

Классификация траекторий динамических систем с помощью физически-информированных нейросетей

А. А. Терентьев¹, С. К. Панченко¹, В. В. Стрижов

terentev.aa@phystech.edu; panchenko.sk@phystech.edu

¹ФПМИ, МФТИ

Целью работы является проверка гипотезы о возможности классификации траекторий физических систем по лагранжианам, а также задачей ставится предложить метод для данной классификации. Идея состоит в том, чтобы сначала восстановить лагранжиан системы, которая могла бы породить данную траекторию. Лагранжиан восстанавливается с помощью физически-информированных Лагранжевых нейронных сетей. Предложена норма, которая вводится на пространстве лагранжианов. Она используется как метрика для метрических методов классификации. В основном для многомерных классификаций используются свёрточные нейронные сети, в предположении, что существуют статистические связи, между ближайшими точками временного ряда. Данные методы хорошо себя показывают во многих задачах, где эти связи являются преобладающими. Но для физических систем такие методы не подходят, и предлагается использовать знания о физических связях систем, для выбранных Лагранжевых нейронных сетей этим знанием является закон сохранения энергии.

Ключевые слова: *физическая система; лагранжиан; лагранжева нейронная сеть; гипотеза компактности*

1 Введение

Одним из способов описания физических объектов является лагранжева динамика. [4] Преимущество описания с помощью лагранжиана является общность. Динамика физической системы описывается лагранжианом, который записывается в произвольных координатах, являющихся параметрами для данной системы. Лагранжиан предлагается использовать для классификации траекторий объектов.

Для использования лагранжевой механики выбираются обобщенные координаты. Выбирается минимальное количество независимых переменных (обобщённых координат), необходимое для полного описания состояния механической системы. Иметь знания о Лагранжиане системы, который определяется как разница между кинетической энергией T и потенциальной энергией V :

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

где $\mathbf{q}(t)$ — обобщенные координаты системы, $\dot{\mathbf{q}}(t)$ — обобщенные скорости системы, L — лагранжиан системы, далее для краткости мы опускаем t .

Динамика системы описывается системой уравнений Эйлера-Лагранжа:

$$\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} = \mathbf{Q}_x = 0.$$

Лагранжиан восстанавливается с помощью Лагранжевых нейронных сетей. Априорные знания о физике системы являются преимуществом данной модели по сравнению с классическими нейронными сетями [1, 2].

Для лагранжианов физических систем утверждается гипотеза компактности: лагранжианы, отвечающие различным движениям, лежат в различных компактных и отделимых

подмножествах-классах пространства моделируемых функций. На основе этого делается вывод, что векторы параметров функций, соответствующих различным классам, также компактны и отделимы в общем пространстве параметров.

Для проверки гипотезы был произведен вычислительный эксперимент на датасете РАМАР2. Датасет содержит траектории движений человека во время обычных дневных активностей: готовка еды, пробежка, поход до магазина и другие.

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, $\mathbf{y}_i^{(j)} = \ddot{\mathbf{q}}_i^{(j)}$ — 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.

Лагранжева динамика

Сведём задачу моделирования лагранжиана системы к задаче регрессии. Регрессионная модель для траектории \mathcal{D}_j выбирается из класса лагранжевых нейронных сетей и представляет собой композицию:

$$\begin{aligned} g_j: (\mathbf{x}|\mathbf{w}) &\rightarrow L, \quad \mathbf{x} \in \mathbb{R}^{2 \times r}, \quad L \in \mathbb{R}, \\ \varkappa: L &\rightarrow \mathbf{y}, \quad \mathbf{y} \in \mathbb{R}^{2 \times r}, \\ \mathbf{f}_j &= \varkappa \circ g_j, \end{aligned}$$

где g_j — параметрическая функция, аппроксимирующая 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:

$$\begin{aligned} \mathcal{L}(\mathbf{w}) &= \mathcal{L}(\mathbf{w}|\mathcal{D}_j) = \frac{1}{m_j} \sum_{i=1}^{m_j} \|\hat{\mathbf{y}}_i^{(j)} - \mathbf{y}_i^{(j)}\|_2^2, \\ \mathbf{w}_j^* &= \arg \min_{\mathbf{w} \in \mathbb{W}} (\mathcal{L}(\mathbf{w})), \end{aligned}$$

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

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. Т.к. достаточно трудно сказать какой алгоритм классификации подойдет для лагранжианов, то были взяты логистическая регрессия, гауссовская классификация, случайный лес.

