Weak Collocation Regression For Parameter Identification Chaotic Dynamics Via Invariant Mesure

BY ZHIJUN ZENG

Date:03/22

1 Dynamic System Setting

Noisy observations are denoted as

$$\mathbf{X}^* = (\mathbf{x}^*(t_0) + \eta_0, \mathbf{x}^*(t_1) + \eta_1, \cdots, \mathbf{x}^*(t_n) + \eta_n) \tag{1}$$

where $x^*$is the solution of the autonomous dynamical system of $\dot{x} = v(x, \theta^*)$, and $\{\eta_0, \eta_1, \ldots, \eta_n\}$are the measurement errors or uncertainties.

We suppress the time variable and consider the state-space distribution of data

$$\rho^* = \frac{\sum_{i=0}^{n} \delta_{x^*(t_i)}}{n+1} \tag{2}$$

The dynamic system admits a physical measure $\rho(\theta)$ if for a Lebesgue positive set of initial condition $x(0) = x$, one has that

$$\rho(\theta) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \delta_{x(t)} \mathrm{dt} \tag{3}$$

Mathematically, statistical properties of can be characterized by the occupation measure $\rho_{x,T}$ defined as

$$\rho_{x,T}(B) = \frac{1}{T} \int_0^T \mathbb{1}_B(\mathbf{x}(s))\, d\, s = \frac{\int_0^T \mathbb{1}_B(\mathbf{x}(s))\, d\, s}{\int_0^T \mathbb{1}_{\mathbb{R}^d}(\mathbf{x}(s))\, d\, s} \tag{4}$$

where B is any Borel Measurable set.If there exist an invariant measure ρ such that aull $\rho_{x,T}$ weakly converges to ρ for all initial condition, then ρ is an physical measure .

By definition of physical measures μ^*, we have

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(x(t))\, d\, t = \int_{\mathbb{R}^d} f(x)\, d\, \mu^*(x), \quad f \in C_c^\infty(\mathbb{R}^d) \tag{5}$$

By taking $f(x) = \nabla \phi(x) \cdot v(x)$ for some $\phi \in C_c^\infty(\mathbb{R}^d)$

$$\begin{aligned} \int_{\mathbb{R}^d} \nabla\, \phi(x) \cdot v(x)\, d\, \mu^*(x) &= \lim_{T \to \infty} \frac{1}{T} \int_0^T \nabla\, \phi(x(t)) \cdot v(x(t))\, d\, t = \lim_{T \to \infty} \frac{1}{T} \int_0^T \nabla\, \phi(x(t)) \cdot \dot{x}(t)\, d\, t \\ &= \lim_{T \to \infty} \frac{1}{T} (\phi(x(T)) - \phi(x(0))) = 0 \end{aligned} \tag{6}$$

This shows that μ^* is the stationary distributional solution

$$\nabla \cdot (v(x, \theta) \rho(x, \theta)) = 0. \tag{7}$$

2 Weak Collocation Regression Method

Using Galerkin method, one can derivate the weak form of (7)

$$\int_{\Omega} v(x,\theta)\cdot\nabla\phi(x)\,\rho(x,\theta)\mathrm{dx} = \mathbb{E}_{\rho(x,\theta)}(v(x,\theta)\cdot\nabla\phi(x)) = 0, \forall\phi\in C^{\infty}(\Omega)$$

Taking $\phi(x) = \{\rho_{\mu,\Sigma}(x)$: the density function of $\{N(\mu_i,\Sigma_i)\}_{i=1}^{N}\}$, we can construct N nonlinear(linear in Linear ODE systems) equations

$$F_j(\theta) = 0, j = 1, 2, \ldots N$$

The equations are solved by Newton method or Gradient descent method. In the second case, the loss function is chosen as $\|\boldsymbol{F}(\theta)\|_2^2$.

Taking Lorentz equation as an example

$$\left\{\begin{array}{ccc}\dot{x} & = & \sigma(y-x)\\ \dot{y} & = & x(\rho-z)-y\\ \dot{z} & = & \mathrm{xy}-\beta z\end{array}\right., v(x,y,z) = \left[\begin{array}{c}\sigma(y-x)\\ x(\rho-z)-y\\ \mathrm{xy}-\beta z\end{array}\right]$$

the equation $\mathbb{E}_{\rho(x,\theta)}(v(x,\theta)\cdot\nabla\phi(x)) = 0$ result in

$$\sigma\left(\int_{\Omega} y\phi_{x_1} - x\phi_{x_1}\right) + \rho\int_{\Omega} x\phi_{x_2} - \beta\int_{\Omega} z\phi_{x_3} = \int_{\Omega}(\mathrm{xz}+y)\phi_{x_2} - \int_{\Omega}\mathrm{xy}\phi_{x_3}$$

