Моделирование динамики физических систем с помощью Physics-Informed Neural Networks

А. А. Терентьев[1]
terentev.aa@phystech.edu
[1]ФПМИ, МФТИ

В работе решается задача выбора оптимальной модели в классе нейронных сетей для предсказания динамики физической системы. Использование классических нейросетевых методов для построения траекторий физических систем не позволяет получить решение, удовлетворяющее законам сохранения. Для их учёта используются различные модификации прямого решения с помощью нейронных сетей, которые бы описывали Лагранжеву динамику и опирались на теорему Нётер. В данной работе предлагается исследовать модели на основе Нётеровских сетей с более сложной архитектурой внутренних слоёв, а также рассмотреть адекватность моделирования систем, подвергающихся внешнему воздействию.

1 Введение

Одной из основных задач механики является задача моделирования динамики различных систем. Эта задача находит применение в различных областях: моделирование траектории, работа механизмов или систем, их реакции на внешние и внутренние воздействия, которые могут привести к нарушению работы или полному выводу их из строя.

Для того, чтобы использовать лагранжеву механику необходимо выбрать обобщенные координаты, которые полностью бы описывали поведения системы, иметь знания о Лагранжиане системы, который определяется как разница между кинетической энергией (T) и потенциальной энергией (V)

$$L = T - V$$

и величине обобщенных сил системы в случае внешних воздействий. Динамика системы в таком случае описывается системой уравнений Эйлера-Лагранжа:

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = Q_x = 0.$$

Мы не всегда можем иметь представление о лагранжиане в случае сложных систем.

Получить решение данной задачи с помощью классической Ньютоновской механики и аппроксимации дифференциальных уравнений нейронными сетями практически не представляется возможным для хаотических систем, таких как двойной маятник, т.к. такие сети не имеют знания о законах сохранения, импульса, момента импульса и других законах физики, и оказываются неспособными получить оптимальные параметры модели.

Одним из способов его нахождения являются Лагранжевы нейронные сети (LNN), которые аппроксимируют лагранжиан системы, после чего оказывается возможным возможным решить возникающие уравнения Эйлера-Лагранжа. Таким образом, нейронная сеть оказывается способной учесть физику системы, а точнее о закон сохранения энергии. Для некоторых систем, используя теорему Нётер и некоторых знаниях о симметриях Лагранжиана системы, могут быть получены более точные решения, учитывающие больше законов сохранения. Например, Нётеровские Лагранжевы нейронные сети (NLNN) способны учитывать трансляционную и вращательную симметрию лагранжиана и, следовательно, находят оптимальное решение для систем, в которых выполняются законы сохранения импульса и момента импульса.

На точность решения влияет не только количество учтенных интегралов движения, но и выбор самой модели, на основе которой находится соответствующая функция динамики системы. В работе исследуются различные архитектуры нейронных сетей в основе NLNN.

Для инженерных задач необходимо не только уметь предсказывать динамику самой системы, но и её реакцию на внешнее воздействие. Так, для гироскопических приборов критическим является их реакция на колебательные воздействия, которые могут приводить к погрешностям или полному выходу системы из строя, причем необходимо исследовать поведение прибора при разных эксплуатационных условиях.

Для вычислительного эксперимента соответственно взята модель гироскопа. В данной системе сохраняется энергия, импульс и момент импульса. Поведение системы исследовалось с помощью NLNN, в том числе при различных внешних воздействиях. Результаты моделирования сравнивались с траекториями, полученными аналитически с помощью метода Рунге-Кутты 4-ого порядка.

2 Постановка задачи регрессии динамики физической системы

Задачу моделирования динамики системы можно свести к задаче регрессии. Пусть дана выборка из m траекторий

$$\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^m,$$

где $\mathbf{x}_i = (\mathbf{q}_i, \dot{\mathbf{q}}_i)$ – координаты траектории движения двойного маятника, $\mathbf{y}_i = \dot{\mathbf{x}}_i = (\dot{\mathbf{q}}_i, \ddot{\mathbf{q}}_i)$ – динамика движения системы двойного маятника, $\mathbf{q}_i \in \mathbb{R}^{r \times n}$, где r – количество координат, n – длина траектории.

Регрессионная модель выбирается из класса нейронных сетей

$$\{\mathbf{f}_k \colon (\mathbf{w}, \mathbf{X}) \to \hat{\mathbf{y}} \mid k \in \mathcal{K}\},$$

где $\mathbf{w} \in \mathbb{W}$ – параметры модели, $\hat{\mathbf{y}} = \mathbf{f}(\mathbf{X}, \mathbf{w}) \in \mathbb{R}^{2 \times r \times n}$, $\mathbf{X} = \bigcup_{i=1}^m \mathbf{x}_i$.

