

Выбор оптимальной модели в задаче моделирования динамики физической системы нейронными сетями*

П. А. Северилов¹, В. В. Стрижов²

Аннотация: Лагранжевы нейронные сети (LNN) были разработаны для получения динамики ограниченного набора физических систем, что делает этот подход относительно узким. В данной работе рассматривается в терминах симметрий теорем Нётер связь моделей LNN с нейронными сетями более общего назначения с долговременной кратковременной памятью (LSTM), которая отлично подходит для прогнозирования последовательных задач, с базовой полносвязной модели нейронной сети (FC). Эксперименты сравнения проводились на задаче моделирования физической системы двойного маятника. В работе показано, что LNN является оптимальной среди моделей FN, LSTM для данной задачи. Также была представлена Нётеровская LNN, учитывающая дополнительные трансляционную и вращательную симметрии. Показано, что более интерпретируемая модель дает более точные результаты для решения задачи моделирования динамики физической системы

Ключевые слова: Лагранжева нейронная сеть, динамика физической системы, Теорема Нётер, симметрия

1 Введение

В классической механике, чтобы получить динамику физической системы, необходимо расписать все силы системы, уравнения сохранения импульса и энергии. Однако

*по

¹Московский физико-технический институт, severilov.pa@phystech.edu

²Вычислительный центр имени А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, Московский физико-технический институт, strijov@phystech.edu

на практике моделирование динамики систем данным способом представляется затруднительным. Например, для системы двойного маятника.

Альтернативным подходом является лагранжева динамика, 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 (U) системы:

$$L = T - U$$

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

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

Теорема Нетер говорит о том, что любая непрерывная симметрия 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 сдвигах во времени, то сохраняется энергия.

Получение динамики физической системы с помощью нейронных сетей обычно не учитывают законы сохранения и симметрию Нетер [1, 1, 1]. Простейшие подходы предсказывают ускорение, получая на вход координаты и скорости системы. Лагранжевы нейронные сети (LNN) используют альтернативный метод: предсказывается лагранжиан системы, из него дифференцированием получают динамику системы. Уравнение Эйлера-Лагранжа сохраняют энергию с течением времени, поэтому эта симметрия встроена в нейронную сеть LNN.

2 Связанные работы

3 Лагранжевы нейронные сети

3.1 Лагранжева механика

- **Проблемы HNN:** гамильтонов формализм требует, чтобы координаты системы были «каноническими» ((\mathbf{q}, \mathbf{p}) должны подчиняться соотношениям, заданным скобками Пуассона)
- **Решение проблемы:** использовать лагранжианы систем (обеспечивают сохранение полной энергии, могут делать это с использованием произвольных координат)

Моделирование динамики системы с помощью лагранжиана

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

3.2 Лагранжевы нейронные сети

3.3 DeLaN

Работа, в котором авторы рассматривают моделирование определенных типов лагранжевых систем. Предполагается, что кинетическая энергия является $T = \dot{q}^T M \dot{q}$, где M — q -зависимая положительно определенная матрица.

Этот подход хорошо работает для динамики твердого тела, которая включает в себя множество систем, встречающихся в робототехнике. Однако многие системы не обладают такой кинетической энергией: заряженная частица в магнитном поле, быстро движущийся объект в СТО.

3.4 LNN

Лагранжевы нейронные сети (LNN) Ключевой идеей является параметризовать нейронной сетью лагранжиан \mathcal{L} , получить выражение ограничения Эйлера-Лагранжа, обратно распространить ошибку через полученные ограничения.

