Sensitivity Analysis of Chaotic System

Jiaxi Zhao

9th October, 2023

Content

- Chaotic Dynamics and Sensitivity Analysis
- Stability Analysis(TO DO)
- Shadowing Lemma
- Least Square Shadowing (LSS)
- Numerical Experiments

Chaos

Chaos are ubiquitous in our life:

- Climate and weather forecasting
- Traffic flow
- Combustion and mixing
- Financial market
- Biological system

Lots of modern engineering problems are about modeling, quantifying, and controling chaos.

Sensitivity analysis of Chaos

Besides simulating the chaos, one of the most important task is sensitivity analysis of the chaos. This is especially important for controlling the chaos.

- Weather control: disaster mitigation, precipitation
- Fluid control: efficient mixing, combustion control

Moreover, sensitivity analysis is a key step to quantify and understand chaos.

Lorenz dynamics

Consider the Lorenz system

$$\begin{aligned}\dot{x} &= \sigma(y-x),\\ \dot{y} &= \rho x - y - xz,\\ \dot{z} &= xy - \beta z,\end{aligned} \tag{1}$$

where σ is the Prandtl number, ρ Rayleigh number. Simple properties of the Lorenz Equations[1]:

- **Nonlinearity**: coming from the nonlinear convection term.
- **Symmetry**: Equations are invariant under $(x, y) \to (-x, -y)$. Moreover, the z-axis is invariant.
- **Fixed point**: $(0, 0, 0)$: stable fixed point for $0 < \rho < 1$, unstable otherwise; $(\sqrt{\beta(\rho-1)}, \sqrt{\beta(\rho-1)}, \rho-1), (-\sqrt{\beta(\rho-1)}, -\sqrt{\beta(\rho-1)}, \rho-1)$ stable fixed points for $1 < \rho < \frac{\sigma(\sigma+\beta+3)}{\sigma-\beta-1} \approx 24.74$.

[1] C. Sparrow, The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors, Springer-Verlag, New York, 1982.

Lorenz dynamics

- **Volume contraction**: The Lorenz system is dissipative

$$\frac{d}{dt} d\,\mathrm{vol} = \mathrm{Tr}(\nabla f) d\,\mathrm{vol} = -(\sigma + \beta + 1) d\,\mathrm{vol}\,. \tag{2}$$

- **Bounded attractor**: Consider the Lyapunov function

$$V(x, y, z) = \rho x^2 + \sigma y^2 + \sigma(z - 2\rho)^2. \tag{3}$$

  Together with the previous property guarantee the existence of a bounded globally attracting set of measure 0. This is the emergence of the strange attractor with fractal dimension.

Lorenz dynamics visualization

Lorenz dynamics has one of the key feature of the chaos:

Sensitive dependence on initial conditions: Given two trajectories, no matter how near they are at the initial condition, they will diverge exponentially fast until their distance reaches the size of the attractor basin. The deviation speed can be quantified by the first Lyapunov exponent.

Definition (Formal definition)

Chaos is aperiodic long-term behaviour in a deterministic system that exhibits sensitive dependence on initial conditions.

Quantifying chaotic dynamics: statistics

For chaotic dynamic, studying the trajectory-wise information is not a good choice due to the sensitive dependence on the inital condition. An alternative choice is to use statistical quantities to describe the chaos, usually calculated along the trajectories

$$\langle g \rangle := \frac{1}{T} \sum_{k=1}^{T} g(\phi^k(\mathbf{u})). \tag{4}$$

Advantanges of statistics

- 1. It is not sensitive to the initial condition via the Birkhoff's ergodic theorem: Let ϕ be an ergodic endomorphism of the probability space X with invariant measure ρ and let $g : X \to \mathbb{R}$ be a real-valued measurable function. Then for almost $\forall \mathbf{u} \in X$, we have

$$\frac{1}{T}\sum_{k=1}^{T} g(\phi^k(\mathbf{u})) = \int_X g(\mathbf{u}) d\rho(\mathbf{u}). \tag{5}$$

- 2. Statistics has important physical and engineering interpretation.
    - Prediction weather $\Longleftrightarrow$ prediction climate
    - Prediction pointwise fluid field $\Longleftrightarrow$ prediction the turbulence energy spectrum
    - Prediction the pressure on the each area of a plane $\Longleftrightarrow$ predicting the average lifting of the plane.

