Федеральное государственное автономное образовательное учреждение
высшего образования
"Московский физико-технический институт
(национальный исследовательский университет)"
Физтех-школа прикладной математики и информатики
Кафедра интеллектуальных систем

Северилов Павел Андреевич

Выбор оптимальной модели в задаче моделирования динамики физической системы

03.04.01 — Прикладные математика и физика

Магистерская диссертация

Научный руководитель:
д. ф.-м. н.
Стрижов Вадим Викторович

Москва

2022 г.

Содержание

Аннотация

В работе решается задача выбора оптимальной модели предсказания динамики физической системы. Под динамикой системы понимается изменение во времени параметров системы. Нейронные сети не имеют априорных знаний о моделируемой системе, что не позволяет получить оптимальные параметры, учитывающие физические законы. Лагранжева нейронная сеть учитывает закон сохранения энергии при моделировании динамики. В данной работе предлагается Нётеровская агранжева нейронная сеть, учитывающая законы сохранения импульса и момента импульса в дополнение к закону сохранения энергии. Показано, что для данной задачи Нётеровская Лагранжева нейронная сеть является оптимальной среди полносвязной модели нейронной сети, нейронной сети с долговременной кратковременной памятью и Лагранжевой нейронной сетью. Сравнение моделирования проводилось на искусственно сгенерированных данных для системы двойного маятника, которая является простейшей хаотической системой. Результаты экспериментов подтверждают гипотезу, что внесение априорного знания о физике системы повышает качество модели.

1 Введение

Моделирование динамики физической системы является одной из основных задач механики [3, 1]. Под динамикой системы понимается изменение во времени параметров этой системы.

В классической механике, чтобы получить динамику системы, необходимо расписать все силы системы, уравнения сохранения импульса и энергии и численно решить полученные дифференциальные уравнения. Однако на практике моделирование динамики данным способом представляется затруднительным. Например, для хаотических систем: системы двойного маятника, системы трех тел.

Существующие модели решения обыкновенных дифференциальных уравнений с помощью нейронных сетей не имеют априорных знаний о моделируемой системе [6, 10]. Это не позволяет получить оптимальные параметры модели, которые учитывают физические законы [4, 7, 9].

Альтернативным классической механике подходом является лагранжева динамика [3], которая переформулирует проблему в терминах набора обобщенных координат, которые полностью описывают возможные движения частицы. Чтобы использовать лагранжеву динамику, необходимо построить лагранжиан L, который определяется как разница между кинетической энергией (T) и потенциальной энергией (V) системы:

$$L = T - V$$

Интеграл L по времени называется *действием.* Истинная траектория системы минимизирует этот интеграл (*принцип наименьшего действия*). Отсюда следует, что действие не может меняться при малых вариациях пути, что эквивалентно уравнению Эйлера-Лагранжа

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

Лагранжевы нейронные сети (LNN) [7, 9] основаны на лагранжевой динамике. Модель аппроксимирует лагранжиан системы, из которого с помощью уравнения Эйлера-Лагранжа получается динамика системы. Таким образом, LNN имет априорные знания о моделируемой системе в виде её лагранжиана.

В модели LNN лагранжиан независим от времени, т.е. учитывается закон сохранения энергии. Но не учитываются другие законы сохранения, как закон сохранения импульса и момента импульса.

Теорема Нетер [11] говорит о том, что любая непрерывная симметрия лагранжиана L связана с сохраняющейся величиной системы. Например, если L не меняется при трансляции (т. е. обладает трансляционной симметрией), то сохраняется импульс. Если L не меняется при поворотах, то сохраняется момент импульса. Таким образом, в терминах теоремы Нётер LNN имеет временную трансляционную симметрию, но не имеет трансляционной и вращательной симметрий.

В данной работе предлагается Нётеровская нейронная сеть, которая является модификацией Лагранжевой нейронной сети, в которой учитываются трансляционные и вращательные симметрии. Модель получает на вход разницу координат вместо самих координат и аппроксимирует потенциальную энергию системы. Лагранжиан системы в таком случае будет разностью известной кинетической энергии и аппроксимированной потенциальной энергией. Показано, что при такой постановке, лагранжиан системы будет инвариантен относительно трансляции и поворотов в пространстве.

Для вычислительного эксперимента была взята простейшая хаотическая система двойного маятника [1]. В данной системе сохраняется энергия, импульс и момент импульса [8]. Для моделирования динамики системы были взяты полносвязная нейронная сеть[12], нейронная сеть с долгой краткосрочной памятью (LSTM) [5], LNN [7] и Нётеровская LNN. Результаты моделирования сравнивались с траекториями маятника, полученными аналитическим решением с помощью метода Рунге-Кутты 4-ого порядка [2]. Показано, что качество моделирования динамики системы повышается при добавлении априорного знания о физике системы.

