Self-supervised hamiltonian mechanics neural networks

Zhan Youqiu

August 30, 2020

Abstract

Abstract abstract.

Keywords — keyword1, keyword2, keyword3

Contents

1 Introduction

Some text [3].

2 Canonical equation and ODE

To build a common sense of the physics theories it is going to involve, the canonical equation is introduced.

The canonical equation is a set of ordinary differential equations (ODE) whose solution depicts the motion of the system. The equation in mathematical form is [4][1, p. 65][6, p. 132]

$$\dot{\mathbf{q}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}}, \tag{1}$$

where $\mathbf{q} \in \mathbb{R}^n$ is the **generalized coordinates**, $\mathbf{p} \in \mathbb{R}^n$ is the **generalized momentum**, and $\mathcal{H}$ is the **hamiltonian** of the system, which is a scalar function w.r.t. t, $\mathbf{q}$, and $\mathbf{p}$. n is the number of degrees of freedom (DOF). A hamiltonian is specific for a specific system.

The tuple $\mathbf{x} := (\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2n}$ is called the **canonical coordinates**. In computer programs, it is convenient to write Equation 1 in form of

$$\dot{\mathbf{x}} = \mathbf{f}\left(t, \mathbf{x}\right), \tag{2}$$

which is the common form of ODE. Here in our specific case,

$$\mathbf{f}\left(t, \mathbf{x}\right) := \boldsymbol{\omega} \nabla_{\mathbf{x}} \mathcal{H}, \tag{3}$$

where the notion $\boldsymbol{\omega}\nabla_{\mathbf{x}}$ denotes the **symplectic gradient** w.r.t. $\mathbf{x}$, whose first n components is the gradient w.r.t. the last n components of $\mathbf{x}$, and the last n components is the negative gradient w.r.t. the first n components of $\mathbf{x}$.

One of properties of the symplectic gradient is that, moving along the symplectic gradient field of a scalar does not change the value of the scalar function, which means that the value of $\mathcal{H}$ is conserved if $\frac{\partial \mathcal{H}}{\partial t} = 0$. In fact, the physical meaning of $\mathcal{H}$ is the energy, so its conservation is obvious.

According to Equation 3, the difference between $\mathbf{x}$ at 2 different times is an integral

$$\mathbf{x}\left(t_2\right) = \mathbf{x}\left(t_1\right) + \int_{t_1}^{t_2} \mathbf{f}\left(t, \mathbf{x}\left(t\right)\right) \mathrm{d}t. \tag{4}$$

The integral can be calculated using the torchdiffeq Python package [2].

3 The training

Our goal is to derive the function $(t, \mathbf{x}) \mapsto \mathcal{H}$ according to the dataset containing a series of samples in form of $(t, \mathbf{x})$ on a series of possible motions of the system.

The dataset does not contain the $\dot{\mathbf{x}}$ infomation, which acts as the ground truth in the supervised model [3]. Our model is self-supervised, and thus does not need the $\dot{\mathbf{x}}$ infomation.

The model uses the loss inspired from Equation 4

$$\mathcal{L} := \mathrm{MSE}\left(\mathbf{x}\left(t_1\right) + \int_{t_1}^{t_2} \boldsymbol{\omega}\nabla_{\mathbf{x}}\mathcal{H}\mathrm{d}t, \mathbf{x}\left(t_2\right)\right), \tag{5}$$

where $(t_1, \mathbf{x}(t_1))$ and $(t_1, \mathbf{x}(t_1))$ are 2 samples from the same motion of the system. The complete process of a traning circle is shown in Figure 2. For comparison, the train circle of the supervised hamiltonian neural network is shown in Figure 1.

The Adam optimizer [5] is used for optimizing the neural network.

4 Tasks

4.1 Free particle

A free particle is a system with 1 DOF whose hamiltonian is

$$\mathcal{H}\left(t, q, p\right) := \frac{p^2}{2m},$$

where m is the mass of the particle. To be simple, we take $m = 1/2$.