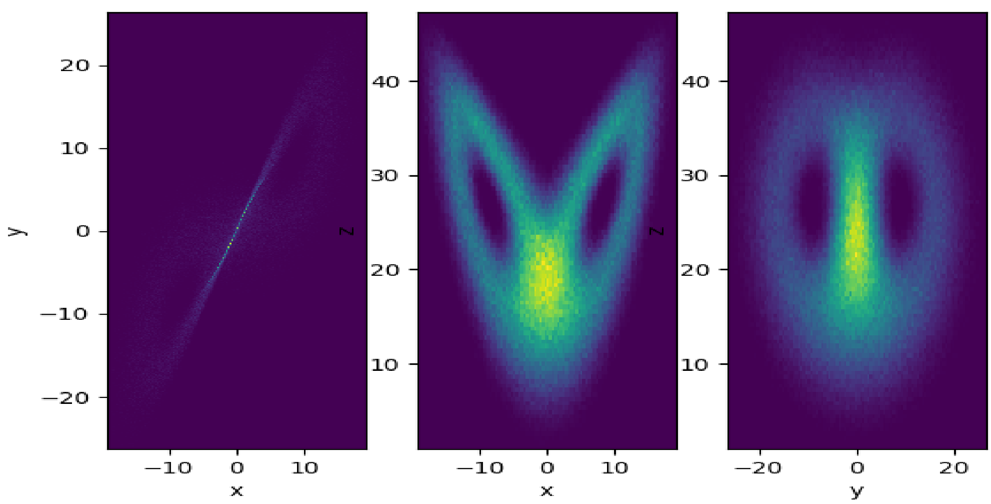
which is indeed a linear equation for each test function ϕ.

3 Examples

3.1 Lorentz System

$$\left\{\begin{array}{ccc}\dot{x} & = & \sigma(y-x)\\ \dot{y} & = & x(\rho-z)-y\\ \dot{z} & = & \mathrm{xy}-\beta z\end{array}\right.,$$

Here $\sigma = 10, \rho = 28, \sigma = 8/3$, initial condition is [1,1,5],$T = 20000$, $\mathrm{dt} = 0.01$.the projection of density map wrt xy,xz,yz direction is below.



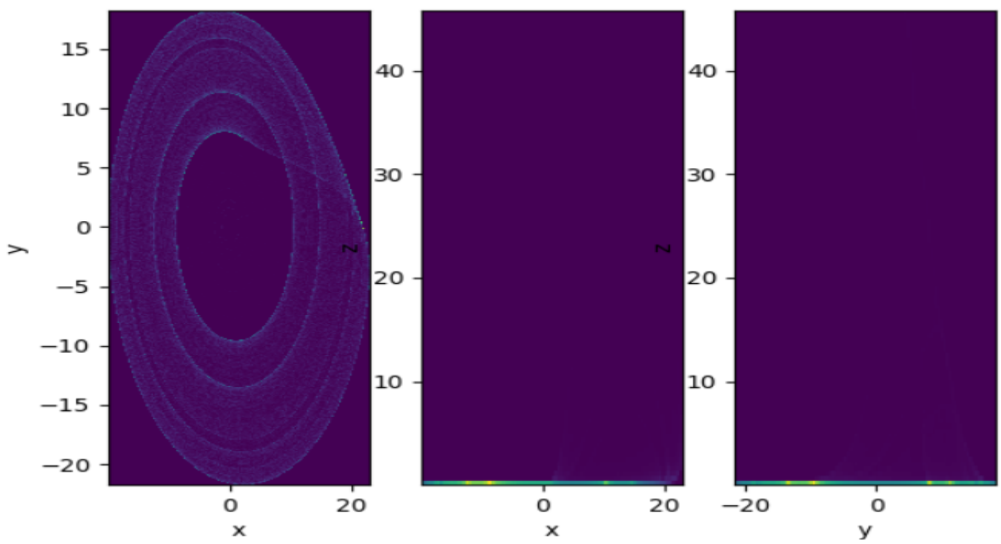
To reconstruct the parameter, we choose 1000 gaussion whose mean is generated by LHS sampling in the domain of data.Using Gradient method with lr = 0.01 and 30000 epoch and Least square method , the reconstruction result is

	σ	ρ	β
GD	10.0567	28.0346	2.71
Error	0.0567	0.0346	0.0453
LS	10.0573	28.0351	2.7115
Error	0.0573	0.0351	0.0448

3.2 Rossler System

$$\begin{cases} \dot{x} &= & -y-z \\ \dot{y} &= & x+\mathrm{ay} \\ \dot{z} &= & b+z(x-c) \end{cases},$$

Here $a=0.1, b=0.1, c=14$. initial condition is [2,2,5],$T=20000$, dt $=0.01$..the projection of density map wrt xy,xz,yz direction is below.