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:

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

3 Лагранжевы нейронные сети

3.1 Лагранжева динамика

Лагранжев формализм [7, 3, 8, 1] 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:

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

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

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

3.2 Лагранжевы нейронные сети

В работе [7] предложено в нейронную сеть

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

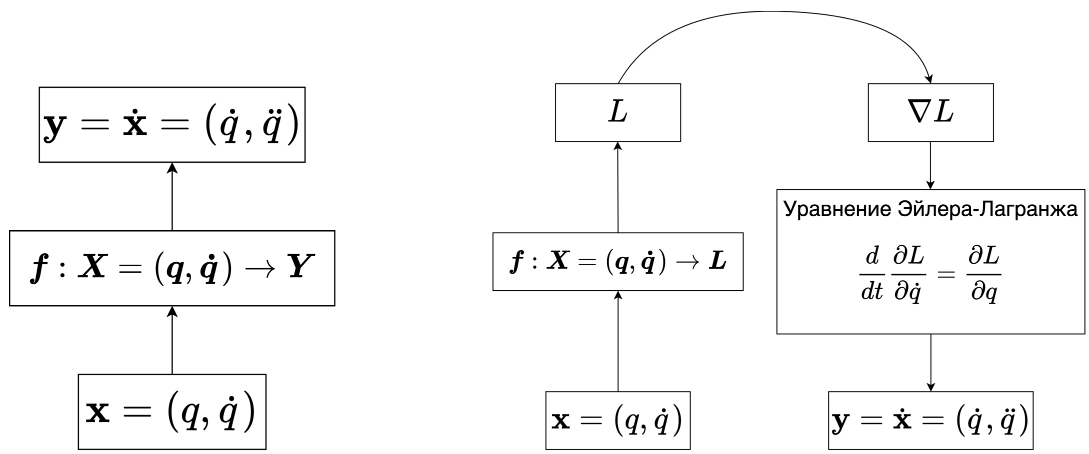
добавить априорные знания о физике системы, учитывая лагранжиан системы:

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

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

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

На рисунке 1 представлены схемы работы базового решения нейронными сетями и LNN для задачи моделирования динамики системы.На основе базового решения основаны работы [10, 6].



(a) Базовое решение нейронными сетями

(b) Решение LNN

Рис. 1: Схемы работы базового решения нейронными сетями (a) и LNN (b) для задачи моделирования динамики физической системы

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 $(\dot{\mathbf{q}}, \ddot{\mathbf{q}})$.

4 Нётеровская Лагранжева нейронная сеть

Лагранжева нейронная сеть не обладает трансляционной и вращательной симметриями в пространстве. Предлагается *Нётеровская Лагранжева нейронная сеть (LNN)*, которая получает на вход разницу между координатами элементов системы

$$\Delta q_{ij} = \sqrt{(\vec{q_i} - \vec{q_j}) \cdot (\vec{q_i} - \vec{q_j})}$$

и аппроксимирует потенциальную энергию системы $\boldsymbol{V}(\Delta \boldsymbol{q})$:

$$f : \boldsymbol{X} = (\boldsymbol{\Delta q}) \to \boldsymbol{V},$$

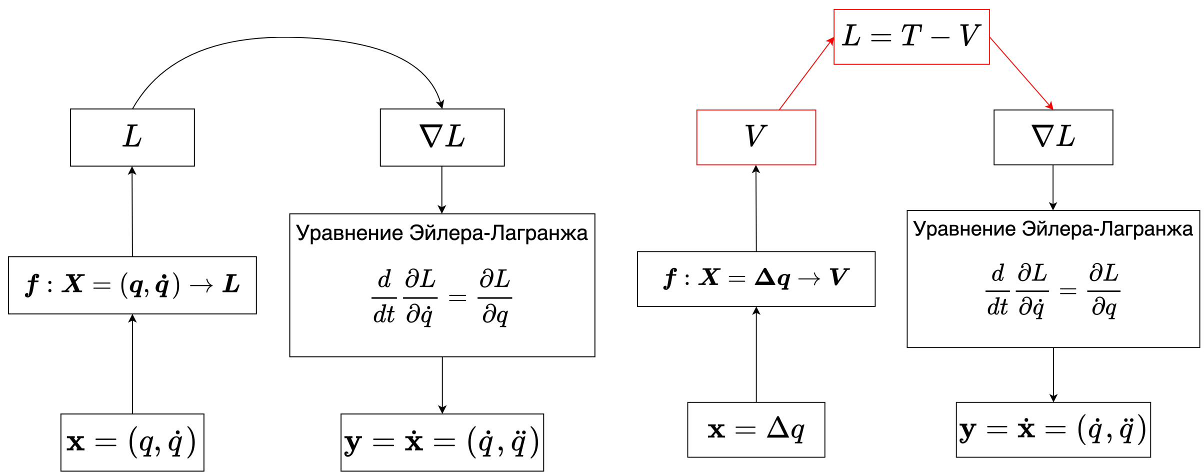
В общем случае, Лагранжиан системы, получаемый Нётеровской LNN представим в виде

$$L(\boldsymbol{\Delta}\mathbf{q}, \dot{\mathbf{q}}) = T(\boldsymbol{\Delta}\mathbf{q}, \dot{\mathbf{q}}) - V(\boldsymbol{\Delta}\mathbf{q})$$

