Neural Operator for Nonlinear Filter

BY ZHIJUN ZENG

2024年4月5日

1 Problem setting

Consider the signal based model

$$\begin{cases} d\,x_t = f(x_t, t)\,d\,t + g(x_t, t)\,d\,v_t \\ d\,y_t = h(x_t, t)\,d\,t + d\,w_t \end{cases}$$

where $x_t \in \mathbb{R}^n$ is the state of system at time t, the initial state $x_0 \sim p_0$ satisfying some initial distribution. The observation $y_t \in \mathbb{R}^m$, initialize at $y_0 = 0$. The noise is vector based Brownian motion $\mathbb{E}[\mathrm{dv}_t \mathrm{dv}_t^{\mathrm{T}}] = Q(t)\mathrm{dt} \in \mathbb{R}^{n \times n}$, and $\mathbb{E}[\mathrm{dw}_t \mathrm{dw}_t^{\mathrm{T}}] = S(t)\mathrm{dt} \in \mathbb{R}^{n \times n}$, here $S(t) > 0$.

The unnormalized conditional density function of the states x_t, denoted by $\sigma(x, t)$ satisfies the following SPDE (DMZ)

$$\begin{cases} d\,\sigma(x, t) = \mathcal{L}\,\sigma(x, t)\,d\,t + \sigma(x, t)\,h^T(x, t)\,S^{-1}\,d\,y_t \\ \sigma(x, 0) = \sigma_0(x) \end{cases}$$

where $\sigma(x, 0) = p_0(x)$, and

$$\mathcal{L}(\cdot) := \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i\, \partial x_j} ((g\,Q\,g^T)_{ij} \cdot) - \sum_{i=1}^{n} \frac{\partial\,(f_i \cdot)}{\partial x_i}$$

Here we call f the drift term, g the diffusion term, h the observation term.

By an invertible exponential transformation

$$\sigma(x, t) = \exp\,(h^T(x, t)\,S^{-1}(t)\,y_t)\,u(x, t)$$

We obtain an PDE with stochastic coefficients

$$\begin{cases} \begin{aligned} \frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial t} (h^T S^{-1})\, y_t\, u(x, t) = & \\ & \exp\,(-h^T\,(x, t)\,S^{-1}(t)\,y_t) \left(\mathcal{L} - \frac{1}{2} h^T S^{-1}\, h \right) \\ & \cdot (\exp\,(-h^T\,(x, t)\,S^{-1}(t)\,y_t)\,u(x, t)) \\ u(x, 0) = \sigma_0(x) & \end{aligned} \end{cases}$$

This equation is called pathwise robust DMZ equation.

Assume that the observation time sequences $0 = t_0 < \cdots < t_{N_t} = T$ are given. Then the observation $\{y_{t_j}\}$ at each observation time t_j are unknown until the online experiment. Therefore, one can freeze the y_t to $y_{t_{j-1}}$ in time $t_{j-1} \leqslant t < t_j$, making the exponential transformation

$$u_j(x, t) = \exp\,(h^T(x, t)\,S^{-1}(t)\,y_{t_{j-1}})\,u(x, t).$$

One obtain that u_j satisfies the FKE

$$\frac{\partial}{\partial t}u_j(x,t) = \left(\mathcal{L} - \frac{1}{2}h^T S^{-1} h\right)u_j(x,t)$$

The finite difference solver:

M. Yueh, W. Lin, and S. T. Yau, "An efficient num-erical method

for solving high-dimensional nonlinear filtering problems," Comˆ

mun. Inf. Syst., vol. 14, no. 4, pp. 243-262, 2014.

2 Numerical Method for NLS

For

$$\begin{cases} d\,x_t = f(x_t,t)\,d\,t + d\,v_t \\ d\,y_t = h(x_t,t)\,d\,t + d\,w_t \end{cases}$$

The FKE becomes

$$\frac{\partial}{\partial t}u_j(x,t) = \frac{1}{2}\Delta\,u_j(x,t) - f(x)\cdot\nabla u_j(x,t) - \left(\nabla\cdot f(x) + \frac{1}{2}\|h(s)\|_2^2\right)u_j(x,t)$$

The initial condition is $u(0,s)=\sigma_0(s)$, once the new measurement $y(t_k)$ arrive we correct the solution $(u_j(x,t_k))$ at t_k by

$$\exp\{[y(t_k) - y(t_{k-1})]\cdot h(x)\}u_j(x,t_k)$$

Initially, we discretize the time interval $[t_{k-1},t_k]$ by taking a uniform partition $[t_{k-1},t_{k-1}+\Delta t,\ldots t_{k-1}+N\Delta t = t_k]$.

For the sake of computations, we restrict the domain $\mathbb{R}^D$ to $[-R,R]^D$ cells and partition it uniformly.