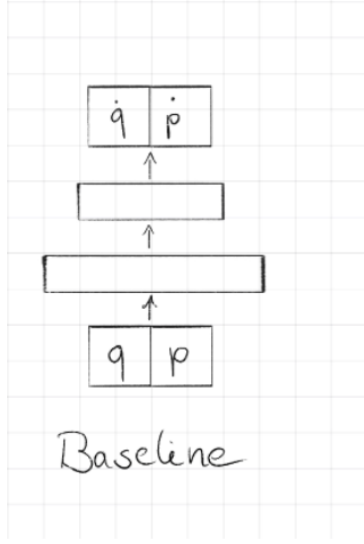
Получение динамики системы из предсказанного лагранжиана

$$\begin{aligned} \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_j} &= \frac{\partial \mathcal{L}}{\partial q_j} \Rightarrow \frac{d}{dt} \nabla_{\dot{q}} \mathcal{L} = \nabla_q \mathcal{L} \\ (\nabla_{\dot{q}} \nabla_{\dot{q}}^\top \mathcal{L}) \ddot{q} + (\nabla_q \nabla_{\dot{q}}^\top \mathcal{L}) \dot{q} &= \nabla_q \mathcal{L} \\ \ddot{q} &= (\nabla_{\dot{q}} \nabla_{\dot{q}}^\top \mathcal{L})^{-1} [\nabla_q \mathcal{L} - (\nabla_q \nabla_{\dot{q}}^\top \mathcal{L}) \dot{q}] \end{aligned}$$

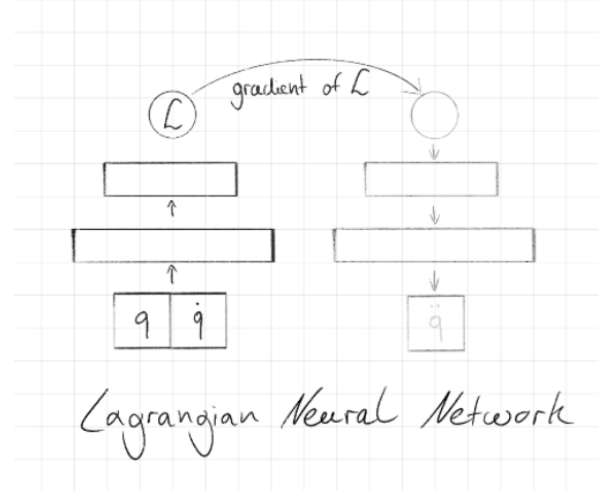
Таким образом, для заданного набора координат $x_t = (q_t, \dot{q}_t)$ получили метод вычисления $\dot{x}_t = (\dot{q}_t, \ddot{q}_t)$ из параметризованного лагранжиана.

Исходя из вышесказанного получаем **функцию ошибки**:

$$\mathcal{L} = \left\| \dot{x}_t^{\mathcal{L}_\theta} - \dot{x}_t^{true} \right\|_2$$



(a) Схема работы базового решения моделирования динамики физической системы нейронными сетями



(b) Схема работы Lagrangian Neural Networks (LNN) моделирования динамики физической системы

4 Теорема Нётер

Определение 1 (Генераторы и инвариантность). Пусть $L(t, x, \dot{x})$ – лагранжиан. Рассматривается преобразование

$$\begin{aligned} t &\mapsto t' = T(t, \mathbf{x}, \epsilon) \\ \mathbf{x} &\mapsto \mathbf{x}' = X(t, \mathbf{x}, \epsilon) \end{aligned} \quad (1)$$

где преобразование является гладким как функция от ϵ и при $\epsilon = 0$ является тождественным и. Пусть

$$\begin{aligned} T(t, \mathbf{x}, \epsilon) &= t + \tau(t, \mathbf{x})\epsilon + O(\epsilon^2) \\ X(t, \mathbf{x}, \epsilon) &= \mathbf{x} + \xi(t, \mathbf{x})\epsilon + O(\epsilon^2) \end{aligned} \quad (2)$$

Члены $\tau(t, x)$ и $\xi(t, x)$ называются генераторами преобразования. Генератор оставляет Лагранжиан инвариантным, если

$$\mathcal{L}(t', \mathbf{x}', \dot{x}') \frac{dt'}{dt} - \mathcal{L}(t, \mathbf{x}, \dot{x}) = O(\epsilon^2) \quad (3)$$