В качестве функция ошибки взята квадратичная ошибка:

$$\mathcal{L}(\mathbf{y}, \mathbf{X}, \mathbf{w}) = \|\hat{\mathbf{y}} - \mathbf{y}\|_2^2.$$

Таким образом, задача моделирования динамики системы представлена в виде задачи минимизации квадратичной ошибки:

$$\mathbf{w}^* = \underset{\mathbf{w} \in \mathbb{W}}{\operatorname{argmin}} \big(\mathcal{L}(\mathbf{w})\big).$$

3 Вычислительный эксперимент

3.1 Цель эксперимента

Построить метод аппроксимации лагранжиана физической системы по одной или нескольким траекториям, которые учитывает известные свойства системы.

3.2 Набор данных для базового эксперимета

Обучение модели производилось на наборе траекторий двойного маятника представленного в работе [1]. Было синтезировано несколько траекторий для разных начальных параметров. Из синтезированных траектории были выбраны нетривиальные. На отобранных данных производилось обучение

3.3 Базаовая модель лагранжевых сетей

Лагранжев формализм моделирует физическую систему с координатами $x_t = (q, \dot{q})$, которая начинается в состоянии x_0 и заканчивается в другом состоянии x_1. Определяется функционал, который называется действием

$$S = \int_{t_0}^{t_1} L dt,$$

где L – лагранжиан системы. Действие определяет путь, по которому координаты x_t пройдут из x_0 в x_1 в промежуток времени от t_0 до t_1. Путь минимизирует действие S, т.е. $\delta S = 0$. Это приводит к уравнению Эйлера-Лагранжа, определяющему динамику системы

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} = \frac{\partial L}{\partial \mathbf{q}}$$

Ускорение каждой компоненты системы $\ddot{\mathbf{q}}$ может быть напрямую получено из данного уравнения:

$$\begin{aligned}
\frac{\partial}{\partial \dot{\mathbf{q}}}\frac{dL}{dt} &= \frac{\partial L}{\partial \mathbf{q}} \\
\frac{\partial}{\partial \dot{\mathbf{q}}}\left(\frac{\partial L}{\partial \mathbf{q}}\frac{dq}{dt} + \frac{\partial L}{\partial \dot{\mathbf{q}}}\frac{d\dot{\mathbf{q}}}{dt}\right) &= \frac{\partial L}{\partial \mathbf{q}} \\
\frac{\partial}{\partial \dot{\mathbf{q}}}\left(\frac{\partial L}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}}\ddot{\mathbf{q}}\right) &= \frac{\partial L}{\partial \mathbf{q}} \\
\frac{\partial}{\partial \dot{\mathbf{q}}}\frac{\partial L}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial}{\partial \dot{\mathbf{q}}}\frac{\partial L}{\partial \dot{\mathbf{q}}}\ddot{\mathbf{q}} &= \frac{\partial L}{\partial \mathbf{q}} \\
\frac{\partial}{\partial \dot{\mathbf{q}}}\frac{\partial L}{\partial \dot{\mathbf{q}}}\ddot{\mathbf{q}} &= \frac{\partial L}{\partial \mathbf{q}} - \frac{\partial}{\partial \dot{\mathbf{q}}}\frac{\partial L}{\partial \mathbf{q}}\dot{\mathbf{q}} \\
\ddot{\mathbf{q}} &= \left(\frac{\partial}{\partial \dot{\mathbf{q}}}\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)^{-1}\left[\frac{\partial L}{\partial \mathbf{q}} - \frac{\partial}{\partial \dot{\mathbf{q}}}\frac{\partial L}{\partial \mathbf{q}}\dot{\mathbf{q}}\right] \\
\ddot{\mathbf{q}} &= \left(\nabla_{\dot{\mathbf{q}}\dot{\mathbf{q}}} L\right)^{-1}\left[\nabla_q L - \left(\nabla_{\dot{\mathbf{q}}\mathbf{q}} L\right)\dot{\mathbf{q}}\right],
\end{aligned}$$

где гессиан $(\nabla_{\dot{\mathbf{q}}\mathbf{q}} L)_{ij} = \frac{\partial^2 L}{\partial q_j \partial \dot{q}_i}$.

