Parameter estimation and modeling of nonlinear dynamical systems based on Runge–Kutta physics-informed neural network

https://doi.org/10.1007/s11071-023-08933-6

RK-PINN

2024-11-08

Adding PINN to RKNN

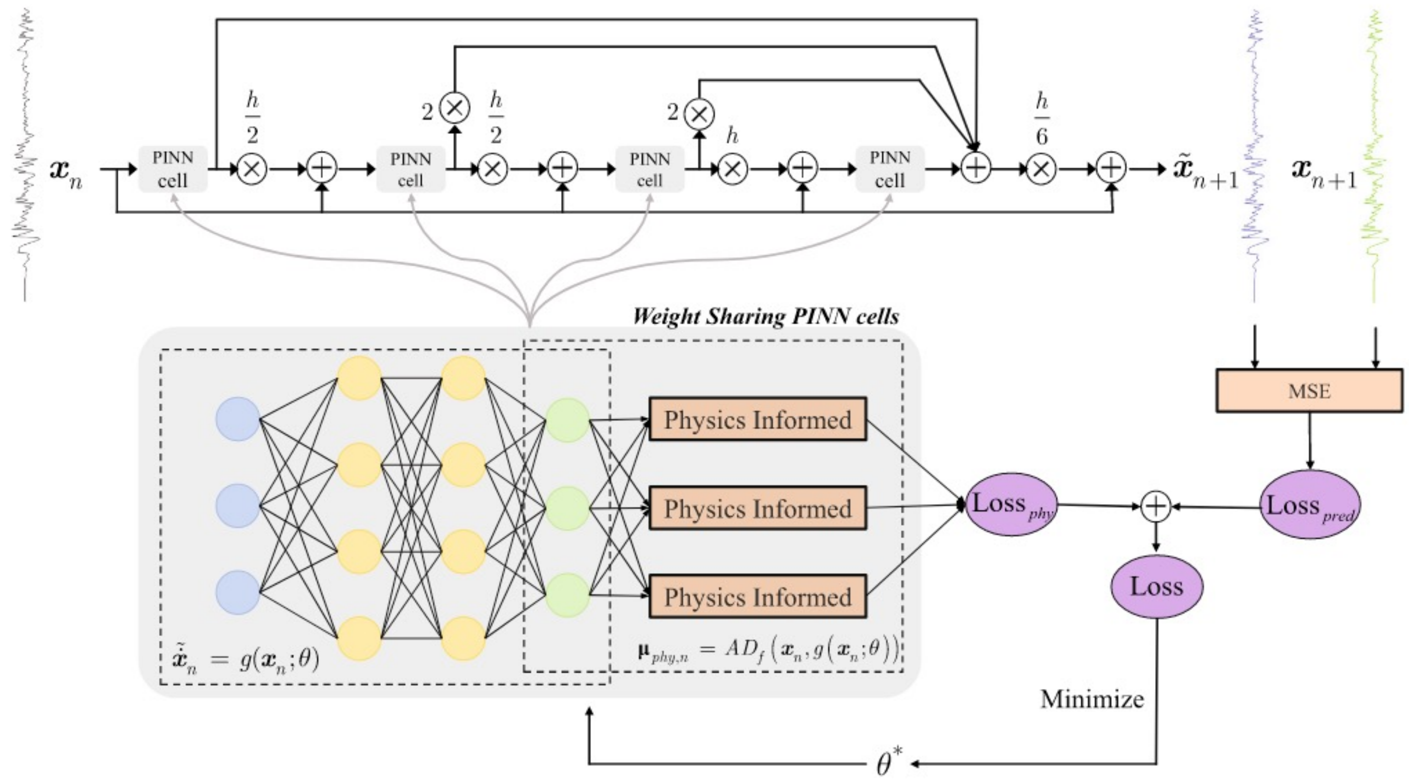

Fig. 2 Flowchart of RK4-PINN

Loss

- $Loss_{pred}$, MSE between prediction and truth
- $Loss_{phy}$
  - F() encodes system physics, for time-invariant system, dF(x)/dx at different time should be the same, i.e. coefficient of variance should be zero.
  - In a more general form, let the distribution of dF(x)/dx be q, if we know the ideal distribution as p from physical knowledge, KL divergence of P from Q should be 0.
  - the KL divergence of P from Q is the expected excess surprise from using Q as a model instead of P when the actual distribution is P

Weighted Loss

- Loss = α $Loss_{pred}$ +β $Loss_{phy}$
- Suppose network is MLP, $Loss_{phy}$ mathmatically equals to:

$$\frac{\partial \tilde{\boldsymbol{x}}_n}{\partial \boldsymbol{x}_n} = \frac{\partial \tilde{\boldsymbol{x}}_n}{\partial \boldsymbol{A}_{m-1}} \cdot \frac{\partial \boldsymbol{A}_{m-1}}{\partial \boldsymbol{Z}_{m-1}} \cdot \frac{\partial \boldsymbol{Z}_{m-1}}{\partial \boldsymbol{A}_{m-2}} \cdots \frac{\partial \boldsymbol{A}_1}{\partial \boldsymbol{Z}_1} \cdot \frac{\partial \boldsymbol{Z}_1}{\partial \boldsymbol{A}_0}$$
$$= \prod_{i=1}^{m-1} \left(\boldsymbol{W}_i^T \cdot \mathrm{diag}\left(1 - \tanh^2(\boldsymbol{Z}_i)\right)\right) \cdot \boldsymbol{W}_m$$

- Minimize $Loss_{phy}$ makes Wi close to 0
  - $Loss_{phy}$ is a regulation term that preventing overfit
  - Alpha should larger than beta

Parameter estimation

$$\tilde{s} = \frac{\partial \tilde{\dot{x}}}{\partial y} = -\frac{\partial \tilde{\dot{x}}}{\partial x}$$

$$\dot{x} = s(y - x)$$

$$\dot{y} = rx - y - xz$$

$$\dot{z} = -qz + xy$$

$$\tilde{r} = \frac{\partial(\tilde{\dot{y}} + y + xz)}{\partial x}$$

$$\tilde{q} = \frac{\partial(-\tilde{\dot{z}} + xy)}{\partial z}$$

System physics, i.e. F()

$$\text{Loss}_{\text{phy}} = c_v^2(\tilde{s}) + c_v^2(\tilde{r}) + c_v^2(\tilde{q})$$

dF(x) / dx