Классификация траекторий динамических систем с помощью физически-информированных нейросетей

Терентьев Александр

Московский физико-технический институт,
Физтех-школа прикладной математики и информатики
Кафедра интеллектуальных систем
Научный руководитель: д.ф.-м.н. Стрижов В.В.

18.05.2024, Москва

Цель работы

Проблема

Существующие методы классификации многомерных рядов не имеют априорных знаний о динамике, изучаемых систем. Методы используют слишком громоздкие модели. Либо они требуют большой обучающей выборки для выделения необходимых признаков и свойств изучаемых систем.

Цель

Целью работы является предложить метод для классификации многомерных временных рядов, являющимися траекториями динамических систем, использующие априорную информацию о физической природе рядов.

Идея

Каждому временному ряду сопоставить динамическую систему. Классифицировать не траектории, а динамические системы.

Постановка задачи классификации многомерных рядов

Дано

Дана выборка $D = \{(X_i, y_i)\}_{i=1}^N$, где X_i – траектории размерности r и длинной T, $y_i \in \overline{1, K}$ – метка i-ой траектории
Траекторией размерности r и длинной T называется $X = [X^1, X^2, , , , , X^T]$ такой, что $X^j = (\mathbf{q}^j), \dot{\mathbf{q}}^j), \dot{\mathbf{q}}^j \in \mathbb{R}^r$ – скорость в j-ый момент времени, $\mathbf{q}^j \in \mathbb{R}^r$ – координата в j-ый момент времени

Найти

Метод классификации $p(\hat{y}|X, D)$, где $X = \{(X_i)\}_{i=1}^N$ – набор длинны N траекторий размерности r и длинной t, D – данная обучающая выборка, $\hat{y} \in \overline{1, K}^N$ – предсказанные метки классов

Критерий

Модели сравниваются с помощью метрики Acuracy = $\frac{1}{N}\sum_{i=1}^N [\hat{y}_i = y_i]$, $\hat{y}_i$ – предсказанные моделью метки траектории, $\hat{y}$ – класс, которому принадлежит траектория

Постановка задачи восстановления трактории

Дано

Дана выборка $D = \{(X_i, \dot{X}_i)\}_{i=1}^N$, где X – траектории размерности r и длинной T, $\mathbf{y}_i = \dot{X}_i = \ddot{\mathbf{q}}_i \in \mathbb{R}^r$ – ускорение в i-ый момент времени.

Найти

Требуется найти функцию $\hat{\mathbf{y}} = f(\mathbf{X} = (\mathbf{q}, \dot{\mathbf{q}})|D)$, где D – данная обучающая выборка, $\hat{Y} = \hat{\dot{X}} = \{\hat{\mathbf{y}}_i = \hat{\ddot{\mathbf{q}}}_i\}_{i=1}^N$ – предсказанная динамика траектории

Критерий

В качестве функции потерь берется средняя квадратичная ошибка

$$\mathcal{L}(\hat{\mathbf{y}}, \mathbf{y}) = \frac{1}{T} \sum_{i=1}^{T} \|\hat{\mathbf{y}}_i - \mathbf{y}_i\|_2^2,$$

Лагарнжева нейронная сеть

$$\mathbf{x}_i = (\mathbf{q}_i, \dot{\mathbf{q}}_i) \xrightarrow{g:\,(\mathbf{x}|\mathbf{w})\to L} L \xrightarrow{\nabla L} \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} = \frac{\partial L}{\partial \mathbf{q}} \xrightarrow{\ddot{\mathbf{q}}_i = f(\mathbf{q}_i, \dot{\mathbf{q}}_i)} \mathbf{y}_i = \ddot{\mathbf{q}}_i$$

Рис.: Схема нейронной сети

Параметризация

$g\colon (\mathbf{x}|\mathbf{w}) \to L$ - полносвязная нейронная сеть с функцией активации SoftPlus, w - параметры нейронный сети Данная сеть по координатам и скоростям восстанавливает значение лагранжиана Иными словами g – приближенный лагранжиан системы

Система для нахождения ускорений

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

Модификация сети

В исходной архитектуре сети в каждой точки траектории мы решаем СЛАУ, проблема в том, что у матрицы $H_L(\mathbf{q}, \dot{\mathbf{q}}) = \nabla_{\dot{\mathbf{q}}\dot{\mathbf{q}}} L$ собственные значения могут быть сколь угодно маленькими.

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

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

$$H\ddot{\mathbf{q}} = b$$

Тогда изменим функцию потерь так, чтобы штрафовать за собственные значения матрицы H меньше 1, а вместо разности ускорений возьмем невязку для полученной СЛАУ

$$\mathcal{L}^{mod}(\mathbf{w}) = \frac{1}{T} \sum_{i=1}^{T} \|\hat{\mathbf{b}}_i - \mathbf{b}_i\|_2^2 + \alpha \text{act}(\beta(\lambda(H_{\ddot{q}}) - 1)).$$

Результат регуляризации

Возьмем синтетический набор данных для траекторий двойного маятника и построим график остатков в случаи регуляризации и без.