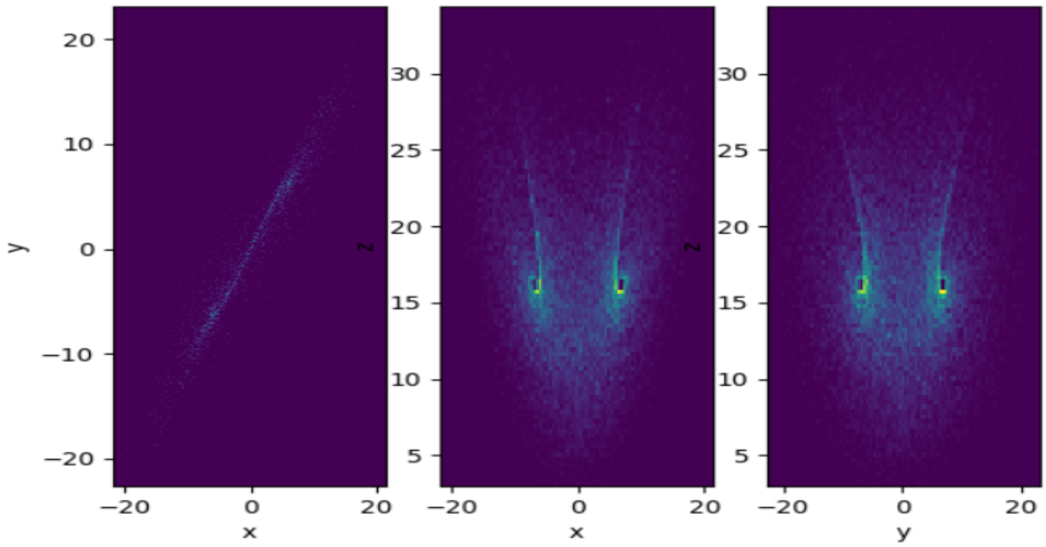
To reconstruct the parameter, we choose 1000 gaussion whose mean is generated by LHS sampling in the domain of data.Using Gradient method with lr = 0.01 and 30000 epoch and Least square method , the reconstruction result is

	a	b	c
GD	0.1051	0.080	13.9053
Error	0.0051	-0.0199	-0.0947
LS	0.1051	0.0801	13.9053
Error	0.0051	-0.0199	-0.0947

3.3 Chen System

$$\begin{cases} \dot{x} &= & a(y-x) \\ \dot{y} &= & (c-a)x-\mathrm{xz}+\mathrm{cy} \\ \dot{z} &= & \mathrm{xy}-\mathrm{bz} \end{cases},$$

Here $a=40, b=3, c=28$. initial condition is [2,2,5],$T=20000$, dt $=0.01$.the projection of density map wrt xy,xz,yz direction is below.