Next step: sensitivity analysis

Besides the value of the statistics $\langle g \rangle$, we are also interested in its dependence over some parameters in order to understand the system better and control the system.

Using Lorenz system as an example, we are interested in the following statistics

$$J(\rho) := \lim_{T \to \infty} \frac{1}{T} \int_0^T z_\rho(t) dt. \tag{6}$$

Notice we involve ρ Rayleigh number as a parameter of this statistics as we want to study the influence of ρ on J. The other quantity of interest is the derivative $\frac{dJ}{d\rho}$ of J over ρ.

Sensitivity analysis: original method

Original method is given by linearizing the dynamics to obtain:

$$\frac{d}{dt}v(t) = \begin{pmatrix} -\sigma & \sigma & 0 \\ \rho - z(t) & -1 & -x(t) \\ y(t) & x(t) & -\beta \end{pmatrix} v(t) + \begin{pmatrix} 0 \\ x(t) \\ 0 \end{pmatrix}, \quad (7)$$

This is method does not work for chaotic dynamics since the derivative operator (over ρ) and integration operator (over t) do not commute!!!

$$\begin{aligned} &\frac{d}{d\rho} \lim_{T\to\infty} \frac{1}{T} \int_0^T z_\rho(t)dt \\ \neq\ & \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{d}{d\rho} z_\rho(t)dt = \lim_{T\to\infty} \frac{1}{T} \int_0^T v_z(t)dt. \end{aligned} \quad (8)$$

Pseudo-orbit

Given a map $f : X \to X$ of a metric space (X, d) to itself, define a ϵ-pseudo-orbit as a sequence of $(\mathbf{u}_n)$ of points such that $\mathbf{u}_{n+1}$ belongs to a ϵ-neighborhood of $f(\mathbf{u}_n)$.

Example (Parameterized family of dynamics)

Consider a family of dynamical systems parameterized by s:

$$\mathbf{u}_{n+1} - f(\mathbf{u}_n, s). \tag{9}$$

Given that f is L-Lipschitz continuous w.r.t. s, any orbit of the dynamical system with parameter $s + \delta s$ is a $L\delta s$-pseudo-orbit of the dynamical system with parameter s.

Pesudo-orbit

Example (Round-off error in numerical simulation)

When numerically solving a dynamical system, due to the machine round-off error, in each step we are not performing the exact calculation but rather an ϵ approximation calculation assuming the state variable is bounded

$$\|\mathbf{u}_{n+1} - f(\mathbf{u}_n)\| < \epsilon f(\mathbf{u}_n). \tag{10}$$

Shadowing

Theorem (Shadowing lemma)

Let Γ be a hyperbolic invariant set of a diffeomorphism f. There exists a neighborhood U of Γ with the following property: for any $\delta > 0$ there exists $\epsilon > 0$, such that any (finite or infinite) ϵ-pseudo-orbit that stays in U also stays in a δ-neighborhood of some true orbit[2].

In a hyperbolic invariant set, the dynamics exhibit a combination of stable and unstable behavior.

Shadowing trajectory demo:

[2] Pilyugin SY. Shadowing in dynamical systems. Lecture notes in mathematics, vol. 1706. Springer; 1999.

Digress: Stability in numerical linear algebra

- A method for computing some result w is forward stable if the computed solution $\widehat{w}$ is "near" the exact solution:

$$\frac{\|\widehat{w} - w\|_2}{\|w\|_2} < \epsilon. \tag{11}$$

- A method is backward stable if $\widehat{w}$ is the exact solution of a perturbed problem; that is, if p is the problem satisfied by w and $\widehat{p}$ is the problem satisfied by $\widehat{w}$, then

$$\frac{\|\widehat{p} - p\|_2}{\|p\|_2} < \epsilon. \tag{12}$$

Lesat square shadowing: Finding shadowing direction

Suppose we have a $s + \delta s$-trajectory:

$$\widehat{u}^{t+1} = f(\widehat{u}^t, s + \delta s), \quad u^t \in \mathbb{R}^n. \tag{13}$$