Теорема 1 (Закон сохранения Нётер). Пусть $\mathcal{L}(t, \mathbf{x}, \dot{x})$ – лагранжиан, инвариантный относительно генераторов $\tau(t, \mathbf{x})$ и $\xi(t, \mathbf{x})$. Пусть $\mathbf{x}(t)$ – экстремальная функция $J[\mathbf{x}] = \int \mathcal{L}(t, \mathbf{x}, \dot{x}) dt$. Тогда

$$\left\langle \frac{\partial \mathcal{L}}{\partial \dot{x}}, \xi \right\rangle + \left(\mathcal{L} - \left\langle \dot{x}, \frac{\partial \mathcal{L}}{\partial \dot{x}} \right\rangle \right) \tau$$

сохраняется по $\mathbf{x}(t)$.

5 Вычислительный эксперимент

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

5.1 Постановка задачи

Пусть дана выборка (\mathbf{X}, \mathbf{Y}) , где $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^\top \in \mathbb{R}^{n \times m}$ — матрица независимых переменных, $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_n]^\top \in \mathbb{R}^{n \times k}$ — матрица целевых переменных.

5.2 Данные: система двойного маятника

В качестве моделируемой физической системы взята система двойного маятника (Рис. 2). 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.

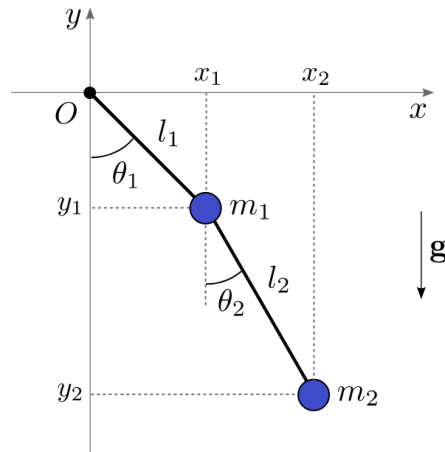


Рис. 2: Схема физической системы двойного маятника

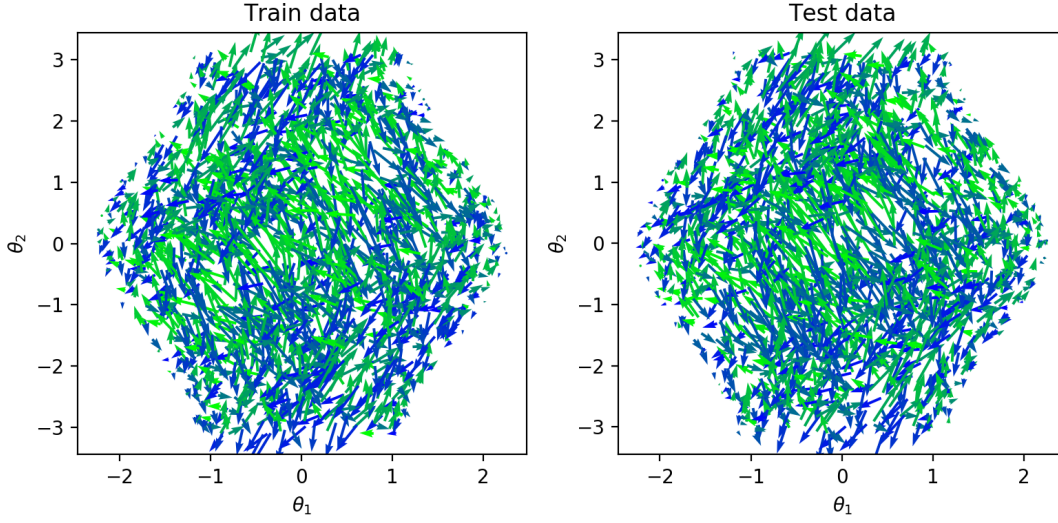


Рис. 3: Визуализация данных канонических координат системы двойного маятника

5.2.1 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 2, 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:

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

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

энергия T определяется выражением:

$$\begin{aligned} T &= \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 \\ &= \frac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2) \\ &= \frac{1}{2}m_1l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2\left[l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1 - \theta_2)\right] \end{aligned}$$

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

$$\begin{aligned} V &= m_1gy_1 + m_2gy_2 \\ &= -m_1gl_1\cos\theta_1 - m_2g(l_1\cos\theta_1 + l_2\cos\theta_2) \\ &= -(m_1 + m_2)gl_1\cos\theta_1 - m_2gl_2\cos\theta_2 \end{aligned}$$

Таким образом, Лагранжиан системы:

$$\begin{aligned} L &= \frac{1}{2}(m_1 + m_2)l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2l_2^2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1 - \theta_2) \\ &\quad + (m_1 + m_2)gl_1\cos\theta_1 + m_2gl_2\cos\theta_2 \end{aligned}$$

Исходя из данного лагранжиана получаем канонические координаты системы (Рис. 3)

5.2.2 Аналитическое решение получения динамики системы

The canonical momenta associated with the coordinates θ_1 and θ_2 can be obtained directly from L :

$$\begin{aligned} p_{\theta_1} &= \frac{\partial L}{\partial \dot{\theta}_1} = (m_1 + m_2)l_1^2\dot{\theta}_1 + m_2l_1l_2\dot{\theta}_2\cos(\theta_1 - \theta_2) \\ p_{\theta_2} &= \frac{\partial L}{\partial \dot{\theta}_2} = m_2l_2^2\dot{\theta}_2 + m_2l_1l_2\dot{\theta}_1\cos(\theta_1 - \theta_2) \end{aligned}$$

The equations of motion of the system are the Euler-Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}_i}\right) - \frac{\partial L}{\partial \theta_i} = 0 \implies \frac{dp_{\theta_i}}{dt} - \frac{\partial L}{\partial \theta_i} = 0 \quad \text{for } i = 1, 2$$

Since:

$$\begin{aligned} \frac{dp_{\theta_1}}{dt} &= (m_1 + m_2)l_1^2\ddot{\theta}_1 + m_2l_1l_2\ddot{\theta}_2\cos(\theta_1 - \theta_2) \\ &\quad - m_2l_1l_2\dot{\theta}_2\dot{\theta}_1\sin(\theta_1 - \theta_2) + m_2l_1l_2\dot{\theta}_2^2\sin(\theta_1 - \theta_2) \\ \frac{dp_{\theta_2}}{dt} &= m_2l_2^2\ddot{\theta}_2 + m_2l_1l_2\ddot{\theta}_1\cos(\theta_1 - \theta_2) \\ &\quad - m_2l_1l_2\dot{\theta}_1^2\sin(\theta_1 - \theta_2) + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_1 - \theta_2) \\ \frac{\partial L}{\partial \theta_1} &= -m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_1 - \theta_2) - (m_1 + m_2)gl_1\sin\theta_1 \\ \frac{\partial L}{\partial \theta_2} &= m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_1 - \theta_2) - m_2gl_2\sin\theta_2 \end{aligned}$$

then equation (10) yields (after dividing by l_1 when $i = 1$ and by $m_2 l_2$ when $i = 2$):

$$(m_1 + m_2) l_1 \ddot{\theta}_1 + m_2 l_2 \ddot{\theta}_2 \cos(\theta_1 - \theta_2) + m_2 l_2 \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) + (m_1 + m_2) g \sin \theta_1 = 0$$

$$l_2 \ddot{\theta}_2 + l_1 \ddot{\theta}_1 \cos(\theta_1 - \theta_2) - l_1 \dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + g \sin \theta_2 = 0$$

The canonical momenta associated with the coordinates θ_1 and θ_2 can be obtained directly from L :

$$p_{\theta_1} = \frac{\partial L}{\partial \dot{\theta}_1} = (m_1 + m_2) l_1^2 \dot{\theta}_1 + m_2 l_1 l_2 \dot{\theta}_2 \cos(\theta_1 - \theta_2)$$

$$p_{\theta_2} = \frac{\partial L}{\partial \dot{\theta}_2} = m_2 l_2^2 \dot{\theta}_2 + m_2 l_1 l_2 \dot{\theta}_1 \cos(\theta_1 - \theta_2)$$