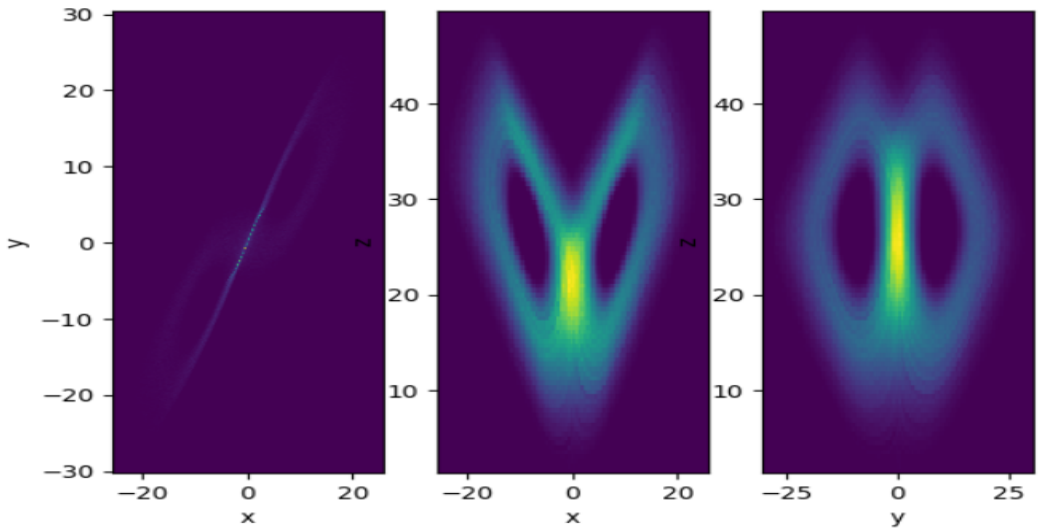


	a	b	c
GD	40.2535	3.0110	28.1462
Error	0.2535	0.0110	0.1462
LS	40.2535	3.0110	28.1462
Error	0.2535	0.0110	0.1462

3.4 Arctan Lorentz System

$$\begin{cases} \dot{x} &= 50\text{arctan}\left(\dfrac{\sigma(y-x)}{50}\right) \\ \dot{y} &= 50\text{arctan}\left(\dfrac{x(\rho-z)-y}{50}\right) \\ \dot{z} &= 50\text{arctan}\left(\dfrac{(\text{xy}-\beta z)}{50}\right) \end{cases},$$

Here $\sigma=10, \rho=28, \sigma=8/3$, initial condition is [1,1,5],$T=20000$, dt$=0.01$.the projection of density map wrt xy,xz,yz direction is below.



Here we use stochastic gradient descent, for each batch we sample 50 gaussian from the 1000 LHS gaussian list.The learning rate is 0.1 and a total 3000 epoch is used in training.

	a	b	c
GD	10.0177	28.0473	2.7052
Error	0.0177	0.0473	0.0386

Also we can see that the system converge in around 700 epoch.

```
step = 0, loss = 13606, W = [[0.2606559097766876], [-0.18592116236686707], [-0.29381808638572693]]
step = 50, loss = 7330.8, W = [[0.23338229954242706], [4.244386672973633], [-1.6133419275283813]]
step = 100, loss = 3267.05, W = [[-0.7432518601417542], [8.373147010803223], [-1.1452692747116089]]
step = 150, loss = 1874.38, W = [[-0.776563286781311], [13.598703384399414], [-1.652927279472351]]
step = 200, loss = 9870.32, W = [[0.17139899730682373], [16.284543991088867], [-0.4884900450706482]]
step = 250, loss = 1133.35, W = [[1.718989610671997], [18.993850708007812], [-0.689951479434967]]
step = 300, loss = 3519.36, W = [[2.7393362522125244], [21.802255630493164], [0.09377888590097427]]
step = 350, loss = 4682.72, W = [[4.563082218170166], [24.928869247436523], [1.011198878288269]]
step = 400, loss = 1050.5, W = [[7.475602149963379], [27.00672149658203], [2.163640260696411]]
step = 450, loss = 180.075, W = [[10.267330169677734], [28.045682907104492], [2.724902629852295]]
step = 500, loss = 229.029, W = [[10.095870018005371], [27.876361846923828], [2.7024264335632324]]
step = 550, loss = 198.879, W = [[9.993453025817871], [28.066232681274414], [2.586441993713379]]
step = 600, loss = 215.098, W = [[9.981377601623535], [28.20126724243164], [2.7300398349761963]]
step = 650, loss = 83.5705, W = [[9.908446311950684], [27.98430824279785], [2.6761767864227295]]
step = 700, loss = 127.472, W = [[10.000760078430176], [28.072101593017578], [2.6652183532714844]]
step = 750, loss = 12.7585, W = [[10.049041748046875], [27.909833908081055], [2.651001214981079]]
step = 800, loss = 34.0581, W = [[10.044719696044922], [28.067684173583984], [2.6968798637390137]]
step = 850, loss = 96.3634, W = [[10.046515464782715], [27.925172805786133], [2.69292950630188]]
step = 900, loss = 24.1463, W = [[9.920768737792969], [28.056562423706055], [2.671236276626587]]
step = 950, loss = 20.2743, W = [[9.971670150756836], [27.984514236450195], [2.6705563068389893]]
step = 1000, loss = 20.6455, W = [[9.991427421569824], [28.039653778076172], [2.6724250316619873]]
step = 1050, loss = 137.069, W = [[10.036661148071289], [28.008699417114258], [2.706580877304077]]
step = 1100, loss = 32.3733, W = [[10.027220726013184], [27.92790985107422], [2.6619715690612793]]
step = 1150, loss = 70.1727, W = [[10.046793937683105], [27.960403442382812], [2.617043972015381]]
step = 1200, loss = 30.4325, W = [[10.12477970123291], [27.953689575195312], [2.6405186653137207]]
step = 1250, loss = 64.829, W = [[9.998767852783203], [27.955413818359375], [2.728026866912842]]
step = 1300, loss = 11.7296, W = [[9.963363647460938], [27.978424072265625], [2.66508412361145]]
step = 1350, loss = 29.1377, W = [[9.940669059753418], [27.8826904296875], [2.657597064971924]]
step = 1400, loss = 50.655, W = [[10.066084861755371], [27.976806640625], [2.71030330657959]]
step = 1450, loss = 32.8353, W = [[9.989806175231934], [28.007421493530273], [2.665088415145874]]
step = 1500, loss = 8.36017, W = [[10.032815933227539], [27.968191146850586], [2.657179594039917]]
step = 1550, loss = 13.306, W = [[10.052788734436035], [28.055137634277344], [2.643028736114502]]
step = 1600, loss = 64.1716, W = [[10.022038459777832], [28.030010223388672], [2.657855987548828]]
step = 1650, loss = 13.7256, W = [[9.959837913513184], [27.931169509887695], [2.730159044265747]]
```