Assume f is uniform Lipschitz w.r.t. s, we know this is a $L\delta s$-pseudo orbit for the dynamics with parameter s. According to shadowing lemma, we could find u^t, s.t.

$$u^{t+1} = f(u^t, s), \quad u^t \in \mathbb{R}^n, \quad \left\| \widehat{u}^t - u^t \right\|_2 < \epsilon, \ \forall t. \tag{14}$$

Lesat square shadowing: vector version

To solve this, we formally write $\widehat{u}^t = u^t + v^t \delta s$ and solve v via optimization

$$\min \sum_{t=1}^{T} \frac{\|\widehat{u}^t - u^t\|_2^2}{\delta s^2} = \sum_{t=1}^{T} v_t^T v_t$$
$$s.t.\ v^{t+1} = \nabla_u f(u^t, s) v^t + \partial_s f(u^t, s),\ \ t \in [T], \tag{15}$$

where the constraint is nothing but the linearized dynamics. One can view this either as a optimization over all the tangent vectors v_t or simply over the initial condtion v_1.

Solve LSS: Least square

Writing the linearized equation as a linear constraint

$$\min \sum_{t=1}^{T} v_t^T v_t$$

$$\begin{pmatrix} \mathbf{I} & -\nabla_u f(u_{T-1}) & \cdots & 0 & 0 \\ 0 & \mathbf{I} & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{I} & -\nabla_u f(u_1) \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \begin{pmatrix} v_T \\ v_{T-1} \\ v_{T-2} \\ \vdots \\ v_2 \\ v_1 \end{pmatrix} = \begin{pmatrix} \partial_s f(u_{T-1}) \\ \partial_s f(u_{T-2}) \\ \partial_s f(u_{T-3}) \\ \vdots \\ \partial_s f(u_1) \\ 0 \end{pmatrix}, \tag{16}$$

This is just a least square problem of size $T \times n$.

Solve LSS: PMP

Even for moderate size system, e.g. simulating the 2D isotropic turbulence using Navier-Stokes equation with grid size 32×32 for 1000 time step would result in a linear system of size 1e6 for each gradient evaluation, which is extremely expensive.

Another method is viewing this as a control problem and solve it using PMP. In each gradient descent step, it requires to solve a forward primal problem followed by a backward adjoint problem. It become expensive when forward problem solving is costable, e.g. LES.

Sensitivity analysis of the Lorenz system

Going back to the Lorenz system, we consider the following statistics:

$$J(\rho) := \lim_{T\to\infty} \frac{1}{T} \int_0^T z_\rho(t)dt. \tag{17}$$

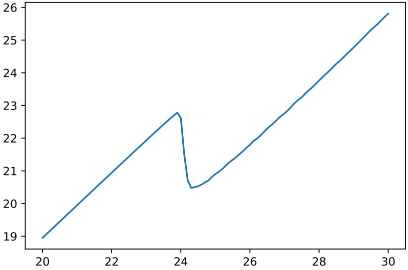
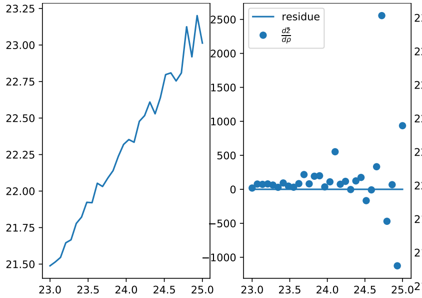
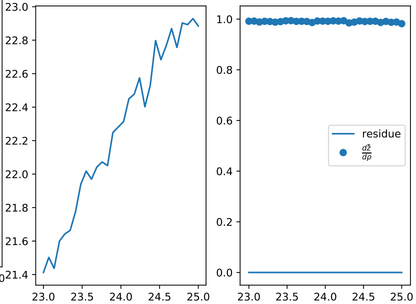


Figure: Dependence of the statistics $J(\rho)$ over parameter ρ. calculate using *solve_ivp* scipy routine with $T = 100$ time unit and 100 random sample trajectories.