В общем виде, кинетическая энергия системы может быть получена аналитически для большинства систем:

$$T(\mathbf{\Delta q}, \dot{\mathbf{q}}) = \sum_{i=1}^{r} \frac{1}{2} m_i \dot{\tilde{q}}_i^2$$

На рисунке 2 представлены схемы работы LNN и Нётеровской LNN для задачи моделирования динамики системы. Красным цветом наглядно показано отличие LNN от Нётеровской модификации: вместо лагранжиана параметризуется потенциальная энергия.



(a) Схема работы LNN (b) Схема работы Нётеровской LNN

Рис. 2: Схемы работы LNN (a) и Нётеровской LNN (b) для задачи моделирования динамики физической системы

Покажем, что данная модификация LNN сохраняет импульс и момент импульса системы двойного маятника.

Теорема 1. *(Северилов, 2022) Нётеровская LNN учитывает трансляционную симметрию (закон сохранения импульса). Т.е. лагранжиан системы не изменится, если подвергнуть элементы системы перемещению* $\boldsymbol{\xi}$*:*

$$L(\mathbf{\Delta\tilde{q}}, \dot{\tilde{\mathbf{q}}}) = L(\mathbf{\Delta q}, \dot{\mathbf{q}}), \quad \tilde{\mathbf{q}} = \mathbf{q} + \boldsymbol{\xi}$$

Доказательство.

Допустим грузы маятника подверглись перемещению $\boldsymbol{\xi}$, т.е. обновленные координаты сисетемы примут вид:

$$\tilde{\mathbf{q}} = \mathbf{q} + \boldsymbol{\xi} = \begin{bmatrix} \vec{q}_1 + \vec{\xi} \\ \dots \\ \vec{q}_r + \vec{\xi} \end{bmatrix},$$

где r – число координат системы.

Тогда, учитывая, что $\dot{\tilde{\mathbf{q}}} = \dot{\mathbf{q}}$, лагранжиан обновленной системы:

$$L(\boldsymbol{\Delta}\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) = L(\boldsymbol{\Delta}\tilde{\mathbf{q}}, \dot{\mathbf{q}}) = T(\boldsymbol{\Delta}\tilde{\mathbf{q}}, \dot{\mathbf{q}}) - V(\boldsymbol{\Delta}\tilde{\mathbf{q}})$$

Таким образом, чтобы показать инвариантность лагранжиана относительно перемещения, необходимо показать, что

$$T(\boldsymbol{\Delta}\tilde{\mathbf{q}}, \dot{\mathbf{q}}) = T(\boldsymbol{\Delta}\mathbf{q}, \dot{\mathbf{q}}), \quad V(\boldsymbol{\Delta}\tilde{\mathbf{q}}) = V(\boldsymbol{\Delta}\mathbf{q})$$

Равенства будут выполнены, если $\boldsymbol{\Delta}\tilde{\mathbf{q}} = \boldsymbol{\Delta}\mathbf{q}$:

$$\begin{aligned} \Delta\tilde{q}_{ij} &= \sqrt{\left[\left(\vec{q}_i + \vec{\xi}\right) - \left(\vec{q}_j + \vec{\xi}\right)\right] \cdot \left[\left(\vec{q}_i + \vec{\xi}\right) - \left(\vec{q}_j + \vec{\xi}\right)\right])} \\ &= \sqrt{(\vec{q}_i - \vec{q}_j) \cdot (\vec{q}_i - \vec{q}_j)} \\ &= \Delta q_{ij} \end{aligned}$$

Получили, что $L(\boldsymbol{\Delta}\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) = L(\boldsymbol{\Delta}\mathbf{q}, \dot{\mathbf{q}})$, т.е. полученный лагранжиан инвариантен относительно перемещения координат – имеет трансляционную симметрию. Тогда по теореме Нётер из трансляционной симметрии следует закон сохранения импульса.

■

Теорема 2. *(Северилов, 2022) Нётеровская LNN учитывает вращательную симметрию (закон сохранения момента импульса). Т.е. лагранжиан системы не изменится, если подвергнуть элементы системы повороту* $\mathbf{Q}$*:*

$$L(\boldsymbol{\Delta}\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) = L(\boldsymbol{\Delta}\mathbf{q}, \dot{\mathbf{q}}), \quad \tilde{\mathbf{q}} = \mathbf{Q}\mathbf{q}$$

Доказательство.