3 Гипотеза компактности

Мы считаем, что для лагранжианов траекторий верна гипотеза компактности, т. е., что лагранжианы лежат в компактно локализованных подмножествах. В основе этого предположения лежит то, что в основе различных движениях человека лежат различные группы мышц. А в схожих соответственно схожие. Поэтому разные по природе движения будут образовывать разные механические системы и наоборот. Отсюда естественно вытекает гипотеза.

Осталось эту гипотезу проверить. Для этого необходимо уточнить термины, упомянутые в первом абзаце.

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

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

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,

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

где \mathbf{w} — параметры функции из линейного пространства всевозможных значений параметров \mathbb{W} . Данное множество функций \mathbb{L}_ε представляет собой множество лагранжианов из параметрического семейства функций, аппроксимирующих лагранжиан, описанного в разделе 2.1, $\mathbb{L}_\varepsilon = \{g^{(\varepsilon)} : (\mathbf{x}|\mathbf{w}) \rightarrow L \mid \mathbf{x} \in \mathbb{R}^{2 \times r}, L \in \mathbb{R}, \mathbf{w} \in \mathbb{W}\}$, $\mathbb{L}_\varepsilon \subset \mathbb{L}$, где \mathbf{w} — параметры функции из линейного пространства всевозможных значений параметров \mathbb{W} .

Для построения нормы получим систему уравнений для вычисления ускорений.

Уравнения Эйлера-Лагранжа.

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 правила цепочки: Распишем производную по времени через производные по координатам:

$$\left(\frac{\partial}{\partial \mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \right) \frac{dq}{dt} + \left(\frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \right) \frac{d\dot{\mathbf{q}}}{dt} = \frac{\partial L}{\partial \mathbf{q}},$$

Перепишем производные по времени для наглядности

$$\frac{\partial}{\partial \dot{\mathbf{q}}} \left(\frac{\partial L}{\partial \mathbf{q}} \dot{\mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}} \ddot{\mathbf{q}} \right) = \frac{\partial L}{\partial \mathbf{q}},$$

Раскроем скобки

$$\frac{\partial}{\partial \mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}} + \frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \ddot{\mathbf{q}} = \frac{\partial L}{\partial \mathbf{q}},$$

Перенесем слагаемые так, чтобы слева остались только неизвестные

$$\frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \ddot{\mathbf{q}} = \frac{\partial L}{\partial \mathbf{q}} - \frac{\partial}{\partial \mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}},$$

Выразим неизвестную переменную разделив обе части на одно выражение

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

Перепишем частые производные в другой нотации

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

где гессиан

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

Замечание 1. $\det H \neq 0$ в любой точке, а значит он положительно определен. Данное условие необходимо, т. к. иначе существует малая вариация, при которой состояние системы не меняется. Следовательно, имеем не вырожденную систему линейных уравнений относительно ускорений.

$$H \ddot{\mathbf{q}} = \frac{\partial L}{\partial \mathbf{q}} - \frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}},$$

Лемма 1. Если лагранжины заданы на компакте, то существуют неотрицательные числа a, b такие, что $\det H \geq a$, а собственные значения матрицы H не меньше b

Доказательство. Предположим, что это не так. H является непрерывной функцией, т.к. ее значения являются вторыми производными L , следовательно, $\det H$ и собственные значения также непрерывны, а значит достигают своего минимума и максимума. Получили противоречия, т.к. матрица невырожденная во всех точках. \square

Замечание 2. Если дополнительно предположить, что система уравнений не является вырожденной в пределе, то утверждение выше верно и для всего пространства

Т.к. уравнения инвариантны относительно умножения Лагранжиана на константу, то мы можем выбрать лагранжины для систем так, чтобы собственные значения матрицы H были не меньше 1.

Теперь видно, что на пространстве \mathbb{L} можно задать норму.

Лемма 2. Оператор $A(L) = \frac{\partial L}{\partial \mathbf{q}} - \frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \mathbf{q}$ является линейным

Доказательство.

$$\begin{aligned} A(L) &= \frac{\partial \alpha L_1 + \beta L_2}{\partial \mathbf{q}} - \frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial \alpha L_1 + \beta L_2}{\partial \dot{\mathbf{q}}} \mathbf{q} = \frac{\partial \alpha L_1}{\partial \mathbf{q}} + \frac{\beta L_2}{\partial \mathbf{q}} - \left(\frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial \alpha L_1}{\partial \dot{\mathbf{q}}} + \frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial \beta L_2}{\partial \dot{\mathbf{q}}} \right) \mathbf{q} = \\ &= \alpha \frac{\partial L_1}{\partial \mathbf{q}} + \beta \frac{\partial L_2}{\partial \mathbf{q}} - \alpha \frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial L_1}{\partial \dot{\mathbf{q}}} \mathbf{q} - \beta \frac{\partial}{\partial \dot{\mathbf{q}}} \frac{\partial L_2}{\partial \dot{\mathbf{q}}} \mathbf{q} = \alpha A(L) + \beta A(L) \end{aligned}$$

□

Теперь рассмотрим \mathbb{L} как пространство L_2 . В нем $\|L\|_L = \|(A(L), H(L))\|_2$ является полунормой, т.к. линейный оператор сохранит абсолютную однородность и неравенство треугольника.