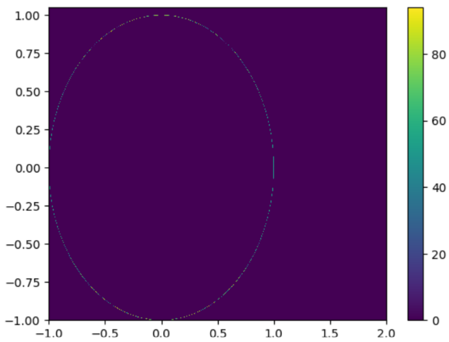
3.5 2D Periodic Attractor

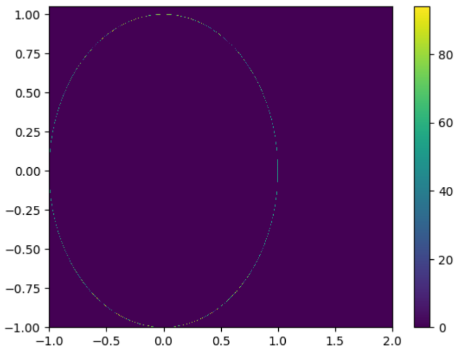
$$\begin{cases} \dot{x} &= x - y - (x^2 + y^2)x \\ \dot{y} &= x + y - (x^2 + y^2)y \end{cases},$$

This is a planar system which has a unique closed orbit γ and is a periodic attractor.

Note that the phase chart of such equation is almost the same as

$$\begin{cases} \dot{x} &= x - 0.3\,y - (x^2 + y^2)x \\ \dot{y} &= 0.3x + y - (x^2 + y^2)y \end{cases},$$