2.1 2D example

$$\begin{cases} d\,x_1 = \cos(x_1)d\,v_1 \\ d\,x_2 = \cos(x_2)d\,v_2 \\ d\,y_1 = x_1^3\,d\,t + d\,w_1 \\ d\,y_2 = x_2^3\,d\,t + d\,w_2 \end{cases}$$

Then

$$f(x) = \begin{bmatrix} \cos(x_1) \\ \cos(x_2) \end{bmatrix}, h(x) = \begin{bmatrix} x_1^3 \\ x_2^3 \end{bmatrix}$$

The strong form of equation will be

$$\frac{\partial}{\partial t}u_j = \frac{1}{2}\Delta\,u_j - \cos(x_1)\partial_{x_1}u_j - \cos(x_2)\partial_{x_2}u_j + \left(\sin(x_1) + \sin(x_2) - \frac{1}{2}(x_1^6 + x_2^6)\right)u_j$$

The weak form of equation will be

$$\frac{<u_j(t+\Delta t), v> - <u_j(t), v>}{\Delta t} = -\frac{1}{2}<\nabla u_j, \nabla v> - <f\cdot\nabla u_j, v> + <a(x)u_j, v>$$

where $a(x)=\sin(x_1)+\sin(x_2)-\frac{1}{2}(x_1^6+x_2^6)$. The boundary condition is dirichlet boundary condition.

Here is some discretization parameter: $\Delta t=0.001, \sigma_0(s)=\exp\{-10|x|^2\}, h=0.1, \Omega=[-5,5]^2$.

3 Numerical Example

3.1 1D

Consider the problem of

$$\begin{cases} \mathrm{dx}_1=\mathrm{dv}_1 \\ \mathrm{dy}_1=x_1(1+0.2\cos(x_1))\mathrm{dt}+\mathrm{dw}_1 \end{cases}$$

where $\mathbb{E}[\mathrm{dw}_t\mathrm{dw}_t]=I_1\mathrm{dt}, \mathbb{E}[\mathrm{dv}_t\mathrm{dv}_t]=0.1I_1\mathrm{dt}$. For filter problem, $x(0)=[1.0]$.

For operator learning problem, $g(t)=1, h(x,t)=x(1+0.2\cos(x)), Q(t)=0.1, S(t)=1$, the FKE becomes

$$\frac{\partial}{\partial t}u_j(x,t)=\left(\frac{1}{2}0.1\frac{\partial^2 u_j}{\partial x^2}-\frac{1}{2}x^2(1+0.2\cos(x))^2 u_j(x)\right)$$

The neural operator problem can be defined as predicting $u_j(x,t+\Delta t)$ from $u_j(x,t)$.

3.2 2D

Consider the problem of

$$\begin{cases} d\,x_1=d\,v_1 \\ d\,x_2=d\,v_2 \\ d\,y_1=x_1\,(1+0.2\cos(x_2))\,d\,t+d\,w_1 \\ d\,y_2=x_2\,(1+0.2\cos(x_1))\,d\,t+d\,w_2 \end{cases}$$

where $\mathbb{E}[\mathrm{dw}_t\mathrm{dw}_t]=I_2\mathrm{dt}, \mathbb{E}[\mathrm{dv}_t\mathrm{dv}_t]=0.1I_2\mathrm{dt}$. For filter problem, $x(0)=[1.0,1.2]^\top$.

For operator learning problem, $g(t)=I_2, h(x,t)=\begin{bmatrix} x_1\,(1+0.2\cos(x_2)) \\ x_2\,(1+0.2\cos(x_1)) \end{bmatrix}, Q(t)=0.1I_2, S(t)=I_2$, the FKE becomes

$$g\,Q\,g^T=0.1I_2, h^T S^{-1}\,h=h^T h=$$

$$\frac{\partial}{\partial t}u_j(x_1,x_2,t)=\frac{1}{2}0.1\left(\frac{\partial^2 u_j}{\partial x_1{}^2}+\frac{\partial^2 u_j}{\partial x_2{}^2}\right)-[x_1^2\,(1+0.2\cos(x_2))^2+x_2^2\,(1+0.2\cos(x_1))^2]u_j(x_1,x_2,t)$$

The neural operator problem can be defined as predicting $u_j(x,t+\Delta t)$ from $u_j(x,t)$.

3.3 Post-process

Obtained $u_j(x,t+\Delta t)$, we can compute

$$u(x,t+\Delta t)=\exp\left(-h^T(x,t)\,S^{-1}(t)\,y_{t_{j-1}}\right)u_j(x,t+\Delta t)$$

And the prediction of $\sigma(x,t)$ is

$$\sigma(x,t+\Delta t)=\exp\left(h^T(x,t)\,S^{-1}(t)\,y_{t_j}\right)u(x,t+\Delta t)$$