Классификация траекторий физических систем с помощью лагранжевых нейронных сетей

Александр Иванович Богданов

Московский физико-технический институт

Курс: Моя первая научная статья

Эксперт: В. В. Стрижов

Консультант: С. К. Панченко

2023

Цель исследования

Цель

Исследование методов классификации траекторий физических систем.

Проблема

Стандартные методы классификации не учитывают физическую связь координат и динамики системы.

Решение

Моделирование лагранжиана с помощью лагранжевых нейронных сетей с последующим использованием полученных коэффициентов в качестве признаковых описаний для определения вида движения.

Публикации по теме

- Северилов Павел. Выбор оптимальной модели в задаче моделирования динамики физической системы. https://github.com/severilov/master-thesis.
- Miles Cranmer, Sam Greydanus, Stephan Hoyer, Peter Battaglia, David Spergel, and Shirley Ho. Lagrangian neural networks. arXiv preprint arXiv:2003.04630, 2020.
- Обработка датасета PAMAP2. https://github.com/andreasKyratzis/PAMAP2-Physical-Activity-Monitoring-Data-Analysis-and-ML/blob/master/pamap2.ipynb

Описание задачи классификации траекторий

Задана выборка с метками из n траекторий

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

- $\mathcal{D}_j = \{\mathbf{x}_i^{(j)}, \mathbf{y}_i^{(j)}\}_{i=1}^{m_j}$ – j-ая траектория,
- $\mathbf{x}_i^{(j)} = (\mathbf{q}_i^{(j)}, \dot{\mathbf{q}}_i^{(j)})$ – координаты j-ой траектории,
- $\mathbf{y}_i^{(j)} = \dot{\mathbf{x}}_i^{(j)} = (\dot{\mathbf{q}}_i^{(j)}, \ddot{\mathbf{q}}_i^{(j)})$ – динамика на 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 – метка j-ой траектории.

Задача регрессии динамики физической системы

Регрессионная модель выбирается из класса нейронных сетей

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

- $\mathbf{w} \in \mathbb{W}$ – параметры модели,
- $\hat{\mathbf{y}}_i^{(j)} = \mathbf{f}_j(\mathbf{x}_i^{(j)}, \mathbf{w}) \in \mathbb{R}^{2\times r}$ – предсказанная динамика j-ой траектории,
- $\mathbf{X}_j = \bigcup_{i=1}^{m_j} \mathbf{x}_i^{(j)}$ – матрица координат j-ой траектории,
- $\mathbf{Y}_j = \bigcup_{i=1}^{m_j} \mathbf{y}_i^{(j)}$ – матрица динамики j-ой траектории,
- $\hat{\mathbf{Y}}_j = \bigcup_{i=1}^{m_j} \hat{\mathbf{y}}_i^{(j)}$ – предсказанная матрица динамики j-ой траектории.

Задача минимизации квадратичной ошибки:

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

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

Задача классификации:

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

$\mathbf{w}_j^*$ – коэффициенты аппроксимированного лагранжиана j-ой траектории.
Для ее решения используются различные методы классификации, среди которых: логистическая регрессия, ядерный метод с гауссовским ядром, случайный лес.

Лагранжева динамика

Лагранжев формализм моделирует физическую систему с координатами траектории $\mathbf{x_t} = (\mathbf{q}, \dot{\mathbf{q}})$. Определяется функционал, называющийся действием:

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

показывающий путь, по которому координаты $\mathbf{x}_t$ пройдут из $\mathbf{x}_0$ в $\mathbf{x}_1$ в промежуток времени от t_0 до t_1. Путь минимизирует действие S, что приводит к уравнению Эйлера-Лагранжа

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

Ускорение каждой компоненты системы $\ddot{\mathbf{q}}$:

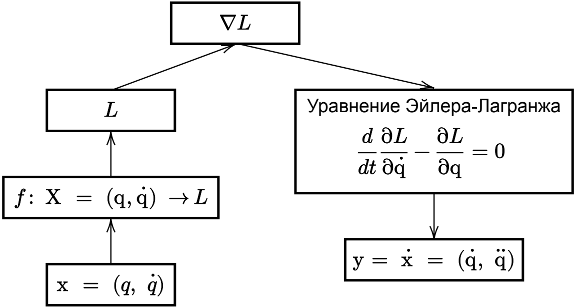
$$\ddot{\mathbf{q}} = \left(\nabla_{\dot{\mathbf{q}}}\nabla_{\dot{\mathbf{q}}}^T L\right)^{-1}\left[\nabla_{\mathbf{q}} L - \left(\nabla_{\dot{\mathbf{q}}}\nabla_{\mathbf{q}}^T L\right)\dot{\mathbf{q}}\right]$$

Лагранжева нейронная сеть

Нейронная сеть

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

Схема работы LNN для задачи моделирования динамики системы



Необходимо параметризовать нейронной сетью лагранжиан L, получить выражение ограничения Эйлера-Лагранжа и обратно распространить ошибку через полученные ограничения.

Подготовка данных

Для эксперимента используется датасет PAMAP2. Проблемы:

- Пропуски данных, связанные с тем, что датчик может пропустить такт. Для восстановления данных использовалась сплайн-интерполяция.
- Наличие скоростей и отсутствие ускорений и координат. Для получения ускорений используется аппроксимация 2-го порядка:

  $$f'(x_i) \approx \frac{f(x_{i+1}) - f(x_{i-1})}{2h}$$

  .

  Для получения координат используется метод Симпсона 3-го порядка аппроксимации::

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

Вычислительный эксперимент

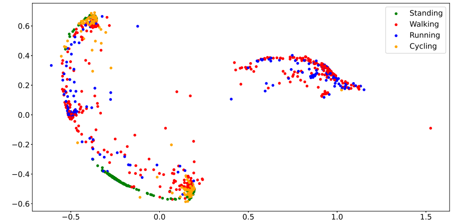
Датасет:

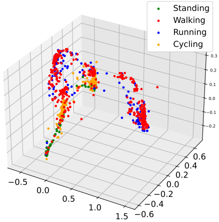
- Акселерометры:
  - На запястье рабочей руки
  - На рабочей ноге
  - На груди
- Частота акселерометров: 100 Гц.
- Количество классов: $K = 24$.

Для эксперимента использовались данные с рабочей ноги и 4 класса, на которых движение этой ноги различно.
Для каждого объекта из подвыборки строится лагранжиан.
Затем параметры этого лагранжиана используются в качестве признаков.

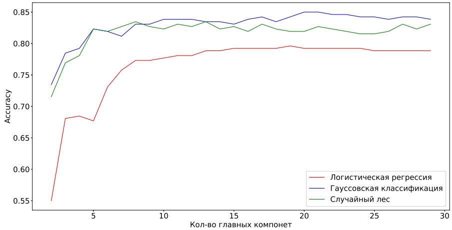
Анализ ошибки

Распределение признаков:





Accuracy:



- LogisticRegression: 79%
- GaussianProcessClassifier: 86%
- RandomForestClassifier: 85%

Заключение

- Подтверждена гипотеза компактности: векторы параметров, соответствующие траекториям различных классов, оказываются разделимы в своём пространстве.
- Предложен метод решения задачи классификации динамических систем с помощью лагранжевой нейронной сети.
- Проведен вычислительный эксперимент на датасате PAMAP2, в которых проведено сравнение различных методов классификации.

Планы

- Доделать эксперимент на весь датасет с использованием всех акселерометров и всех классов.
- Сделать свой датасет.