The equations of motion of the system are the Euler-Lagrange equations:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}_i} \right) - \frac{\partial L}{\partial \theta_i} = 0 \implies \frac{dp_{\theta_i}}{dt} - \frac{\partial L}{\partial \theta_i} = 0 \quad \text{for } i = 1, 2$$

Since:

$$\begin{aligned} \frac{dp_{\theta_1}}{dt} &= (m_1 + m_2) l_1^2 \ddot{\theta}_1 + m_2 l_1 l_2 \ddot{\theta}_2 \cos(\theta_1 - \theta_2) \\ &\quad - m_2 l_1 l_2 \dot{\theta}_2 \dot{\theta}_1 \sin(\theta_1 - \theta_2) + m_2 l_1 l_2 \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) \\ \frac{dp_{\theta_2}}{dt} &= m_2 l_2^2 \ddot{\theta}_2 + m_2 l_1 l_2 \ddot{\theta}_1 \cos(\theta_1 - \theta_2) \\ &\quad - m_2 l_1 l_2 \dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) \\ \frac{\partial L}{\partial \theta_1} &= -m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) - (m_1 + m_2) g l_1 \sin \theta_1 \\ \frac{\partial L}{\partial \theta_2} &= m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) - m_2 g l_2 \sin \theta_2 \end{aligned}$$

then equation (10) yields (after dividing by l_1 when $i = 1$ and by $m_2 l_2$ when $i = 2$):

$$(m_1 + m_2) l_1 \ddot{\theta}_1 + m_2 l_2 \ddot{\theta}_2 \cos(\theta_1 - \theta_2) + m_2 l_2 \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) + (m_1 + m_2) g \sin \theta_1 = 0$$

$$l_2 \ddot{\theta}_2 + l_1 \ddot{\theta}_1 \cos(\theta_1 - \theta_2) - l_1 \dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + g \sin \theta_2 = 0$$

Equations (15) and (16) form a system of coupled second-order nonlinear differential equations. Dividing equation (15) by $(m_1 + m_2) l_1$ and equation (16) by l_2 and also moving all terms which do not involve $\ddot{\theta}_1$ and $\ddot{\theta}_2$ to the right-hand side, we obtain:

$$\begin{aligned} \ddot{\theta}_1 + \alpha_1(\theta_1, \theta_2) \ddot{\theta}_2 &= f_1(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) \\ \ddot{\theta}_2 + \alpha_2(\theta_1, \theta_2) \ddot{\theta}_1 &= f_2(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) \end{aligned}$$

where:

$$\alpha_1(\theta_1, \theta_2) := \frac{l_2}{l_1} \left(\frac{m_2}{m_1 + m_2} \right)$$

$$\alpha_2(\theta_1, \theta_2) := \frac{l_1}{l_2} \cos(\theta_1 - \theta_2)$$

and:

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

Interestingly, f_1 does not depend on $\dot{\theta}_1$ and f_2 does not depend on $\dot{\theta}_2$. Equations (17) and (18) can be combined into a single equation:

$$A \begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} = \begin{pmatrix} 1 & \alpha_1 \\ \alpha_2 & 1 \end{pmatrix} \begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$$

where the matrix A depends on θ_1 and θ_2 since α_1 and α_2 depend on these variables. Being a 2×2 matrix, A can be inverted directly:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{pmatrix} = \frac{1}{1 - \alpha_1 \alpha_2} \begin{pmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{pmatrix}$$

