Классификация траекторий динамических систем с помощью физически-информированных нейросетей

А. А. Терентьев[1], С. К. Панченко[1], В.В. Стрижов
terentev.aa@phystech.edu; panchenko.sk@phystech.edu
[1]ФПМИ, МФТИ

В работе решается задача классификации траекторий физических систем. Динамика систем, соответствующих траекториям, моделируется лагранжианом, который аппроксимируется с помощью Лагранжевых нейронных сетей. Показано, что для динамики физических систем, соответствующим различным движениям выполняется гипотеза компактности и отделимости функциональных классов лагранжианов. Эксперимент производился на датасете PAMAP2, его результат подтвердил, что параматрезация с помощью Лагранжевых нейронных сетей является достаточно общей для описания физических систем, а параметры информативным признаковым описанием для задачи классификации.

Ключевые слова: *физическая система; лагранжиан; лагарнжева нейронная сеть*

1 Введение

Одним из способов описания динамики физических систем является лагранжиан. Его удобство является его общность, у каждой физической системы есть лагранжиан, он отвечает только за ее динамику, его можно записывать в любых координатах, которые являются параметрами для данной системы. Поэтому для классификации траекторий предлагается использовать именно лагранжиан.

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)

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

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

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

Лагранжиан аппроксимируется с помощью Лагранжевых нейронных сетей. Они позволяют аппроксимировать довольно широкий класс функций. А также точнее остальных сетей моделируют его, т.к. обладают априорными знаниями о физике системы[2]. Также к ним можно применить модификации добавляющие дополнительные знания о физике системых[1].

Для лагранжианов физических систем утверждается гипотеза компактности: лагранжианы, отвечающие различным движениям, лежат в различных компактных и отделимых подпространствах(классах) пространства моделируемых функций. В нашем конкретном случае это означает, что векторы параметров функций, лежащих в этих подпространствах, также еомпактны и отделимы.

Для проверки гипотезы был произведен вычислительный эксперпимент на датасете PAMAP2, который содержит траетории движений человека во время различных активностей.

2 Постановка задачи классификации траекторий физических систем

Задана выборка с метками из n траекторий:

$$\{\mathcal{D}_j, z_j\}_{j=1}^{n},$$

где:

$\mathcal{D}_j = \{\mathbf{x}_i^{(j)}, \mathbf{y}_i^{(j)}\}_{i=1}^{m_j}$ – j-ая траектория,
$\mathbf{x}_i^{(j)} = (\mathbf{q}_i^{(j)}, \dot{\mathbf{q}}_i^{(j)})$ – i-ые координаты и скорости j-ой траектории,
$\mathbf{y}_i^{(j)} = \dot{\mathbf{x}}_i^{(j)} = (\dot{\mathbf{q}}_i^{(j)}, \ddot{\mathbf{q}}_i^{(j)})$ – i-ые скорости и ускорения j-ой траектории,
$\mathbf{q}_i^{(j)} \in \mathbb{R}^r$ – вектор обобщенных координат,
r – количество координат,
m_j – длина j-ой траектории,
$z_j \in \overline{1, K}$ – метка j-ой траектории.

2.1 Задачи регрессии восстановления лагранжиана

Сведём задачу моделирования лагранжиана системы к задаче регрессии. Регрессионная модель для траектории D_j выбирается из класса лагранженных нейронных сетей и представляет собой композицию:

$$g_j\colon (\mathbf{x}|\mathbf{w}) \to L, \quad \mathbf{x} \in \mathbb{R}^{2\times r},\ L \in \mathbb{R},$$

$$\varkappa : L \to \mathbf{y}, \quad \mathbf{y} \in \mathbb{R}^{2\times r},$$

$$\mathbf{f}_j = \varkappa \circ g_j,$$

где:

g_j – параметрическая функция, аппроксимириующая лагранжиан системы,
$\mathbf{x} \in \mathbb{R}^{2\times r}$ – элемент траектории (r обобщенных координат и r обобщенных скоростей),
$\mathbf{w} \in \mathbb{W}$ – параметры аппроксимирующей лагранжиан модели,
$L \in \mathbb{R}$ – значение восстановленного лагранжиана,
$\varkappa$ – функция, реализующая уравнения Эйлера-Лагранжа для нахождения значения обобщенных скоростей и ускорений $\mathbf{y}$,
$\mathbf{y} \in \mathbb{R}^{2\times r}$ – целевой вектор в задаче восстановления лагранжиана (r обобщенных скоростей и r обобщенных ускорений)
$\mathbf{f}_j$ – финальная композиция g_j и $\varkappa$, представляющая собой Лагранжеву нейронную сеть.