Таким образом, алгоритм моделирования динамики системы в лагранжевой динамике:

1. Найти аналитические выражения для кинетической (T) и потенциальной энергии (V)
2. Получить лагранжиан $\mathcal{L} = T - V$
3. Применить ограничение Эйлера-Лагранжа $\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} = \frac{\partial L}{\partial \mathbf{q}}$
4. Решить получившуюся систему дифференциальных уравнений

3.4 Лагранжевы нейронные сети

В работе [2] предложено в нейронную сеть

$$f : \mathbf{X} = (\mathbf{q}, \dot{\mathbf{q}}) \to \mathbf{Y}$$

добавить априорные знания о физике системы, учитывая лагранжиан системы:

$$f : \mathbf{X} = (\mathbf{q}, \dot{\mathbf{q}}) \to \mathbf{L};$$

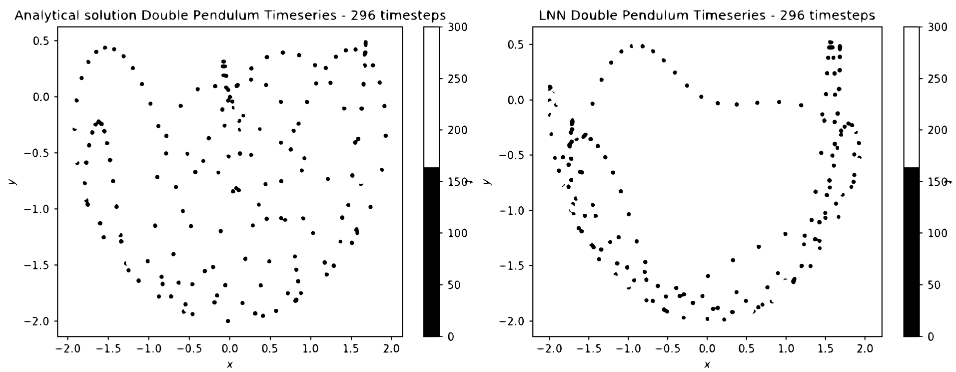
Т.е. ключевой идеей является параметризовать нейронной сетью лагранжиан L, получить выражение ограничения Эйлера-Лагранжа и обратно распространить ошибку через полученные ограничения

$$\ddot{\mathbf{q}} = (\nabla_{\dot{\mathbf{q}}\dot{\mathbf{q}}} L)^{-1} \left[\nabla_q L - (\nabla_{\dot{\mathbf{q}}\mathbf{q}} L)\, \dot{\mathbf{q}}\right],$$

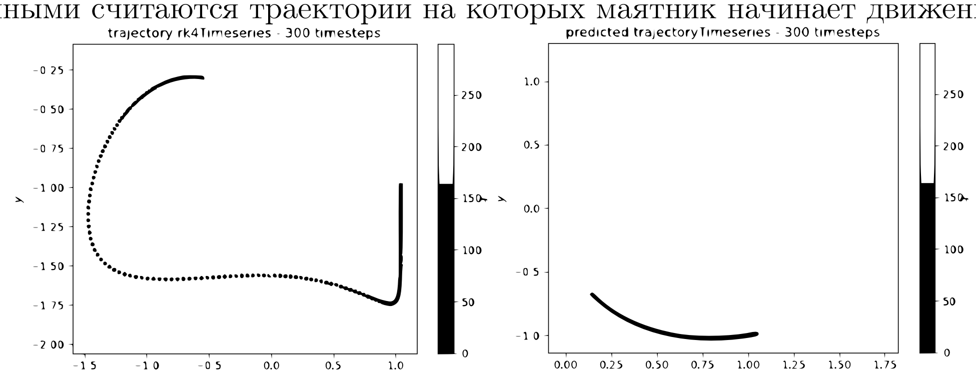
В качестве нейронной сети $f : \mathbf{X} = (\mathbf{q}, \dot{\mathbf{q}}) \to \mathbf{L}$ берется полносвязная сеть с 3-мя слоями. Таким образом, для заданных координат $(\mathbf{q}, \dot{\mathbf{q}})$ имеем модель с априорными знанями о законе сохранения энергии, которой можем получить динамику параметров $(\dot{\mathbf{q}}, \ddot{\mathbf{q}})$.

3.5 Результаты

Ниже приведены результы измерения работы модели на одной сгенерированной траектории



Перед применением алгоритмы была проведена нормализация данных, на отрезок [0, 2π]. Поэтому траектории, у которых угол больше считаются антифизичными. Также нефизичными считаются траектории на которых маятник начинает движение вверх.



Как видно из второй траектории присутствуют нефизичные траектории, что является проблемой данной модели и требует доработки

Литература

[1] *Северилов П. А.* Выбор оптимальной модели в задаче моделирования динамики физической системы // URL: `https://github.com/severilov/master-thesis/blob/main/doc/Severilov2022MasterThesis_rus.pdf`.

[2] *M. Cranmer, S. Greydanus, S. Hoyer et al.* Lagrangian neural networks // ArXiv preprint., 2020. doi: `http://dx.doi.org/Vol.abs/2003.04630`.

Поступила в редакцию