Но множество $\|L\|_L = 0$ также является подпространством L_2 , т.к. $A(L)$, $H(L)$ - линейные операторы, а значит их ядра линейные пространства. Тогда можно ввести факторпространство, в котором отношением эквивалентности будет равенство 0 полунормы $\|L\|_L$.

Рассмотрим обоснованность введенного отношения эквивалентности.

Лемма 3. $\|L\|_L = 0 \Leftrightarrow$ п.в. $\delta \ddot{q} = 0$, где $\delta L, \delta \ddot{q}$ являются вариациями лагранжиана и ускорения.

Доказательство.

$$\|(A(\delta L), H(\delta L))\|_2 = 0 \Leftrightarrow \text{п.в. } A(\delta L) = 0, H(\delta L) = 0$$

$$A(L) = H(L) \ddot{q}$$

$$A(L + \delta L) = H(L + \delta L)(\ddot{q} + \delta \ddot{q})$$

$$A(\delta L) = H(\delta L)(\ddot{q}) + H(\delta L + \delta L)(\delta \ddot{q})$$

$$0 = H(\delta L + \delta L)(\delta \ddot{q}) \Leftrightarrow \delta \ddot{q} = 0$$

□

Замечание 3. Заметим, что в большинстве случаев $\|H(\delta L)\|_2 = 0$, т.к. всегда можно выбрать координаты так, с помощью преобразования Лежандра, чтобы матрица Гессе была единичной, для этого необходимо перейти в канонические координаты, а значит достаточно сравнивать только первую часть нормы $\|A(L)\|_2$

Данную норму вычислить точно в общем случае невозможно, т.к. для этого нужно посчитать интеграл функции на всей числовой прямой, но его можно приблизить

$$\|(A(L), H(L))\|_2 = \sqrt{\int |A(L)(\mathbf{q}, \dot{\mathbf{q}})|^2 + \|H(L)(\mathbf{q}, \dot{\mathbf{q}})\|_2^2 d\Omega} \approx$$

$$\sum_{i_1=1, \dots, i_n=1}^{N, \dots, N} |A(L)(\mathbf{q}, \dot{\mathbf{q}})|^2 + \|H(L)(\mathbf{q}, \dot{\mathbf{q}})\|_2^2 \frac{\Delta X_1 \cdot \dots \cdot X_n}{N \cdot \dots \cdot N} =$$

$$C \cdot \sum_{i_1=1, \dots, i_n=1}^{N, \dots, N} \frac{A(L)(\mathbf{q}, \dot{\mathbf{q}})^2 + \|H(L)(\mathbf{q}, \dot{\mathbf{q}})\|_2^2}{N \cdot \dots \cdot N} = C \cdot \overline{|A(L)(\mathbf{q}, \dot{\mathbf{q}})|^2 + \|H(L)(\mathbf{q}, \dot{\mathbf{q}})\|_2^2}$$

Т.к. мы решаем задачу классификации, то сопоставим каждому лагранжиану g его класс \mathcal{A}_i . Таким образом мы имеем конечное семейство непересекающихся замкнутых выпуклых множеств \mathcal{A} в пространстве \mathbb{L} лагранжианов, где каждое семейство представляет класс, соответствующий движению человека

Теорема 1. *Пусть есть конечное семейство непересекающихся замкнутых выпуклых множеств \mathcal{A} в нормированном пространстве \mathbb{L} , тогда существует $\varepsilon > 0$, что для любого преобразования пространства φ такое, что $\|\varphi(\mathcal{A}_i) - \mathcal{A}_i\| < \varepsilon$ множества из семейства $\varphi\mathcal{A}$ попарно сильно отделимы.*

Доказательство. Согласно первой теореме Хана-Банаха об отделимости множества из семейства \mathcal{A} попарно отделимы

$$\langle a, x \rangle_{x \in \mathcal{A}_i} \leq \gamma_1 < \gamma_2 \leq \langle a, y \rangle_{y \in \mathcal{A}_j}.$$