Задача моделирования динамики системы представлена в виде задачи минимизации квадратичной ошибки:

$$\mathcal{L}(\mathbf{w}) = \mathcal{L}(\mathbf{w}|\mathcal{D}_j) = \frac{1}{m_j}\sum_{i=1}^{m_j} \|\hat{\mathbf{y}}_i^{(j)} - \mathbf{y}_i^{(j)}\|_2^2,$$

$$\mathbf{w}_j^* = \arg\min_{\mathbf{w}\in\mathbb{W}} \left(\mathcal{L}(\mathbf{w})\right),$$

где $\hat{\mathbf{y}}_i^{(j)} = \mathbf{f}_j(\mathbf{x}_i^{(j)}|\mathbf{w}) = \varkappa\big(g_j(\mathbf{x}_i^{(j)}|\mathbf{w})\big) \in \mathbb{R}^{2\times r}$ – предсказанная динамика движения физической системы на j-ой траектории.

2.2 Задачи классификации траекторий по лагранжианам

После решения задачи моделирования лагранжиана получаем задачу классификации:

$$\{\mathbf{w}_j^*, z_j\}_{j=1}^n,$$

где $\mathbf{w}_j^*$ – коэффициенты аппроксимированного лагранжиана j-ой траектории.

Для ее решения используются различные методы классификации, среди которых: логистическая регрессия, гауссовская классификация, случайный лес.

3 Гипотеза о компактности

Пусть $\mathbb{L}$ – функциональное линейное пространство лагранжианов всевозможных систем:

$$g : \mathbf{x} \to L, \quad g \in \mathbb{L},$$

где $\mathbf{x} \in \mathbb{R}^{2\times r}$ – вектор обобщенных координат и скоростей, $L \in \mathbb{R}$ – значение лагранжиана системы. Пусть также имеется параметрическое семейство функций, аппроксимирующих лагранжиан, описанное в разделе 2.1,

$$\mathbb{L}_\varepsilon = \{g^{(\varepsilon)}\colon (\mathbf{x}|\mathbf{w}) \to L \mid \mathbf{x} \in \mathbb{R}^{2\times r},\ L \in \mathbb{R},\ \mathbf{w} \in \mathbb{W}\}, \quad \mathbb{L}_\varepsilon \subset \mathbb{L},$$

где $\mathbf{w}$ – параметры функции из линейного пространства всевозможных значений параметров $\mathbb{W}$. Данное множество функций $\mathbb{L}_\varepsilon$ представляет собой ε-сеть в $\mathbb{L}$ при достаточной сложности модели в силу аппроксимационный свойств нейросетей.

Пусть каждой из $j = \overline{1,n}$ траекторий $\mathcal{D}_j$, соответствующих различным динамическим системам, описывающимся истинным Лагранжианом $g_j \in \mathbb{L}$ и принадлежащих различным классам $z_j \in \overline{1,K}$, сопоставлена аппроксимация лагранжиана $g_j^{(\varepsilon)}(\mathbf{x}|\mathbf{w}_j^*) \in \mathbb{L}_\varepsilon$ с помощью алгоритма, описанного в разделе 2.1.

Основа классификации траекторий физических систем лежит в следующей гипотезе:

Гипотеза 1. В пространстве $\mathbb{L}$ мера пересечения различных классов мала

$$\exists\delta \ > \ 0, \ \forall i, \ j, \ i \neq j \ \mu(z_i \cap z_j) \ < \ \delta$$

Следующая же теорема позволит связать это с классификацией по параметрам нейронных сетей

Гипотеза 2. Пусть выполняются предпосылки гипотезы. А таже z_i имеют границы суммарной длинной не более R. Тогда также в пространстве $\mathbb{L}_\varepsilon$ при ε стремящемся к нулю мера пересечения различных классов мала

$$\exists\delta \ > \ 0, \ \forall i, \ j, \ i \neq j \ \mu(z_i \cap z_j) \ < \ \delta$$

Доказательство $\mu(z_i \cap z_j) \ < \ \nu(z_i \cap z_j) + \varepsilon * R < 2 \ * \ \delta$

Если для системы $\mathbb{S}$ выполняется гипотеза компктности с точностью $\delta \ll min(\mu(z_i))$, то она поддается классификации.