Before we continue, notice that A is always invertible since:

$$\det(A) = 1 - \alpha_1 \alpha_2 = 1 - \left(\frac{m_2}{m_1 + m_2} \right) \cos^2(\theta_1 - \theta_2) > 0$$

because $m_2/(m_1 + m_2) < 1$ and $\cos^2(x) \leq 1$ for all real values of x . From equations (23) and (24) we obtain:

$$\begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} = A^{-1} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \frac{1}{1 - \alpha_1 \alpha_2} \begin{pmatrix} f_1 - \alpha_1 f_2 \\ -\alpha_2 f_1 + f_2 \end{pmatrix}$$

Finally, letting $\omega_1 := \dot{\theta}_1$ and $\omega_2 := \dot{\theta}_2$, we can write the equations of motion of the double pendulum as a system of coupled first order differential equations on the variables $\theta_1, \theta_2, \omega_1, \omega_2$:

$$\frac{d}{dt} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ g_1(\theta_1, \theta_2, \omega_1, \omega_2) \\ g_2(\theta_1, \theta_2, \omega_1, \omega_2) \end{pmatrix}$$

$$g_1 := \frac{f_1 - \alpha_1 f_2}{1 - \alpha_1 \alpha_2} \quad g_2 := \frac{-\alpha_2 f_1 + f_2}{1 - \alpha_1 \alpha_2}$$

where $\alpha_i = \alpha_i(\theta_1, \theta_2)$ and $f_i = f_i(\theta_1, \theta_2, \omega_1, \omega_2)$ for $i = 1, 2$ are given on equations (19)(22).

Equation (27) can be solved numerically using a Runge-Kutta (RK) method. A simulator based on the fourth-order RK method can be found here.

5.3 Исследуемые модели нейронных сетей

5.3.1 Полносвязная нейронная сеть

$$f_{\mathbf{w}}(\mathbf{x}) := \sigma_K \left(W^{(K)} \sigma_{K-1} \left(W^{(K-1)} \dots \sigma_1 \left(W^{(1)} \mathbf{x} \right) \right) \right) \quad (4)$$

	FC	LSTM	LNN	Нётеровская LNN
MSE	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01

Таблица 1: Средняя ошибка MSE между предсказанной динамикой системы нейронной сетью и динамикой системы, полученной аналитическим решением.

5.3.2 LSTM

5.3.3 LNN

5.3.4 LNN, учитывающая трансляционную и вращательную симметрии

Нётеровская Лагранжева нейронная сеть LNN, получающая на вход разницу между каноническими координатами $\delta\theta_{12} = \theta_1 - \theta_2$ и аппроксимирующая потенциальную энергию системы $V(\delta\theta_{12})$

Аппроксимируемый лагранжиан примет вид:

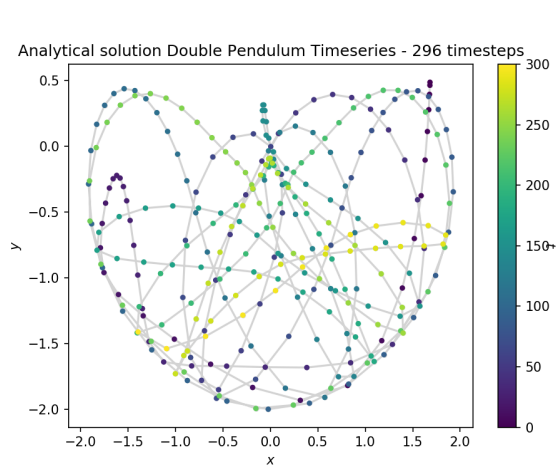
$$L = \frac{1}{2} (m_1 + m_2) l_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 l_2^2 \dot{\theta}_2^2 + m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\delta\theta_{12}) + V(\delta\theta_{12})$$

Теорема 1 Нётеровская LNN учитывает трансляционную симметрию. **Доказательство** аааааа