Т.к. функционал a непрерывен, то для любого $\varepsilon > 0$ существует $U(0)$ — окрестность нуля, такая, что

$$\langle a, y \rangle_{y \in U(0)} \leq \varepsilon.$$

Пусть

$$\varepsilon_{ij} = \frac{|\gamma_1 - \gamma_2|}{3},$$

тогда

$$\langle a, x \rangle_{x \in \mathcal{A}_i + U(0)} \leq \gamma_1 + \varepsilon < \gamma_2 + \varepsilon \leq \langle a, y \rangle_{y \in \mathcal{A}_j + U(0)}.$$

Если взять $\varepsilon = \min_{i,j} \varepsilon_{ij}$, то тогда множества \mathcal{A}_i попарно сильно отделимы. □

Эта теорема позволяет исследовать свойства лагранжианов на их приближениях, что заметно упрощает задачу.

Исходя из данной теоремы, можно задать достаточное свойство для алгоритма аппроксимации. Должно существовать $\varepsilon > 0$, такое, что для любого $\mathcal{L} \in \mathbb{L}$ было верно $\|\mathcal{L} - \mathcal{L}_\varepsilon\| < \varepsilon$. Иными словами нужно, чтобы расстояние от аппроксимированного лагранжиана до истинного было небольшим. Т.е. требуется чтобы нейросеть не переобучалась и с достаточной точностью приближала лагранжиан. Данное условие выполняется для корректно построенной сети.

Уже из Теоремы 1 следует первый метод классификации лагранжиана. Т.к. в пространстве \mathbb{L} мы ввели норму, то она порождает метрику. Следовательно, можно использовать метрические методы классификации. В данной же работе предлагается использовать для классификации параметры нейронной сети, аппроксимирующей лагранжиан. В этом нам помогает следующая теорема.

Теорема 2. *Пусть есть семейство конечное семейство множеств \mathcal{A} в нормированном пространстве \mathbb{L} , $\varepsilon > 0$, и преобразования пространства φ такое, что $\|\varphi(\mathcal{A}_i) - \mathcal{A}_i\| < \varepsilon$ множества из семейства $\varphi\mathcal{A}$ попарно сильно отделимы, а также взаимно-однозначное непрерывное отображение $\psi : \mathbb{L}_\varepsilon \rightarrow \mathbb{R}^n$, обратное отображение непрерывно и выпукло. Тогда множества \mathcal{A}_i попарно сильно отделимы.*

Доказательство. Погрузимся в пространство \mathbb{R}^n . В нем $\varphi\psi\mathcal{A}_i$ попарно сильно отделимы. Тогда каждое множество ограничено соответствующими гиперплоскостями $\langle a, x \rangle \leq \gamma_1$
 $x \in \varphi\psi\mathcal{A}_i$

Расширим каждое множество до гиперплоскостей и будем рассматривать уже расширенные. Каждое из таких множеств компактно. Т.к. ψ^{-1} выпукло и непрерывно, то образ расширенных множеств также выпуклый и компактен. Следовательно, расширенные множества попарно сильно отделимы. \square

Теорема дает достаточные условия для того, чтобы для изначальных множеств выполнялась гипотеза о компактности. Но ее условия достаточно сильные. Для нее требуется, чтобы веса были непрерывны, что может не выполняться. Данный метод удобен тем, что можно не использовать много признаков, но он не подходит для доказательства того, что для набора данных не выполняется гипотеза. В случае ее невыполнения лучше использовать методы основанные на метрике, заданной нормой.

4 Лагранжевы нейронные сети.

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.

Базовая модель Лагранжевой нейронной сети.

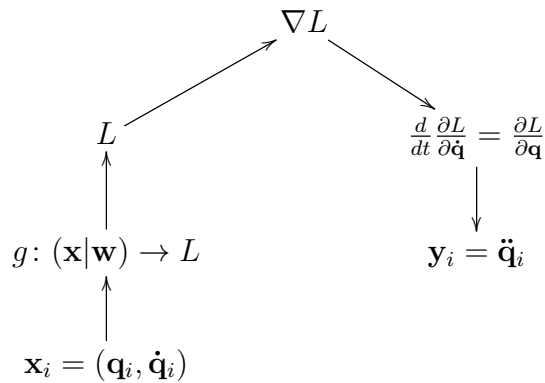


Рис. 1 Схема нейронной сети)