Для физических систем это гипотеза обычно выполняется, т. к. обычно они представлены в какого-то семейства параметризированных функций

4 Постановка задачи классификации траекторий физической системы

Выборка $\mathcal{H} = \{(\{\mathbf{x_i}\}, y_i)\}_{i=1}^{l}$ - множество объектов с известными метками классов.

- Каждый $\mathbf{x_i} = (\mathbf{q_i^{(j)}}, \mathbf{\dot{q}_i^{(j)}}, \mathbf{\ddot{q}_i^{(j)}}) \in \mathbf{X}$ объект временной ряд
- $\mathbf{X}$ - множество временных рядов возможно различной длины
- $y_i \in 1, K$ - метка i-ой траектории
- r - кол-во координат

4.1 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

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

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

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

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

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 $\frac{d}{dt}\frac{\partial L}{\partial \mathbf{\dot{q}}} = \frac{\partial L}{\partial \mathbf{q}}$
4. Решить получившуюся систему дифференциальных уравнений

4.2 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

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

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

5 Вычислительный эксперимент

5.1 Цель эксперимента

- Проверить способность Лагранжевой нейросети моделировать физические системы достаточно эффективно по времени
- Проверить правильность гипотензы о компактности и используя ее классифицировать траектории физические системы.
- Подобрать подходящий под данное распределения алгоритм классификации

5.2 Набор данных для базового эксперимета

Набор данных с акселерометров из датасета PAMAP2

- Количество акселерометров: T = 3
  - Закрепленный на запастье преобладающей руки
  - Закрепленный на груди
  - Закрепленный на локте преобладающей руки
- Количество классов: K = 24
- Количество испытуемых: M = 9
- Каждая активность длилась минимум 1 час, частота сбора данных 100 Гц

5.3 Подготовка данных

Для данных полученных с аксселерометров существует несколько проблем

- Пропуски во временных рядах
- Временные ряды состоят только из ускорений, либо скоростей

Поэтому требуется восстановить пропущенные данные, импользуя интерполяцию, численное диффенцирование и интегрирование

- Датчики иногда могут пропускать один или два такта. Для востановления ряда используется сплайн-интерполяция, кторая позволяет получить квадратичную точность от размеров сетки.

– Для восстановления ускорения, скорости и координат применется численное дифференцирования и интергрирования, соответсвенно методы направленной разности и Рунге-Кутты второго порядка

Для дифференцирования используется центральная разность

$$f'(1) \ = \ \frac{f(2) \ - \ f(0)}{2h} \ - \ \frac{h^2}{6} f^{(3)}$$

Для интегрирования Формула Кортеса

$$\int_a^b f(x)dx \approx \frac{h}{3} \sum_{k \ = \ 1}^{N \ - \ 1} f(x_{k-1}) + 4f(x_k) + f(x_{k+1})$$

,

$$|E(f)| \ \leqslant \ \frac{(b-a)}{2880} h^4 max \left| f^{(4)}\left(x\right) \right|$$

5.4 Эксперимент

Берется подвыборка из датасета PAMAP2. Для каждого объекта приеняется Лагранжева нейро сеть и с помощью нее параметризуется лагранжиан.

Для проверки гипотезы построены 2D и 3D графики для соответстенно двух и трех параметров, полученных понижением размерности методом главных компонент

Some 2D Grafic

Some 3D Grafic

Для использования классификаторов также понизим размерность с помощью метода главных компонент.

Some acc grafic

Из графика видно, что достаточнео всего 10 главных компонент, а дальнейшее повышение размерности не имеет смысла

Из всех методов классификаций лучше всего себя показала гауссовская классификация. Её и выберем за основной алгоритм

– Логистическая регрессия – 76%
– Ядерный метод – 88%
– Случайный лес – 84%

5.5 Заключение

Предложен метод классификации физических тракторий, не зависящий от физически не значимых ей изменениийю Экспериментально показано, что классы траекторий отделимы по параметризации Лагранжианаю Предложен метод классификации Лагранжианов. Предлагается использовать Гаусовский процесс для разделения по классам, как самый точный. Проведена классификация реальных траеторий движения частей тела людей при различных активностях

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