なぜその影響が生まれるか.
2. なぜ，Fig.2 の（逆）ピークを与える σ 帯と Fig.6(c) の（逆）ピークを与える σ 帯が重なるのか.
3. III で，Machine learning における resonance が生まれる Physical reason を挙げているが，これは対象を正しく説明できているか．extraordinarily complicated な hidden layer の中身を解析することなく，physical reason を与えることが，なにを説明しているのか/なにを説明していないのか.

1.1.6 関連する文献

1.1.7 用語まとめ

1.1.8 Abstract

1. stochastic/coherence resonance:
2. nonlinear dynamical system:
3. regularizer/regularization:
4. reservoir computing:
5. state variables/attractor:
6. hyperparameters:

1.1.9 I. Introduction

1. model-free/data-driven:
2. oscillatoin/Lyapnov times:
3. trajectory:
4. basin boundary:
5. robustness:
6. Baysian optimization:

1.1.10 II. Result

1. SURROGATEOPT function (MATLAB):
2. surrogate approximation function:
3. objective function:
4. global minimum:
5. sampling/updating:
6. radial basis function:
7. Mackey-Glass (MG) system:
8. spatiotemporal chaotic Kuramoto-Sivashinsky (KS) system:

1.1.11 A. Emergence of a resonance from short-term prediction

1. transient behavior:
2. z-score normalization:
3. periodic boundary condition:
4. Prediction horizon/stability:

1.1.12 B. Emergence of a resonance from long-term prediction

1. collapse:
2. wider/narrower resonance:

1.1.13 III. HEURISTIC REASON FOR THE OCCURRENCE OF A RESONANCE

1. time-scale match:
2. the mean first-passage time:
3. nonlinear activation:
4. linear reservoir computing:
5. noise-enhanced temporal regularity:
6. vector autoregressive process (VAR):

1.1.14 IV. DISCUSSION

1. magnitude:

1.1.15 Appendix A

1. recurrent neural network(RNN):
2. input/hidden/output layer:
3. linear regression:
4. adjacency matrix:
5. state vector:

6. dynamical state/evolution:
7. neuron:
8. leakage parameter α:
9. link probability p:
10. spectral radius:

2 Lectures

1. Reservoir Computing
   (a) Reservoir Computing for SDEs(Josef Teichmann:)
   (b) Reservoir Computing & Dynamical Systems - Second Sumposium on Machine Learning and Dynamical Systems(Josef Teichmann)
   (c) Introduction to Next Generation Reservoir Computing(Daniel Gauthier)
2. Machine Learning in general
   (a) Machine Learning in Finance(Josef Teichmann)

2.1 Josef Teichmann: Reservoir Computing for SDEs

Access from here.

1. We consider differential equations of the form

$$dY_t = \sum_i V_i\left(Y_t\right) du_t^i, Y_0 = y \in E$$

   to construction evolutions in state space E (could be a manifold of finite or infinite dimension) depending on local characteristics, initial value $y \in E$ and the control u.

2. Theorem (Universality) Let Evol be a smooth evolution operator on a convenient vector space E which satisfies (again the time derivative is taken with respect to the forward variable t) a controlled ordinary differential equation

$$d\mathrm{Evol}_{s,t}(x) = \sum_{i=1}^{d} V_i\left(\mathrm{Evol}_{s,t}(x)\right) du^i(t).$$

   Then for any smooth (test) function $f : E \to \mathbb{R}$ and for every $M \geq 0$ there is a time-homogenous linear $W = W\left(V_1, \ldots, V_d, f, M, x\right)$ from $\mathbb{A}_d^M$ to the real numbers $\mathbb{R}$ such that

$$f\left(\mathrm{Evol}_{s,t}(x)\right) = W\left(\pi_M\left(\mathrm{Sig}_{s,t}(1)\right)\right) + \mathcal{O}\left((t-s)^{M+1}\right)$$

   for $s \leq t$

3. Signature as universal dynamical system

(a) This explains that any solution can be represented - up to a linear readout - by a universal reservoir, namely signature. Similar constructions can be done in regularity structures, too (branched rough paths, etc).

(b) This is used in many instances of provable machine learning by, e.g., groups in Oxford (Harald Oberhauser, Terry Lyons, etc), and also ...

(c) ... at JP Morgan, in particular great recent work on 'Nonparametric pricing and hedging of exotic derivatives' by Terry Lyons, Sina Nejad and Imanol Perez Arribas.

(d) in contrast to reservoir computing: signature is high dimensional (i.e. infinite dimensional) and a precisely defined, non-random object.

(e) Can we approximate signature by a lower dimensional random object with similar properties?

3 TODOs

1. reservoirpy 関連

   (a) Understand and optimize ESN hyperparameters などの Tutorial ページを読む.

2. Reservoir Computing について学ぶ.

   (a) レクチャーノートなどを通じて、知識を準備する。

   (b) Github 環境を整備する。