we parametrize the equation as follow

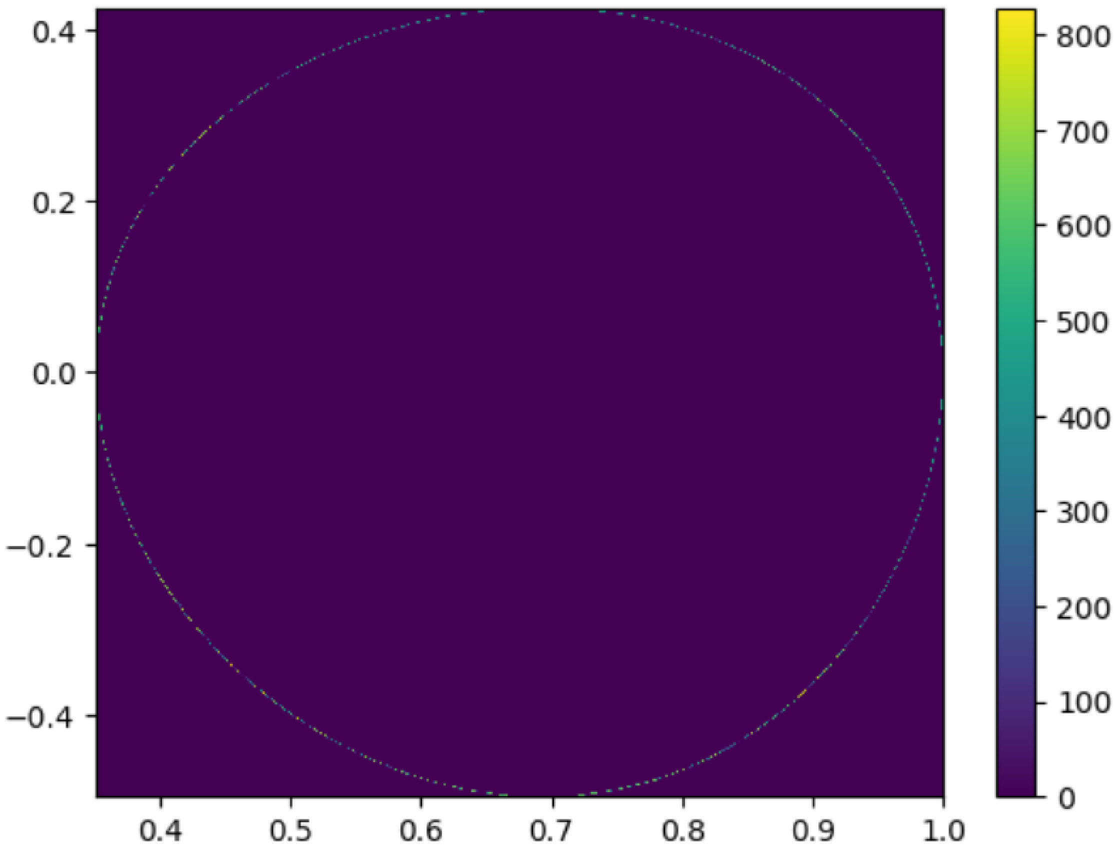
$$\begin{cases} \dot{x} &= x - y - \mathrm{a}x^3 - \mathrm{b}xy^2 \\ \dot{y} &= x + y - \mathrm{c}x^2 y - y^3 \end{cases},$$

The reference result is [1,1,1,1]

	a	b	c	d
LS	0.9982	1.0031	0.9954	1.0021
Error	0.0018	0.0031	0.0046	0.0021

3.6 Lotka-Volterra System

$$\begin{cases} \dot{x} &= 1 - \exp y \\ \dot{y} &= \exp x - 2 \end{cases},$$



we parametrize the equation as follow

$$\left\{\begin{array}{lcl} \dot{x} & = & A - \mathrm{B}\exp y \\ \dot{y} & = & \mathrm{C}\exp x - 2 \end{array}\right. ,$$

The reference result is [1,1,1]

	A	B	C
LS	1.0001	1.0001	1.000
Error	0.0001	0.0001	0.000

4 Teleportation Regularization for WCR

4.1 有限体积法

Yunan Yang的文章重对于正问题方程：

$$\left\{\begin{array}{c} \nabla\cdot(v(x,\theta)\rho(x,\theta)) = 0. \\ \int_\Omega \rho(x,\theta) = 1 \end{array}\right.$$

利用有限体积法得到离散问题：

$$\left\{\begin{array}{c} K_{\mathrm{mat}}(\theta)\rho(x_i,\theta) = \mathbf{0} \\ \rho(x_i,\theta)\cdot\mathbf{1} = 1 \end{array}\right.$$

记$M(\theta) = I + \mathrm{cK_{mat}}$,原离散问题变为

$$\left\{\begin{array}{c} M(\theta)\rho(x_i,\theta) = \rho(x_i,\theta) \\ \rho(x_i,\theta)\cdot\mathbf{1} = 1 \end{array}\right.$$

求解$\rho(x_i,\theta)$即求解矩阵1特征值的特征向量。在离散问题中可能面临特征空间维数大于1的情况，为了避免这点引入 teleportation regularization:

$$M_\varepsilon(\theta) = (1-\varepsilon)M + \varepsilon n^{-1}\mathbf{1}\mathbf{1}^T = (1-\varepsilon)(I + \mathrm{cK_{mat}}(\theta)) + \frac{\varepsilon}{n}\mathbf{1}\mathbf{1}^T$$

最后方程变为：

$$M_\varepsilon(\theta)\rho = \rho,\ \rho\cdot\mathbf{1} = 1$$

4.2 Galerkin

考虑上面的离散矩阵$M_\varepsilon(\theta)$的方程形式：

$$(1-\varepsilon)(\rho(x,\theta) + \nabla\cdot(v(x,\theta)\rho(x,\theta))) + \varepsilon = \rho(x,\theta)$$

化简得到：

$$\nabla\cdot(v(x,\theta)\rho(x,\theta)) = \varepsilon\rho(x,\theta) - \varepsilon$$