Допустим, элементы системы подверглись повороту матрицей поворота $\mathbf{Q}$, т.е. обновленные координаты сисетемы примут вид $\tilde{\mathbf{q}} = \mathbf{Q}\mathbf{q}$. Аналогично теореме 1, достаточно показать, что

$$T(\boldsymbol{\Delta}\tilde{\mathbf{q}}, \dot{\mathbf{q}}) = T(\boldsymbol{\Delta}\mathbf{q}, \dot{\mathbf{q}}), \quad V(\boldsymbol{\Delta}\tilde{\mathbf{q}}) = V(\boldsymbol{\Delta}\mathbf{q}),$$

т.е., что $\mathbf{\Delta\tilde{q}} = \mathbf{\Delta q}$.

Т.к. $\mathbf{Q}$ является матрицей поворота, то $\mathbf{Q}$ – ортогональная, т.е. $\mathbf{Q}\mathbf{Q}^\mathsf{T} = \mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}$. Тогда:

$$\begin{aligned}
\Delta\tilde{q}_{ij} &= \sqrt{(\mathbf{Q}\vec{q_i} - \mathbf{Q}\vec{q_j}) \cdot (\mathbf{Q}\vec{q_i} - \mathbf{Q}\vec{q_j})} \\
&= \sqrt{\mathbf{Q}\left(\vec{q_i} - \vec{q_j}\right) \cdot \mathbf{Q}\left(\vec{q_i} - \vec{q_j}\right)} \\
&= \sqrt{\mathbf{Q}\mathbf{Q}^T \left(\vec{q_i} - \vec{q_j}\right) \cdot \left(\vec{q_i} - \vec{q_j}\right)} \\
&= \sqrt{(\vec{q_i} - \vec{q_j}) \cdot (\vec{q_i} - \vec{q_j})} \\
&= \Delta q_{ij}
\end{aligned}$$

Получили, что $L(\mathbf{\Delta\tilde{q}}, \dot{\tilde{\mathbf{q}}}) = L(\mathbf{\Delta q}, \dot{\mathbf{q}})$, т.е. полученный лагранжиан инвариантен относительно поворота координат – имеет вращательную симметрию. Тогда по теореме Нётер из вращательной симметрии следует закон сохранения момента импульса. ■

В качестве нейронной сети $f : \mathbf{X} = (\mathbf{\Delta q}) \to \mathbf{V}$ аналогично LNN 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, импульса и момента импульса, которой можем получить динамику параметров $(\dot{\mathbf{q}}, \ddot{\mathbf{q}})$.

5 Система двойного маятника

В качестве моделируемой физической системы взята система двойного маятника (Рис. 3). Двойной маятник образуется путем присоединения одного маятника к другому. Каждый маятник состоит из груза, соединенного с безмассовым жестким стержнем, который может двигаться только в вертикальной плоскости. Ось первого маятника закреплена в точке O. Все движения без трения.

В данном случае координатами $\mathbf{q}$ являются углы между стержнями маятника и вертикальной осью

$$\mathbf{q} = (\theta_1, \theta_2)$$

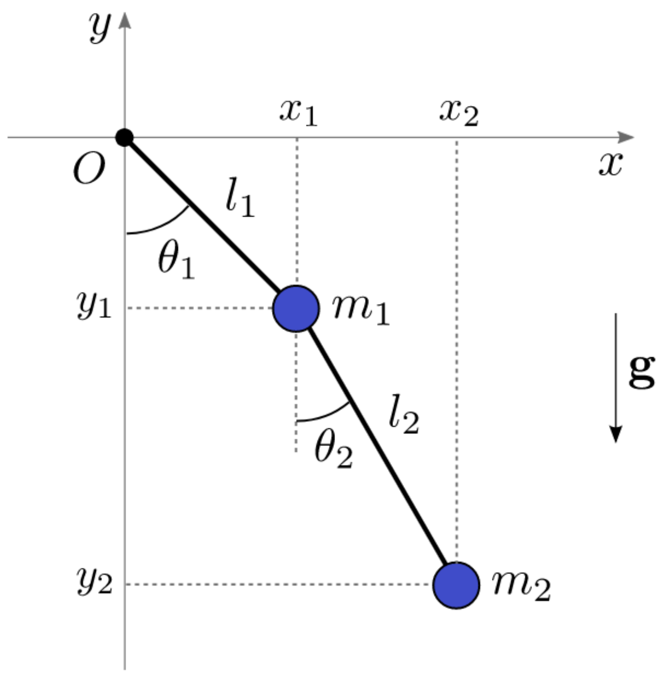


Рис. 3: Схема физической системы двойного маятника

5.1 Лагранжиан системы

Примем точку O за начало декартовой системы координат с осью x, направленной вдоль горизонтального направления, и осью y, направленной вертикально вверх. Пусть θ_1 и θ_2 – углы, которые первый и второй стержни образуют с вертикальным направлением соответственно. Как видно на рисунке 3, положения грузов задаются следующим образом:

$$
\begin{aligned}
x_1 &= l_1 \sin\theta_1, & y_1 &= -l_1 \cos\theta_1, \\
x_2 &= l_1 \sin\theta_1 + l_2 \sin\theta_2, & y_2 &= -l_1 \cos\theta_1 - l_2 \cos\theta_2
\end{aligned}
$$