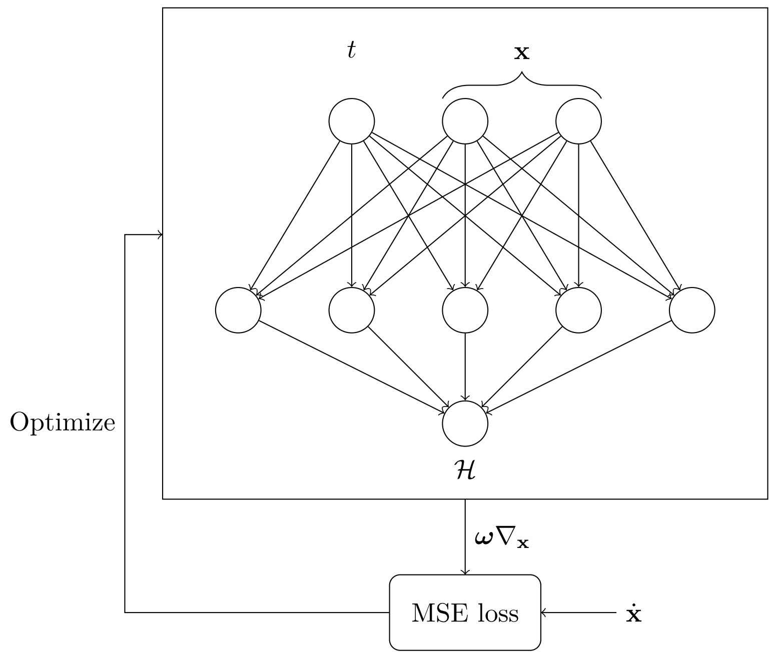

Figure 1: The train circle of a supervised hamiltonian neural network [3]

4.2 Harmonic oscillator

References

[1] V. I. Arnol d, K. Vogtmann, and A. Weinstein. *Mathematical methods of classical mechanics*. Springer, 2nd edition, 1989.

[2] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud. Neural ordinary differential equations. *Advances in Neural Information Processing Systems*, 2018.

[3] S. Greydanus, M. Dzamba, and J. Yosinski. Hamiltonian neural networks, 2019.

[4] L. N. Hand and J. D. Finch. *Analytical mechanics*. Cambridge University Press, 2008.

[5] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization, 2017.

[6] L. D. Landau, L. E. Mikhaĭlovich, J. B. Sykes, and J. S. Bell. *Mechanics*. Butterworth-Heinemann, 3rd edition, 1976.

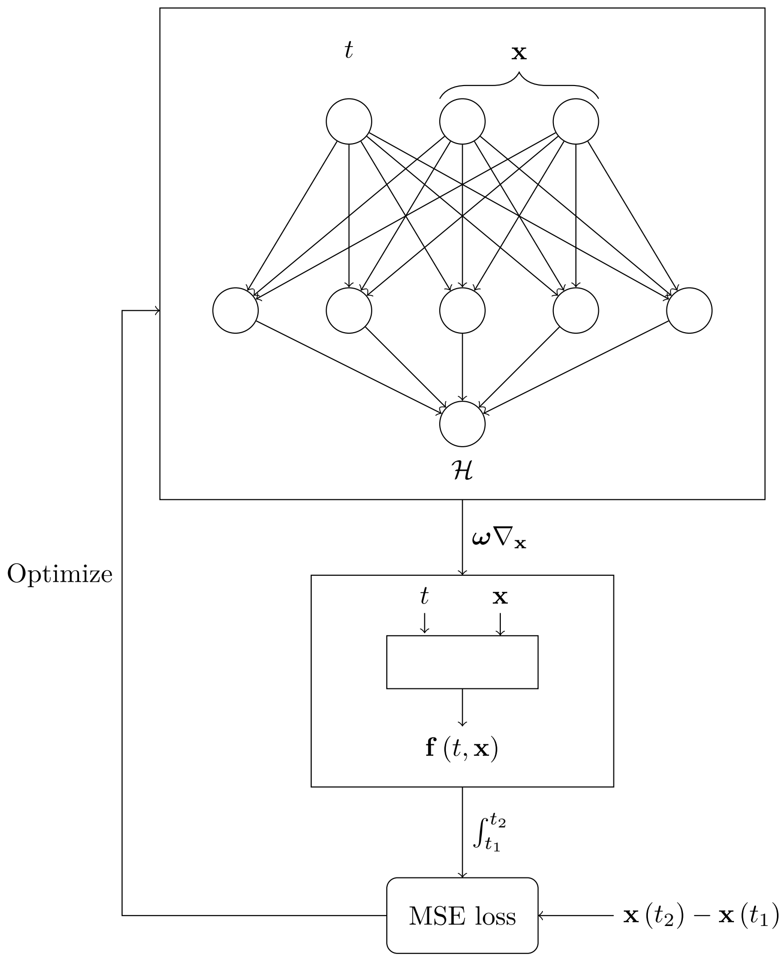

Figure 2: The train circle of a self-supervised hamiltonian neural network