В работе [2] предложено в нейронную сеть

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

добавить априорные знания о физике системы, моделируя лагранжиан:

$$f : \mathbf{X} = (\mathbf{q}, \dot{\mathbf{q}}) \rightarrow \mathbf{L}.$$

Ключевой идеей является параметризация нейронной сетью лагранжиана L , получение выражения ограничения Эйлера-Лагранжа и обратное распространение ошибки через полученные ограничения

$$\ddot{\mathbf{q}} = (\nabla_{\dot{\mathbf{q}}\dot{\mathbf{q}}} L)^{-1} [\nabla_{\mathbf{q}} L - (\nabla_{\dot{\mathbf{q}}\mathbf{q}} L) \dot{\mathbf{q}}].$$

В качестве нейронной сети берется полносвязная сеть с тремя слоями. Таким образом, для заданных координат $(\mathbf{q}, \dot{\mathbf{q}})$ имеем модель с априорными знаниями о законе сохранения энергии, с помощью которой можем получить динамику параметров $(\dot{\mathbf{q}}, \ddot{\mathbf{q}})$.

5 Вычислительный эксперимент

Цель.

Целью эксперимента является классификация многомерных временных рядов, представленных в виде траекторий физических систем. Как подзадачи необходимо подтвердить гипотезу о компактности классов, подтвердить на практике выдвинутую теорию, что пространство лагранжианов можно нормировать предложенной нормой и сэмплировать ее заранее заданным набором точек и с помощью нее построить классификатор.

Набор данных для базового эксперимента.

Исследования проводились на датасете Physical Activity Monitoring [5]. Набор данных с акселерометров из датасета PAMAP2. Три гироскопа: закрепленный на запястье преобладающей руки, закрепленный на груди, закрепленный на локте преобладающей руки. Количество классов: $K = 24$. Количество испытуемых: $M = 9$. Каждая активность длилась пару минут, частота сбора данных 100 Гц.

Обработка данных.

Из всех датчиков представленных в датасете были выбраны акселерометры, как наиболее подходящие для классификации ходьбы, бега, поездки на велосипеде и неподвижного состояния обработка данных состоит в восстановлении пропусков в данных, связанных с отсутствием сигналов от гироскопов, а также получению значения углов и угловых ускорений из угловых скоростей, полученных от гироскопов. Для восстановления ряда используется сплайн-интерполяция, которая позволяет получить квадратичную точность от размеров сетки. Для восстановления ускорения, скорости и координат применяется численное дифференцирование и интегрирование.

Для дифференцирования используется центральная разность:

$$f'(x) = \frac{f(x_{k+1}) - f(x_{k-1}))}{2h}, |E(f)| \leq \frac{h^2}{6} f^{(3)}.$$

Для интегрирования используется формула Кортеса:

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

$$|E(f)| \leq \frac{(b-a)}{2880} h^4 \max |f^{(4)}(x)|.$$

Обучение LNN-сетей.

Полученные траектории передаются как объекты $\mathbf{X} = (\mathbf{q}, \dot{\mathbf{q}})(t)$ и признаки $\mathbf{Y} = \ddot{\mathbf{q}}(t)$. Нейросеть представленная в работе [2] соответственно обучается на каждой траектории отдельно и соответственно получается набор лагранжианов, соответствующих траекториям.

Эти лагранжианы представляют динамику физических систем и позволяет восстановить ускорение $\ddot{\mathbf{q}}(t)$.

Для проверки способности нейросети восстанавливать траектории с необходимой нам точностью построим предсказанные LNN ускорения $\ddot{\mathbf{q}}(t)$ для траекторий $(\mathbf{q}, \dot{\mathbf{q}})(t)$ и сравним их с реальными.