Дифференцируя приведенные выше величины по времени, получаем скорости грузов:

$$
\begin{aligned}
\dot{x}_1 &= l_1 \dot{\theta}_1 \cos\theta_1 & \dot{y}_1 &= l_1 \dot{\theta}_1 \sin\theta_1 \\
\dot{x}_2 &= l_1 \dot{\theta}_1 \cos\theta_1 + l_2 \dot{\theta}_2 \cos\theta_2 & \dot{y}_2 &= l_1 \dot{\theta}_1 \sin\theta_1 + l_2 \dot{\theta}_2 \sin\theta_2
\end{aligned}
$$

Данная система имеет лагранжиан:

$$
\begin{aligned}
L = \frac{1}{2}\left(m_1 + m_2\right) l_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 l_2^2 \dot{\theta}_2^2 + m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos\left(\theta_1 - \theta_2\right) \\
+ \left(m_1 + m_2\right) g l_1 \cos\theta_1 + m_2 g l_2 \cos\theta_2
\end{aligned}
$$

Лагранжиан для двойного маятника определяется выражением $L = T - V$, где T и V – кинетическая и потенциальная энергии системы соответственно. Кинетическая энергия T определяется выражением:

$$\begin{aligned} T &= \frac{1}{2}m_1 v_1^2 + \frac{1}{2}m_2 v_2^2 \\ &= \frac{1}{2}m_1\left(\dot{x}_1^2 + \dot{y}_1^2\right) + \frac{1}{2}m_2\left(\dot{x}_2^2 + \dot{y}_2^2\right) \\ &= \frac{1}{2}m_1 l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2\left[l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1 l_2\dot{\theta}_1\dot{\theta}_2\cos\left(\theta_1 - \theta_2\right)\right] \end{aligned}$$

В выражении выше использован факт, что $\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 = \cos\left(\theta_1 - \theta_2\right)$. Потенциальная энергия V определяется выражением:

$$\begin{aligned} V &= m_1 g y_1 + m_2 g y_2 \\ &= -m_1 g l_1\cos\theta_1 - m_2 g\left(l_1\cos\theta_1 + l_2\cos\theta_2\right) \\ &= -\left(m_1 + m_2\right) g l_1\cos\theta_1 - m_2 g l_2\cos\theta_2 \end{aligned}$$

Таким образом, Лагранжиан системы:

$$\begin{aligned} L = \frac{1}{2}\left(m_1 + m_2\right) l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2 l_2^2\dot{\theta}_2^2 + m_2 l_1 l_2\dot{\theta}_1\dot{\theta}_2\cos\left(\theta_1 - \theta_2\right) \\ + \left(m_1 + m_2\right) g l_1\cos\theta_1 + m_2 g l_2\cos\theta_2 \end{aligned}$$

5.2 Аналитическое решение получения динамики системы

После подставления полученного лагранжиана в уравнение Эйлера-Лагранжа выражение сводится к системе дифференциальных уравнений [14]

$$\frac{d}{dt}\begin{pmatrix}\theta_1\\ \theta_2\\ \dot{\theta}_1\\ \dot{\theta}_2\end{pmatrix} = \begin{pmatrix}\omega_1\\ \omega_2\\ g_1\left(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2\right)\\ g_2\left(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2\right)\end{pmatrix}$$

где $g_1 = \frac{f_1 - \alpha_1 f_2}{1 - \alpha_1\alpha_2}$, $g_2 = \frac{-\alpha_2 f_1 + f_2}{1 - \alpha_1\alpha_2}$, $\alpha_i = \alpha_i\left(\theta_1, \theta_2\right)$ и $f_i = f_i\left(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2\right)$ для $i = 1, 2$

$$\begin{aligned} l\alpha_1\left(\theta_1, \theta_2\right) &= \frac{l_2}{l_1}\left(\frac{m_2}{m_1 + m_2}\right) \\ \alpha_2\left(\theta_1, \theta_2\right) &= \frac{l_1}{l_2}\cos\left(\theta_1 - \theta_2\right) \\ f_1\left(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2\right) &= -\frac{l_2}{l_1}\left(\frac{m_2}{m_1 + m_2}\right)\dot{\theta}_2^2\sin\left(\theta_1 - \theta_2\right) - \frac{g}{l_1}\sin\theta_1 \\ f_2\left(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2\right) &= \frac{l_1}{l_2}\dot{\theta}_1^2\sin\left(\theta_1 - \theta_2\right) - \frac{g}{l_2}\sin\theta_2 \end{aligned}$$

Полученная система дифференциальных уравнений может быть решена численно с помощью метода Рунге-Кутты 4 порядка.