Sensitivity analysis of the Lorenz system



(a) Gradient calcuated using naive method



(b) Gradient calcuated using LSS

Figure: Comparison of naive and LSS methods for sensitivity analysis

Subgrid modeling of the Kuramoto–Sivashinsky equation

Subgrid modeling is very important in the Large-eddy simulation. Here, we consider a simple version of subgrid modeling for 1D KS equation:

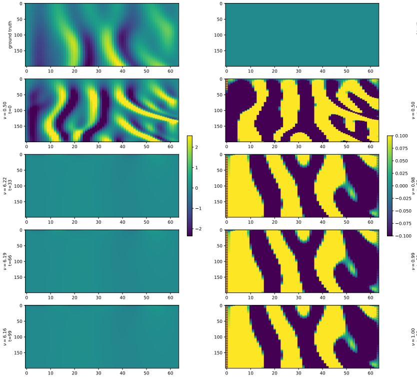
$$\partial_t u + u u_x + u_{xx} + \nu u_{xxxx} = 0, \quad x \in [0, L], \tag{18}$$

If we consider a coarse-grid model $\bar{u} = u * G$,

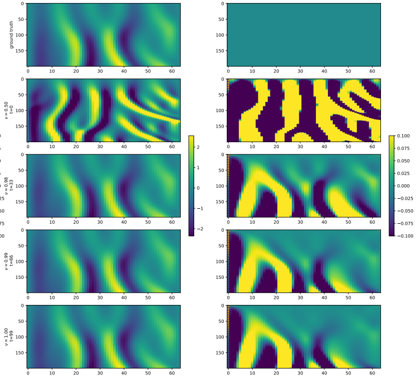
$$\partial_t \bar{u} + \bar{u}\bar{u}_x + \bar{u}_{xx} + (\nu + \Delta\nu)\bar{u}_{xxxx} = 0, \quad x \in [0, L], \tag{19}$$

Due to the nonlinearity, we need to add $\Delta\nu$.

Subgrid modeling of the Kuramoto–Sivashinsky equation



(a) Adjoint method



(b) LSS method

Figure: LSS for KS subgrid modeling

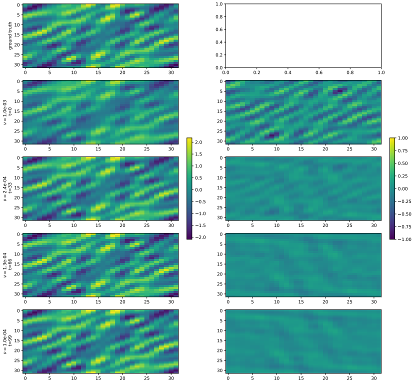
Subgrid modeling of the 2D Navier-Stokes equation

Here, we consider a simple version of subgrid modeling for 2D NS equation:

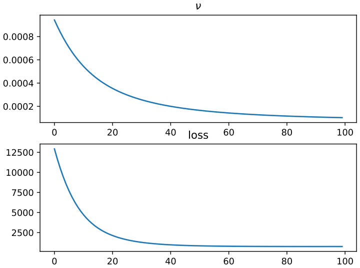
$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} + \nu \Delta \mathbf{u} = 0, \quad \nabla \cdot \mathbf{u} = 0, \tag{20}$$

This corresponds to estimate the eddy-viscosity ν_t of LES subgrid modeling.

Subgrid modeling of the Kuramoto–Sivashinsky equation



(a) Simulated configuration



(b) Error

Figure: LSS for NS subgrid modeling

What is next?

- Applying LSS (possibly via the help of neural network) to 2D isotropic turbulence to study the subgrid modeling and sensitivity analysis of several model parameters. Moreover, implement scalable method to work on the case of real world LES problem, e.g. 3D, more than 10M grid size, and a reasonable long trajectory.
- Applying LSS to more complicated statistics which may also depends on the initial contion (current method assume that the statistics is insensitive w.r.t. the inition condition).
- Viewing deep NN as a dynamics where layer depth is the time index, can we use LSS to propose an algorithms to help calculating the gradient? Moreover, is there any other issue except for gradient vanishing and blow-up for deep NN?

Reference

- 1. C. Sparrow, The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors, Springer-Verlag, New York, 1982.
- 2. Wang Q, Hui R, Blonigan P. Least squares shadowing sensitivity analysis of chaotic limit cycle oscillations. J Comput Phys 2014;267: 210–4.