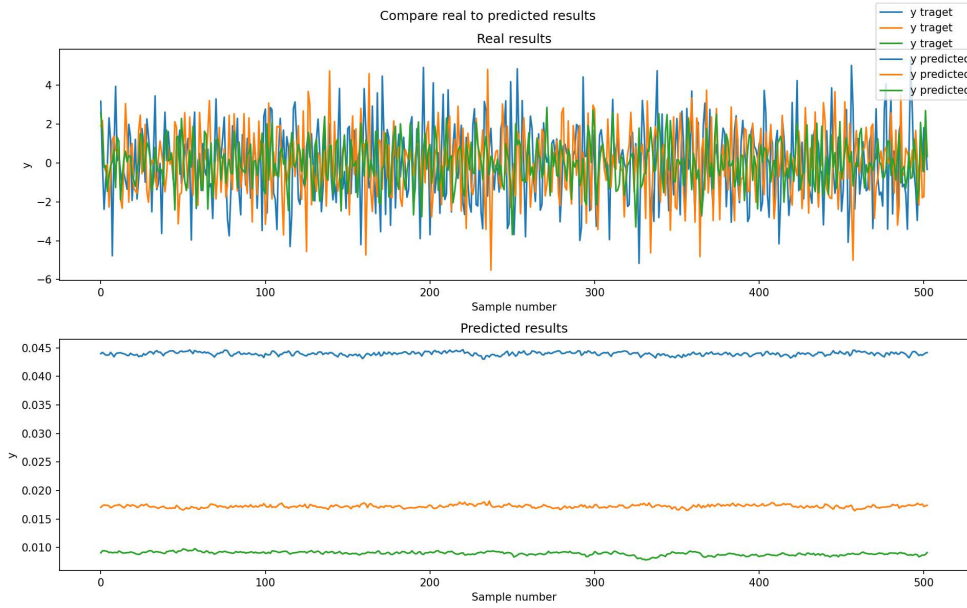


Рис. 2 Временной ряд предсказанного ускорения в зависимости от времени для траектории 3

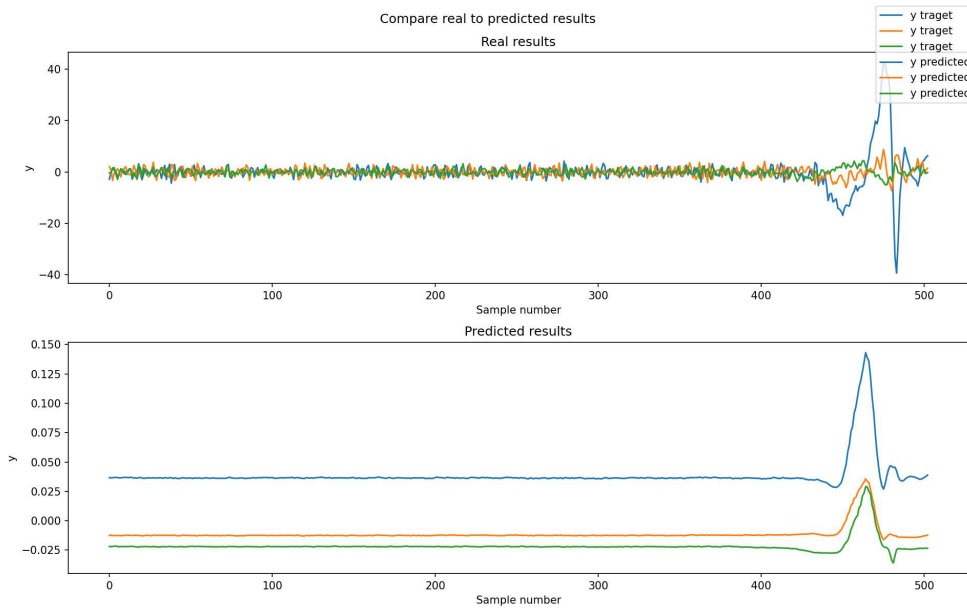


Рис. 3 Временной ряд предсказанного ускорения в зависимости от времени для траектории 6

Лагранжевы сети ищут динамику систему, поэтому они хорошо справляются с шумами, которые не выражают физическую суть системы, особенно когда системы имеют

константное ускорение. Поэтому не требуется дополнительно очищать предварительно данные от шумов, которые в данной выборке связаны с сильными усилиями при выполнении некоторых действий, и можно переходить сразу к обучению LNN-сети.

Получение признаков.

После того как лагранжианы получены мы можем восстанавливать лагранжиан системы в любой точке $(\mathbf{q}, \dot{\mathbf{q}})$. Соответственно мы можем сэмплировать норму, на основе которой строим метрику, а на основе этой метрики уже классифицировать траектории базовыми алгоритмами.

Поэтому для начала мы получаем набор точек в которых будем получать значения лагранжиана. Этот набор мы получаем до классификации, и он един для всех траекторий. Соответственно он является гиперпараметром для нашего классификатора. Мы получаем набор $S = \{s_i\}$ из $N = 20000$ точек $(\mathbf{q}, \dot{\mathbf{q}})$ равномерно распределенных на кубе $-10 \leq x_i \leq 10$. Размер куба подбирается так, чтобы он покрывал точки, на которых мы обучали лагранжианы. Кол-во точек для сэмплирования должно быть подобрана так, чтобы оценка нормы менялась слабо по сравнению с ее оценкой.