5.3 Данные

Для обучения моделей были сгенерированы 20 траекторий движениями двойного маятника из разных начальных состояний $\mathbf{x}_0$ длиной 1500 моментов времени по 0.01 секунде. Длины стержней и массы грузов приняты равными 1. Генерация данных производилась на основе решения системы дифференциальных уравнений, описывающих динамику системы, с помощью метода рунге-Кутты 4 порядка.

На рисунке 4 представлена визуализация сгнерированных координат θ_1, θ_2 для одной из траекторий. Цвет означает угол наклона нижнего стержня маятника: зеленый, если угол $\theta_2 > 0$, синий, если $\theta_2 < 0$.

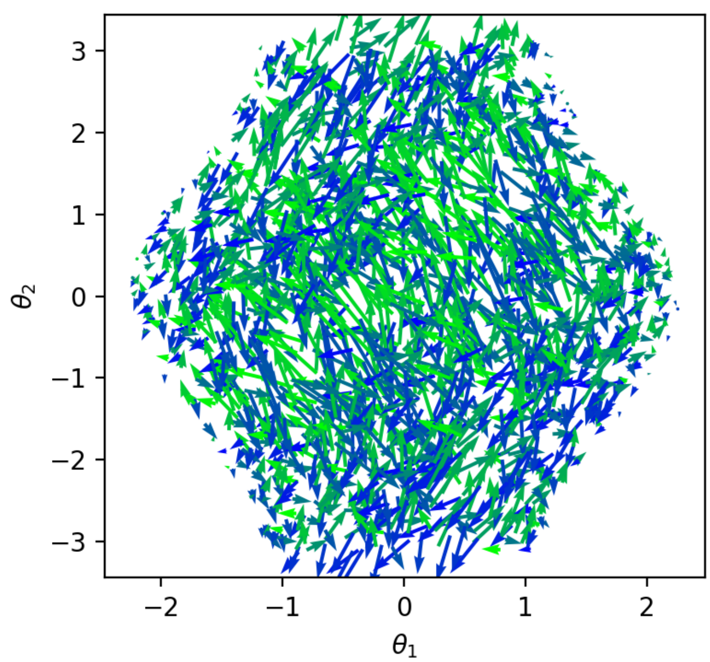


Рис. 4: Визуализация данных координат θ_1, θ_2 системы двойного маятника

6 Вычислительный эксперимент

В рамках вычислительного эксперимента написан программный комплекс для решения поставленных задач [13].

6.1 Исследуемые модели нейронных сетей

В качестве исследуемых моделей для сравнения моделирования динамики системы были использованы полносвязная нейронная сеть (FC), сеть с долговременной кратковременной памятью (LSTM), Лагранжева нейронная сеть и Нётеровская LNN, учитывающая трансляционную и вращательную симметрии. В качестве истинной динамики системы взята динамика, полученная аналитическим решением методом Рунге-Кутты 4 порядка.

В таблице **??** представлены сравниваемые модели. FC является простейшей нейронной сетью с малым количеством параметров. LSTM имеет значительно большее количество параметров, чем полносвязный сеть, но больше подходит для работы с последовательностями (в нашем случае траекториями) [5]. LNN является полносвязной сетью, но имеет априорное знание о законе сохранения энергии. Нётеровская LNN тоже является полносвязной сетью, но имеющей априорное знание о трех законах сохранения: энергии, имупльса, момента импульса.

Аналитическое решение	Метод Рунге-Кутты 4 порядка
Полносвязная нейронная сеть (FC)	$f : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{Y}$, $f = \sigma(...\boldsymbol{W}^T \sigma(\boldsymbol{W}^T \boldsymbol{x} + \boldsymbol{b}))$
Нейронная сеть с долгой краткосрочной памятью (LSTM)	$h : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{Y}$, $f_t = \sigma_g\left(W_f x_t + U_f h_{t-1} + b_f\right)$ $i_t = \sigma_g\left(W_i x_t + U_i h_{t-1} + b_i\right)$ $o_t = \sigma_g\left(W_o x_t + U_o h_{t-1} + b_o\right)$ $c_t = f_t \circ c_{t-1} + i_t \circ \sigma_c\left(W_c x_t + U_c h_{t-1} + b_c\right)$ $h_t = o_t \circ \sigma_h\left(c_t\right)$
LNN	$f : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{L}$, $\ddot{q} = \left(\nabla_{\dot{q}} \nabla_{\dot{q}}^{\top} L\right)^{-1} \left[\nabla_q L - \left(\nabla_q \nabla_{\dot{q}}^{\top} L\right) \dot{q}\right]$
Нётеровская LNN	$f : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{V}$, $L = T - V$, $\ddot{q} = \left(\nabla_{\dot{q}} \nabla_{\dot{q}}^{\top} L\right)^{-1} \left[\nabla_q L - \left(\nabla_q \nabla_{\dot{q}}^{\top} L\right) \dot{q}\right]$

