Neural ODE

2 сентября 2022 г.

1 Задача

Имеется наблюдения некоторого процесса:

$$Z = (z_1, t_1), (z_2, t_2), ..., (z_N, t_N)$$

И система подчиняется ОДУ с неизвестной правой частью

$$\frac{dz}{dt} = f(z(t), t).$$

Хотим найти оценку $f(z, t, \theta)$.

2 Интерес к задаче

Решение данной задачи позволяет моделировать динамику системы даже если правая часть не имеет аналитического вида. Также метод позволяет получить непрерывный аналог решения. И в дальнейшем моделировать физичекие системы лучше(Lagrangian NN, Hamiltonial NN)

3 Методы решения

Метод решения один для обучения всех нейросетей: найти градиент функции потери по параметру θ. Но задача немного сложнее, так как функция потерь зависит от решалки ОДУ:

$$L(z(t_1)) = L(ODESolver(z_0, t, \theta, f)).$$

Для нахождения $dL/d\theta$ воспользуемся методом сопряженных состояний.

Lagrangian NN, Hamiltonian NN

4 Постановка задачи

Хотим решать задачу моделировать эволюцию физической системы во времени с какой-то физически сохраняющейся величиной(энергия) с помощью нахождения оценки гамильтониана/лагранжиана.

5 Интерес к задаче

Позволяет учесть физичность системы и манипулировать с сохраняющимися величинами. Плюсом лагранжевой сети является то, что для неё не нужны канонические координаты.

6 Методы

1) Для гамильтоновой сети будем находить оценку гамильтониана исходя из уравнений Гамильтона, т.е. минимизировать следующий функционал:

$$L_{HNN} = ||\frac{\partial H_\theta}{\partial \boldsymbol{p}} - \frac{\partial \boldsymbol{q}}{\partial t}||_2 + ||\frac{\partial H_\theta}{\partial \boldsymbol{q}} + \frac{\partial \boldsymbol{p}}{\partial t}||_2$$

2) Для лагранжевой придется переписать уравнение Эйлера-Лагранжа в векторном виде и минимизировать этот функционал

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} = \frac{\partial \mathcal{L}}{\partial q_j}$$

$$\frac{d}{dt}\nabla_{\dot{q}}\mathcal{L} = \nabla_q \mathcal{L}$$

$$(\nabla_{\dot{q}}\nabla_{\dot{q}}^\top \mathcal{L})\ddot{q} + (\nabla_q \nabla_{\dot{q}}^\top \mathcal{L})\dot{q} = \nabla_q \mathcal{L} \qquad \text{раскрывает производную по времени } \frac{d}{dt} \quad ($$

$$\ddot{q} = (\nabla_{\dot{q}}\nabla_{\dot{q}}^\top \mathcal{L})^{-1}[\nabla_q \mathcal{L} - (\nabla_q \nabla_{\dot{q}}^\top \mathcal{L})\dot{q}] \qquad \text{обращаем матрицу} \ddot{q} \quad ($$