Классификация траекторий физических систем с помощью лагранжевых нейронных сетей

А. И. Богданов, С. К. Панченко, В. В. Стрижов
bogdanov.ai@phystech.edu; panchenko.sk@phystech.edu; strijov@phystech.edu

В работе решается задача классификации траекторий физических систем. Квазипериодическая динамика системы аппроксимируется лагранжианом, который восстанавливается по обобщённым координатам с помощью Лагранжевой нейронной сети. Показано, что для параметров таких сетей выполняется гипотеза компактности: векторы параметров, соответствующие траекториям различных классов, оказываются разделимы в своём пространстве. Проводится эксперимент на датасете PAMAP2, результаты которого подтверждают, что параметры лагранжевых нейронных сетей действительно являются информативным признаковым описанием для задачи классификации.

Ключевые слова: *физическая система, лагранжиан, лагранжева нейронная сеть*

1 Введение

Для моделирования динамики физических систем обычно используется Лагранжева динамика [3]. В соответствии с ней строится лагранжиан L, который определяется как разница между кинетической энергией (T) и потенциальной энергией (V) системы:

$$L = T - V.$$

Для получения уравнений движения нужно воспользоваться уравнением Эйлера-Лагранжа:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} - \frac{\partial L}{\partial \mathbf{q}} = 0,$$

полученным из принципа *наименьшего действия*.

Нейронные сети упрощают моделирование динамики физических систем и классификацию ее траекторий, так как для для их использования не требуются знания лагранжиана и решение сложных систем дифференциальных уравнений.

[2 предложения о лагранжевой сети]

Лагранжева нейронная сеть (LNN) [2] аппроксимируют обобщенный лагранжиан системы, с помощью которого, используя различные классификаторы, можно определить метку траектории исходя из предположения, что у различных меток сильно различаются коэффициенты лагранжиана.

[гипотеза]

[использование параметров лагранжевой сети в качестве признаков для решения задачи классификации]

[мб этот абзац в другой раздел?]

Для вычислительного эксперимента используется датасет PAMAP2 [1], который содержит траектории движений человека во время различных активностей и их метки.

2 Постановка задачи классификации траекторий физических систем

Задана выборка с метками из n траекторий:

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

где [список]

$$\mathbf{x}_i = (\mathbf{q}_i, \dot{\mathbf{q}}_i)$$

– координаты траектории движения физической системы,

$$\mathbf{y}_i = \dot{\mathbf{x}}_i = (\dot{\mathbf{q}}_i, \ddot{\mathbf{q}}_i)$$

– динамика движения физической системы, $\mathbf{q}_i \in \mathbb{R}^{r\times n}$, где r – количество координат, m – длина траектории, z_j – метка класса.

2.1 Задачи регрессии восстановления лагранжиана

Сведём задачу моделирования лагранжиана системы к задаче регрессии.

Регрессионная модель [для траектории D_j ?] выбирается из класса нейронных сетей:

$$\mathbf{f_j} \colon (\mathbf{w}, \mathbf{X}) \to \mathbf{y},$$

где $\mathbf{w} \in \mathbb{W}$ – параметры модели,

$$\hat{\mathbf{y}}_j = \mathbf{f}_j(\mathbf{X}_j, \mathbf{w}) \in \mathbb{R}^{2\times r\times m}, \quad \mathbf{X}_j = \bigcup_{i=1}^{m} \mathbf{x}_i, \quad \mathbf{y}_j = ...$$

.

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

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

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

2.2 Задачи классификации траекторий по лагранжианам

После решения задачи регрессии получаем задачу классификации:

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

где $\mathbf{w}_j^*$ – коэффициенты обобщенного лагранжиана j-ой траектории, z_j – метка класса этой траектории.

Для ее решения используются различные методы классификации, среди которых: LogisticRegression, GaussianProcessClassifier, RandomForestClassifier [перевести].

3 Лагранжевы нейронные сети

3.1 Лагранжева динамика

Лагранжев формализм [**?**, **?**, **?**, 2] моделирует физическую систему с координатами траектории $x_t = (q, \dot{q})$, которая начинается в состоянии x_0 и заканчивается в другом состоянии x_1. Определяется функционал, называющийся действием:

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

определяющий 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:

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

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

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 Лагранжева нейронная сеть

В работе [?] 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

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

На рисунке 3 представлены схемы работы базового решения нейронными сетями и LNN для задачи моделирования динамики системы.На основе базового решения основаны работы [**?**, **?**].

[b]0.49 ~~baseline_nn_scheme.png~~

Рис. 1 Базовое решение нейронными сетями

[b]0.49

~~lnn_scheme.png~~

Рис. 2 Решение LNN

Рис. 3 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 Вычислительный эксперимент

[Цель эксперимента: сравнение качества моделирования механической системы различными нейронными сетями, качество будем сверять с аналитическим решением]

4.1 Данные

4.2 Обработка данных

5 Результаты

6 Заключение

Литература

[1] Датасет pamap2. *https://www.kaggle.com/datasets/phamson/pamap2*.

[2] Miles Cranmer, Sam Greydanus, Stephan Hoyer, Peter Battaglia, David Spergel, and Shirley Ho. Lagrangian neural networks. *arXiv preprint arXiv:2003.04630*, 2020.

[3] L.D. Landau, E.M. Lifshitz, J.B. Sykes, and J.S. Bell. *Mechanics: Volume 1*. Course of theoretical physics. Elsevier Science, 1976.

Isachenko R.V., Strijov V.V. Quadratic programming feature selection for multicorrelated signal decoding with partial least squares // Expert Systems with Applications, Volume 207, 30 November 2022, 117967