Таблица 1: Сравниваемые методы моделирования динамики системы двойного маятника

6.2 Результаты

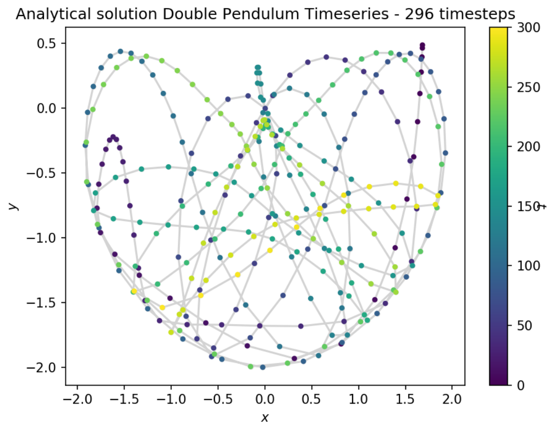
Результаты моделирования динамики системы двойного маятника различными видами нейронных сетей представлены в таблице **??**. Выборка состоит из 10 траекторий, сгенерированных на основе решения методом Рунге-Кутты 4 порядка.

FC	1.57 ± 0.53
LSTM	2.42 ± 0.79
LNN	1.32 ± 0.91
Нётеровская LNN	**1.28** ± 0.66

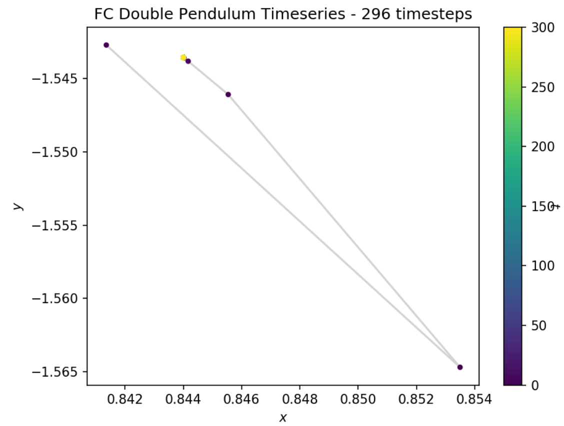
Таблица 2: Средняя ошибка RMSE между предсказанной динамикой системы нейронной сетью и динамикой системы, полученной аналитическим решением.

Сравнивания результаты LNN и FC, LSTM, можем сделать вывод, что модель, имеющая априорные знания о физике системы имеет выше точность моделирования динамики системы. Сравнивая результаты LNN и Нётеровской LNN, можем сделать вывод, что добавление дополнительных априорных знаний о физике системы повышает точность моделирования динамики системы

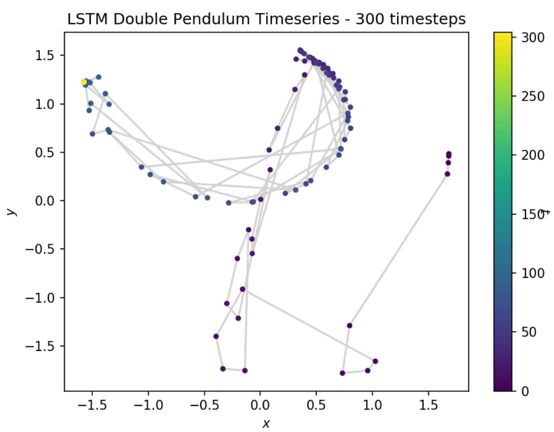
На рис 5 представлены траектории движения нижнего груза двойного маятника, полученные аналитическим решением и рассматриваемыми нейронными сетями. Динамика, предсказанная моделями LNN и Нётеровской LNN, наиболее близка к динамике, полученной аналитическим решением. Это подтверждает гипотезу, что добавление априорных знаний о лагранжиане системы повышает точность предсказания динамики системы (траектории движения маятника).