其弱形式为

$$-\int_\Omega v(x,\theta)\cdot\nabla\phi(x)\ \rho(x,\theta)\mathrm{dx} = \varepsilon\int_\Omega \rho(x,\theta)\phi(x)\mathrm{dx} - \varepsilon\int_\Omega \phi(x)\mathrm{dx}$$

取$\phi(x)$ 为高斯分布的密度函数得到：

$$\mathbb{E}_{x\sim\rho}[v(x,\theta)\nabla\phi(x)] = -\varepsilon\mathbb{E}_{x\sim\rho(x)}[\phi(x)] + \varepsilon\mathbb{E}_{x\sim\phi}[1] = -\varepsilon\mathbb{E}_{x\sim\rho(x)}[\phi(x)] + \varepsilon$$

此时右端项非0.

4.3 数值实践

首先对于自洽系统， 至多可以在乘以一个常数的意义下确定其参数： 对于方程$\frac{\mathrm{dx}}{\mathrm{dt}} = v(x,\theta)$,对于$\frac{\mathrm{dx}}{\mathrm{dt}} = \mathrm{kv}(x,\theta)$,令$t' = \mathrm{kt}$

$$\frac{\mathrm{dx}}{\mathrm{dt}'} = \frac{\mathrm{dx}}{\mathrm{dt}}\frac{\mathrm{dt}}{\mathrm{dt}'} = \mathrm{kv}(x,\theta)\frac{1}{k} = v(x,\theta)$$

所以在时间趋于无穷的情况下得到的invariant measure 相同。

故数值实践中

1. 考虑符号回归类方法，假设自洽系统的右端项均在一个Dictionary中（如完整3阶多项式系统）

2. 将$v(x,\theta)$进行Dictionary的参数化， 通过Teleportation Regularization方法， 找到最可能属于字典的其中一项（比如多项式字典中， 回归得到系数相对大小最大的一项x）， 将它的系数设为1移到右端

3. 利用一般WCR去求解最小二乘问题（如STrigde方法）

4.4 Teleportation Regularization 数值算例

4.4.1 Lorentz

首先用完整2阶多项式字典对原问题的Teleportation Regularization问题进行反演，$\varepsilon =$

$$\begin{cases} \dot{x} &= \sigma(y-x) \\ \dot{y} &= x(\rho-z)-y \\ \dot{z} &= \mathrm{xy}-\beta z \end{cases},$$

字典顺序为$\{1,x,y,z,x^2,\mathrm{xy},\mathrm{xz},y^2,\mathrm{yz},z^2\}$,反演参考为$\sigma=10$, $\rho=28$, $\beta=8/3$

取gaussian sigma = 8，$\varepsilon = 2e-2$,得到系数回归结果如下

```
tensor([[ 4.1279e-02, -1.5794e+00,  1.5579e+00,  7.1639e-03,  2.1076e-03,
         -6.0172e-04,  2.2723e-03, -2.2599e-04, -6.7123e-04, -4.5122e-04],
        [ 1.6689e-01,  4.0327e+00, -8.1171e-03, -3.4494e-03,  2.5010e-03,
         -3.3246e-03, -1.4643e-01,  1.3954e-03, -2.5878e-03, -2.1737e-04],
        [-3.3081e+00,  4.4388e-02, -1.1903e-02, -2.1318e-02,  4.9486e-02,
          1.0762e-01, -1.3884e-03,  8.6351e-03,  5.0898e-04, -1.0614e-02]])
```

其中最大项为第二分量的x，取它为右端项，系数为28（与方程setting对齐），(sigma = 1,threshold = 0.2)得到反演结果为

$$\begin{cases} \dot{x} &= -0.9987x+10.0009y \\ \dot{y} &= 28x-0,9774y-0.9999\mathrm{xz} \\ \dot{z} &= 0.9975\mathrm{xy}-2.6562z \end{cases},$$

字典顺序为$\{x,y,z,x^2,\mathrm{xy},\mathrm{xz},y^2,\mathrm{yz},z^2\}$其中三行分别为v的分量，列为字典顺序对应系数，此时第二分量的x结果最大，取它为右端项

```
tensor([[ 1.5303e-01, -1.3508e-01,  1.6313e-03,  8.4141e-05,  3.0035e-04,
          4.1891e-04, -2.2908e-04, -7.3108e-04, -4.7781e-05],
        [-5.2171e-01,  9.4771e-02,  3.0426e-03,  1.4463e-03, -5.2073e-04,
          1.7912e-02, -1.8009e-04, -2.0765e-03, -1.5189e-04],
        [ 4.2944e-02, -1.9131e-02,  2.5969e-02, -2.5170e-03, -1.5881e-02,
         -1.0613e-03,  2.0424e-03,  3.1047e-04,  7.1841e-04]])
```