Теперь для каждого объекта получаем значения $X_i = \{AL(s_i)\}$ набор его признаков. Норма в этом пространстве будет соответствовать обычной евклидовой норме, поэтому мы можем применить любой известный нам базовый алгоритм классификации.

$$A(L) = \frac{\partial L}{\partial \mathbf{q}} - \frac{\partial}{\partial \mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} \dot{\mathbf{q}}$$

Для регуляризации и уменьшения размерности применим метод главных векторов для преобразования признаков. После этого для исследования набора представим объекты по первым трем главным признакам на 2D и 3D-графиках.

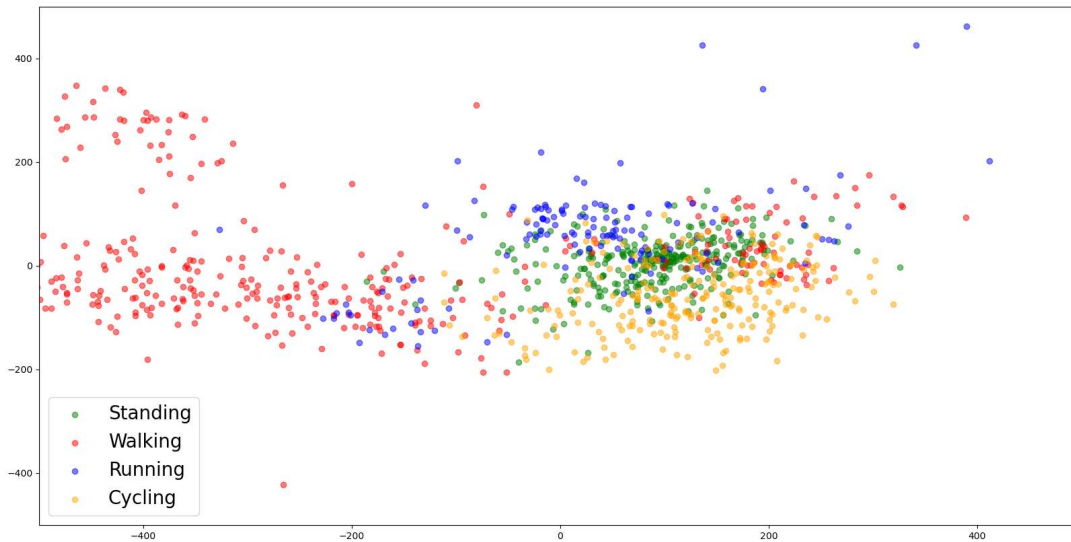


Рис. 4 Распределения данных в 2D

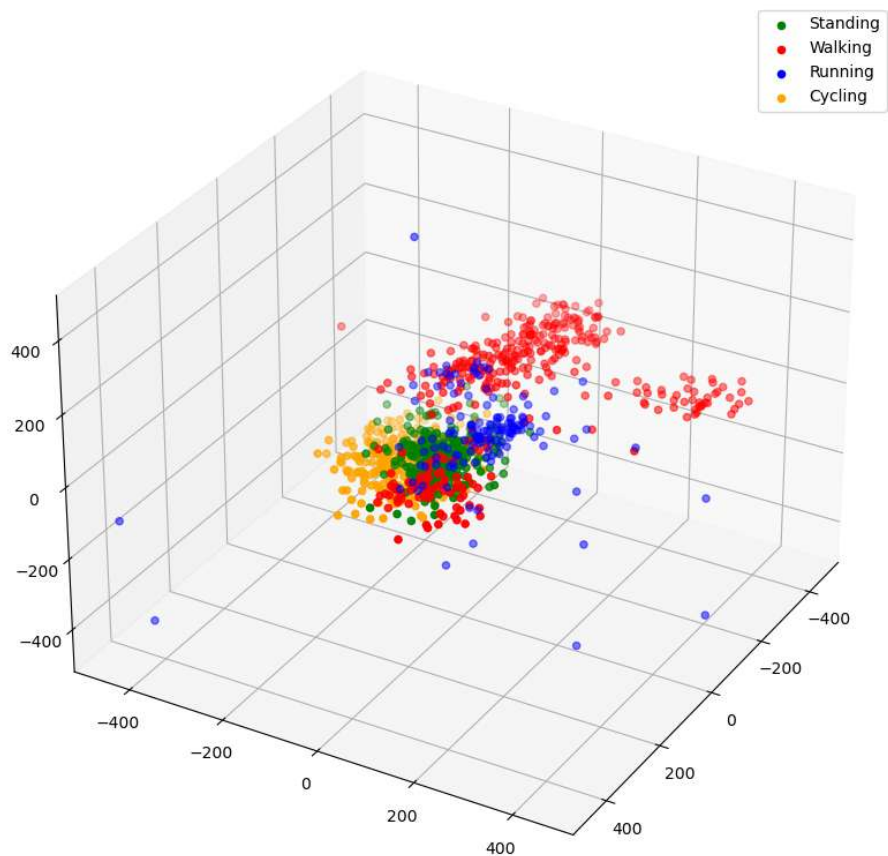


Рис. 5 Распределения данных в 3D

Из графиков следуют, что данные разделяются на различные кластеры, распределенные по нормальному закону, и предполагается, что они разделяются метрическими методами с гауссовским ядром.