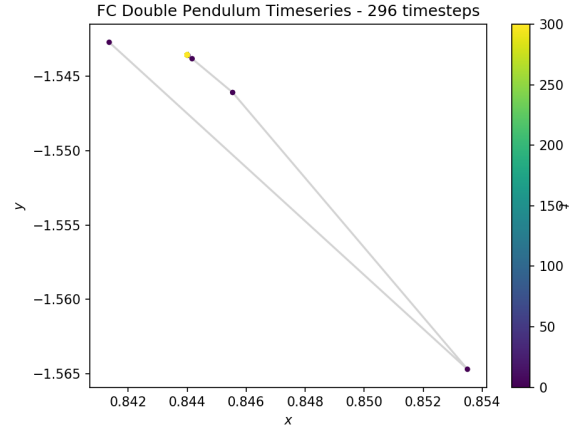
Теорема 2 Нётеровская LNN учитывает вращательную симметрию. **Доказательство** аааааа

5.4 Результаты

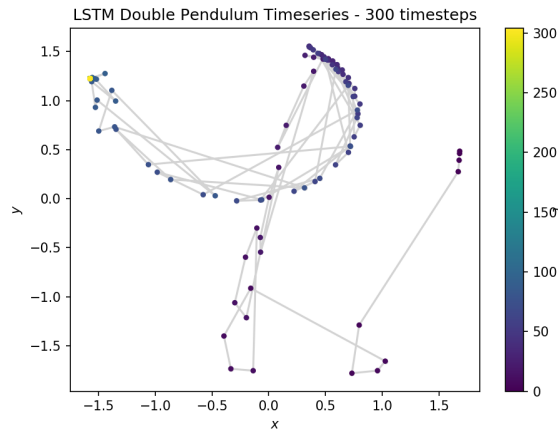
Результаты моделирования динамики системы двойного маятника различными видами нейронных сетей представлены на рисунке 4



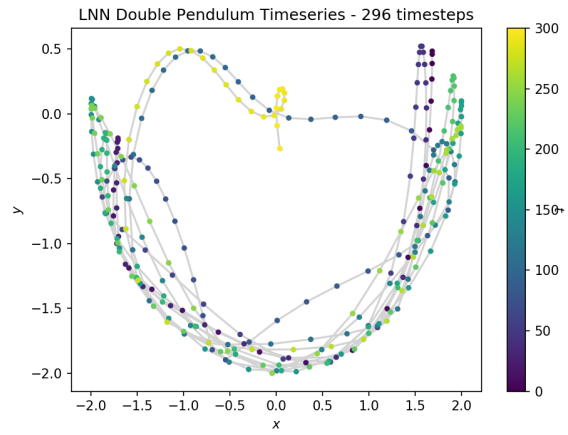
(a) Аналитическое решение



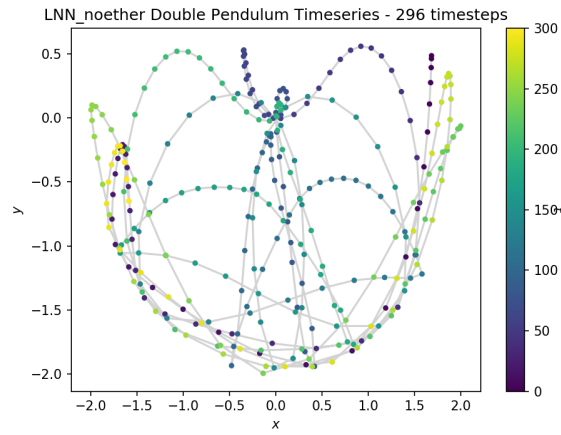
(b) Динамика системы с помощью полносвязной нейронной сети



(c) Динамика системы с помощью LSTM



(d) Динамика системы с помощью LNN



(e) Динамика системы с помощью LNN

Рис. 4: Моделирование динамики системы двойного маятника различными видами нейронных сетей: аналитическое решение, полносвязная нейронная сеть, LSTM, LNN.

6 Заключение

В работе рассмотрена задача

- Показано, что LNN является оптимальной среди моделей FN, LSTM
- Представлена Нётеровская LNN, учитывающая дополнительные симметрии кроме закона сохранения энергии
- Показано, что более интерпретируемая модель дает более точные результаты для решения задачи моделирования динамики физической системы

Список литературы

- [1] Samuel Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.