最后，直接用Teleporation Regularization Problem 反演得到的结果中，第三列的常数未能剔除，但前两位准确

```
tensor([[  0.0000,   5.9633,  -5.6619,   0.0000,   0.0000,   0.0000,   0.0000,
           0.0000,   0.0000,   0.0000],
        [  0.0000, -17.1209,   1.3846,   0.0000,   0.0000,   0.0000,   0.5879,
           0.0000,   0.0000,   0.0000],
        [ -3.6466,   0.0000,   0.0000,   1.6937,   0.0000,  -0.5558,   0.0000,
           0.0000,   0.0000,   0.0000]])
```

4.4.2 Rossler

$$\begin{cases} \dot{x} &= -y-z \\ \dot{y} &= x+\mathrm{ay} \\ \dot{z} &= b+z(x-c) \end{cases},$$

（此时方程有常数项）

字典顺序为$\{1,x,y,z,x^2,\mathrm{xy},\mathrm{xz},y^2,\mathrm{yz},z^2\}$,反演参考为$a=0.1,b=0.1,c=14$

```
tensor([[-1.3493e-04, -3.4332e-05,  6.6266e-03,  6.4965e-03, -1.6510e-06,
         -4.1418e-05, -1.9580e-05, -5.7243e-05,  1.1313e-05, -4.7359e-06],
        [-1.2476e-04, -6.8512e-03, -5.6615e-04,  5.4660e-04,  5.2438e-05,
          5.0311e-05, -4.1061e-05,  3.8604e-06, -5.9305e-05,  4.0475e-06],
        [-2.4169e-03,  7.0269e-06, -5.9525e-05,  9.1364e-02, -2.6246e-06,
         -1.4885e-06, -6.2352e-03, -3.3528e-06, -3.1649e-04, -8.4410e-05]])
```

得到最大值在第三分量的z元素，将其设置为-14移到右端项，得到反演结果为

$$\begin{cases} \dot{x} &= -1.0275y-0.9709z \\ \dot{y} &= 1.0273x+0.1082y \\ \dot{z} &= 0.0935+1.0037z(x-0.1) \end{cases},$$

4.4.3 Chen

$$\begin{cases} \dot{x} &= a(y-x) \\ \dot{y} &= (c-a)x-\mathrm{xz}+\mathrm{cy} \\ \dot{z} &= \mathrm{xy}-\mathrm{bz} \end{cases},$$

字典顺序为$\{1,x,y,z,x^2,\mathrm{xy},\mathrm{xz},y^2,\mathrm{yz},z^2\}$,反演参考为 $a=40,b=3,c=28$

取gaussian sigma = 8，$\varepsilon=2e-3$,得到系数回归结果如下

```
tensor([[ 1.3474e-02,  2.8500e+00, -2.8576e+00, -1.5275e-03,  3.3119e-05,
         -2.6158e-05,  1.5786e-03,  3.3330e-06, -1.1081e-03,  4.0735e-05],
        [ 1.3281e-02,  8.1246e-01, -1.9790e+00, -1.2502e-03,  1.0544e-05,
          2.1594e-04,  7.3971e-02, -1.9881e-04, -1.1357e-03,  2.2826e-05],
        [-9.4666e-02,  2.1607e-04,  5.3141e-04,  2.2911e-01,  5.1140e-03,
         -8.0984e-02,  2.3372e-05,  3.8023e-03, -6.0923e-05, -4.1401e-04]])
```

其中最大项为第一分量的y，取它为右端项，系数为40（与方程setting对齐）。

并且在这个例子中的Teleportation Regularization Problem直接回归也可以得到相当准确的结果

```
tensor([[  0.0000,  18.8767, -18.8759,   0.0000,   0.0000,   0.0000,   0.0000,
           0.0000,   0.0000,   0.0000],
        [  0.0000,   5.6363, -13.1956,   0.0000,   0.0000,   0.0000,   0.4724,
           0.0000,   0.0000,   0.0000],
        [  0.0000,   0.0000,   0.0000,   1.4163,   0.0000,  -0.4720,   0.0000,
           0.0000,   0.0000,   0.0000]])
```

$$\begin{cases} \dot{x} &= -39.9995y + 40x \\ \dot{y} &= 12.0002x + 28.0006y - \mathrm{xz} \\ \dot{z} &= 1.0001\mathrm{xy} - 3.0003z \end{cases},$$