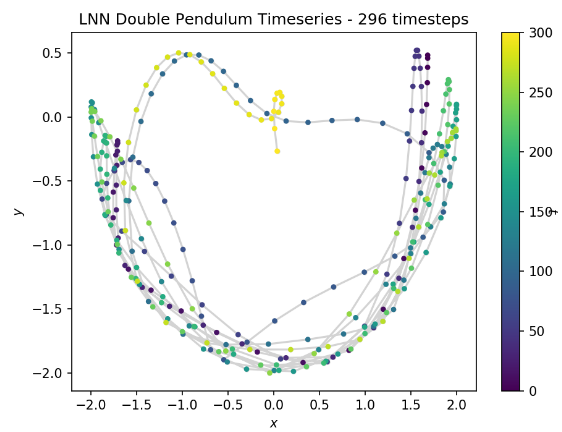
(a) Аналитическое решение



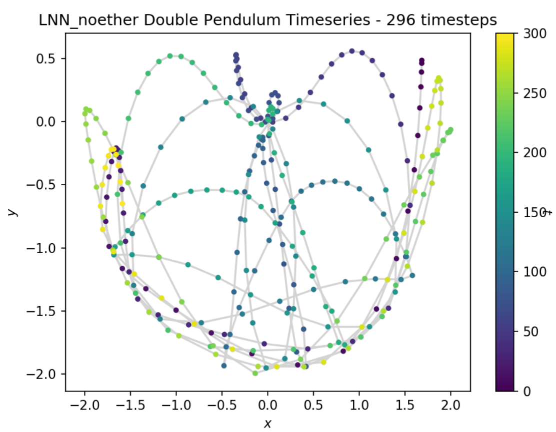
(b) FC



(c) LSTM



(d) LNN



(e) Нётеровская LNN

Рис. 5: Моделирование динамики системы двойного маятника различными видами нейронных сетей: аналитическое решение, полносвязная нейронная сеть, LSTM, LNN и Нётеровская LNN.

7 Заключение

В данной работе предложена Нётеровская LNN, которая в дополнение к закону сохранения энергии учитывает закон сохранения импульса и момента импульса. Также были доказаны теоремы, подтверждающие корректность учета закона сохранения импульса и момента импульса. На основе экспериментов по моделированию динамики системы двойного маятника показано, что добавление априорных знаний о физике системы повышает точность моделирования динамики физической системы.

Данные результаты показывают адекватность использования моделей с априорными знаниями о физике системы и открывает новые пути развития данных моделей. В качестве дальнейших исследований предлагается использование более сложных моделей нейронных сетей в основе LNN: использование реккурентных, сверточных сетей вместо только полносвязных.

Список литературы

[1] *Arnold V. I.* Problem in Mathematical Methods of Classical Mechanics, 2nd ed. — New York: Springer-Verlag, 1989.

[2] *Atkinson K. A.* An Introduction to Numerical Analysis (2nd ed.). — New York: John Wiley and Sons, 1989.

[3] *Goldstein H.* Classical Mechanics. — Addison-Wesley, 1980.

[4] *Greydanus S., Dzamba M., Yosinski J.* Hamiltonian neural networks. — 2019. — Vol. 32. https://proceedings.neurips.cc/paper/2019/file/26cd8ecadce0d4efd6cc8a8725cbd1f8-Paper.pdf.

[5] *Hochreiter S., Schmidhuber J.* Long short-term memory // *Neural computation.* — 1997. — 12. — Vol. 9. — Pp. 1735–80.

[6] *Lagaris I., Likas A., Fotiadis D.* Artificial neural networks for solving ordinary and partial differential equations // *IEEE Transactions on Neural Networks.* — 1998. — Vol. 9, no. 5. — Pp. 987–1000.

[7] Lagrangian neural networks / M. Cranmer, S. Greydanus, S. Hoyer et al. // *ArXiv preprint.* — 2020. — Vol. abs/2003.04630.

[8] *Leonard Susskind G. H.* Classical Mechanics: The Theoretical Minimum. — Penguin Books, 2014.

[9] *Lutter M., Listmann K., Peters J.* Deep lagrangian networks for end-to-end learning of energy-based control for under-actuated systems. — 2019. — 11.

[10] Neural ordinary differential equations / R. T. Q. Chen, Y. Rubanova, J. Bettencourt, D. K. Duvenaud // Advances in Neural Information Processing Systems / Ed. by S. Bengio, H. Wallach, H. Larochelle et al. — Vol. 31. — Curran Associates, Inc., 2018. https://proceedings.neurips.cc/paper/2018/file/69386f6bb1dfed68692a24c8686939b9-Paper.pdf.

[11] *Noether E.* Invariant variation problems // *Transport Theory and Statistical Physics.* — 1971. — Vol. 1, no. 3. — Pp. 186–207. https://doi.org/10.1080/00411457108231446.

[12] *Schmidhuber J.* Deep learning in neural networks: An overview // *Neural Networks.* — 2015. — Vol. 61. — Pp. 85–117. `https://www.sciencedirect.com/science/article/pii/S0893608014002135`.

[13] *Severilov P.* Код экспериментов для Нётеровской lnn. `https://github.com/severilov/master-thesis`.

[14] *Svirin A.* Double pendulum article at math24.net. `https://math24.net/double-pendulum.html`.