Исследование моделей.

Для исследования был выбран случайный лес, т.к. он хорошо работает с любыми сложными наборами данных, линейная регрессия и гауссовский процесс, как достаточно простые классификаторы которые быстро учатся и соответственно полезно знать работают ли они на наборе данных, а также SVC с гауссовским ядром, т.к. по визуальному анализу предполагается, что данные распределены по нормальному закону.

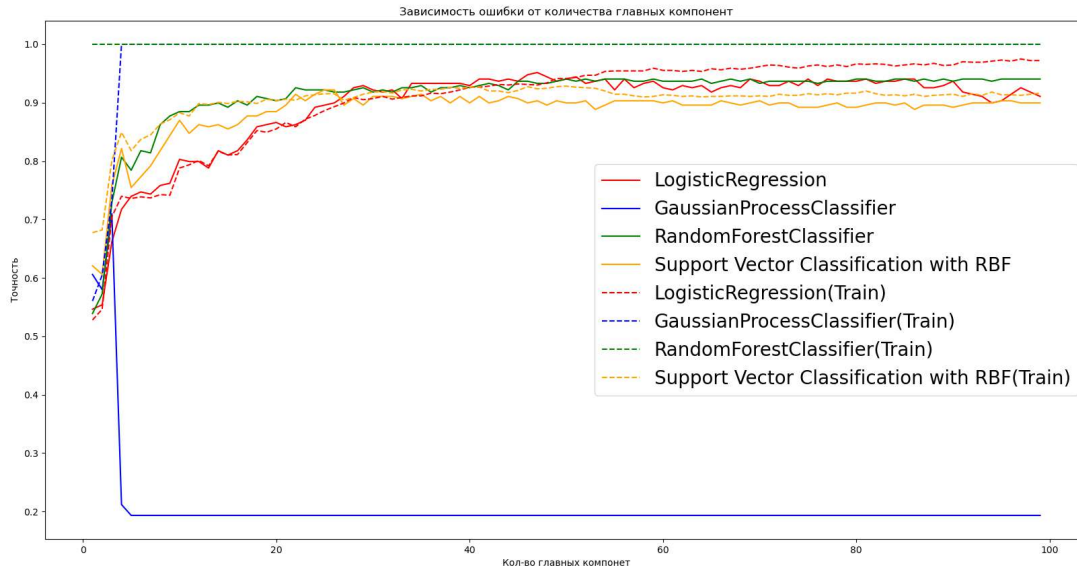


Рис. 6 Точность классификации выбранных методов в зависимости от количества главных компонент

Стабильнее всего обучается SVC, на нем отсутствует переобучение, и он показывает результаты сравнимые с остальными классификаторами, гауссовская переобучается, поэтому абсолютно не подходит под модель. Очень хорошо показывает себя случайный лес, что ожидаемо, т.к. модель данных сложная.

Классификация.

Выберем случайный лес, т.к. он лучше всего подходит под сложные модели данных, а также потому что он показывает лучшие результаты в классификации в соответствии эксперименту.

На основании экспериментального подбора наиболее оптимальным по точности и времени вычисления является $n = 300$ деревьев в лесу. Его и используем. На данном наборе данных полученный LNN-классификатор имеет точность в 95%.

Сравнение с другими классификаторами

Также были проведены сравнения с существующими методами классификации временных рядов. Метод случайного леса для временных рядов TimeSeriesForestClassifier, представленный в библиотеке Sktime показывает на том же наборе данных точность в 78%, что намного хуже LNN-классификатора.

6 Заключение

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>.
- [3] *А.Г.Сергеев* Лекции по Функциональному анализу // URL: https://mi-ras.ru/noc/13_14/2/sergeev/funkan.pdf.
- [4] *Гантмахер Ф. Р.* Лекции по аналитической механике: Учебное пособие для вузов // Под ред. Е. С. Пятницкого. — 3-е изд. — М.: Физматлит, 2005. — 264 с.
- [5] Датасет Physical Activity Monitoring // URL: <https://archive.ics.uci.edu/dataset/231/ramap2+physical+activity+monitoring>