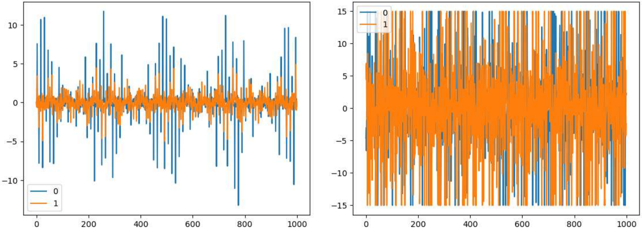


Рис.: График остатков для траектории двойного маятника, слева с регуляризацией, справа без

Теоретические результаты о эквивалентности нахождения минимума невязки и отклонения ускорений

Lemma

Если лагранжианы заданы на компакте, то существуют неотрицательные числа a, b *такие, что* $detH \geq a$, *а собственные значения матрицы* H *не меньше* b

$$A(L) = \nabla_q L - \left(\nabla_{\dot{\mathbf{q}}\mathbf{q}} L\right) \dot{\mathbf{q}}, H(L) = \nabla_{\dot{\mathbf{q}}\dot{\mathbf{q}}} L$$

$$\|L\|_L = \|(A(L), H(L))\|_2$$

Lemma

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

Аппроксимация нормы

$$A(L) = \nabla_q L - \left(\nabla_{\dot{\mathbf{q}}\mathbf{q}} L\right) \dot{\mathbf{q}}, H(L) = \nabla_{\dot{\mathbf{q}}\dot{\mathbf{q}}} L$$

$$\|L\|_L = \|(A(L), H(L))\|_2$$

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

$$C \cdot \overline{|A(L)\left(\mathbf{q}, \dot{\mathbf{q}}\right)|^2 + \|H(L)\left(\mathbf{q}, \dot{\mathbf{q}}\right)\|_2^2}$$

$$A(L) = \nabla_q L - \left(\nabla_{\dot{\mathbf{q}}\mathbf{q}} L\right) \dot{\mathbf{q}},\, H(L) = \nabla_{\dot{\mathbf{q}}\dot{\mathbf{q}}} L$$

$$X \xrightarrow{LNN} \hat{L} = g(x|w) \xrightarrow{L_N^i = A(L)(\mathbf{q}_i, \dot{\mathbf{q}}_i)} L_N \in \mathbb{R}^N \xrightarrow{\text{classifier}: R^N \to K} \hat{y}$$

$$Q_N \sim U([-L, L]^{2r})$$

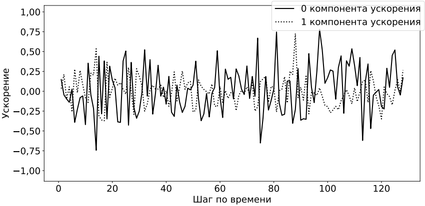
Рис.: Схема нейронной сети

Аппроксимация нормы

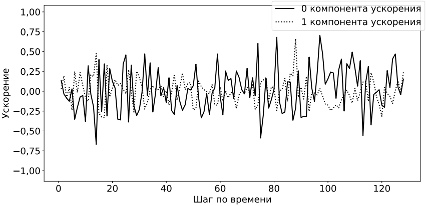
$$\overline{|A(L)\,(\mathbf{q}, \dot{\mathbf{q}})\,|^2 + \|H(L)\,(\mathbf{q}, \dot{\mathbf{q}})\,\|_2^2}$$

$$\|H(L)\,(\mathbf{q}, \dot{\mathbf{q}})\,\|_2^2 \approx const$$

Вычислительный эксперимент



(a) Исходный ряд



(b) Предсказанный ряд

Рис.: Временной ряд зависимости ускорения от времени для тестовой выборки

Кривые обучения

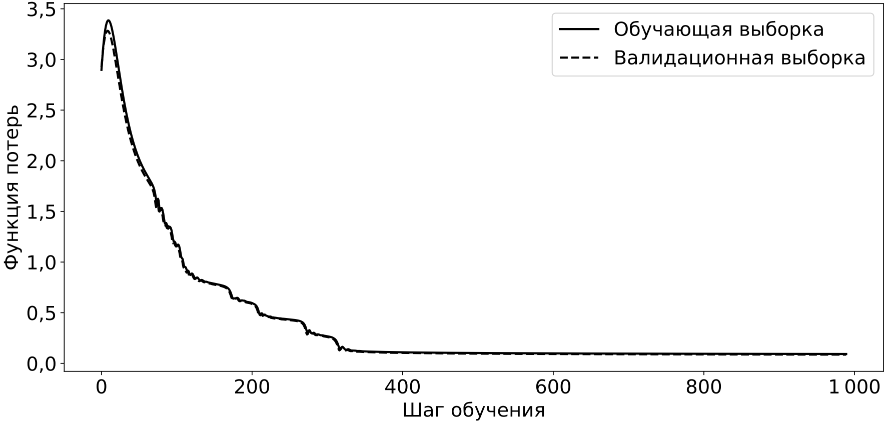


Рис.: График обучения модели

Вычислительный эксперимент

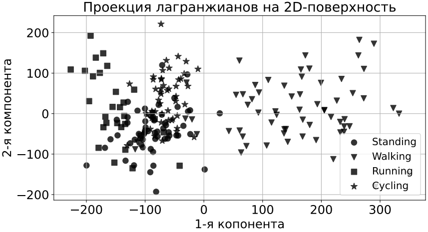
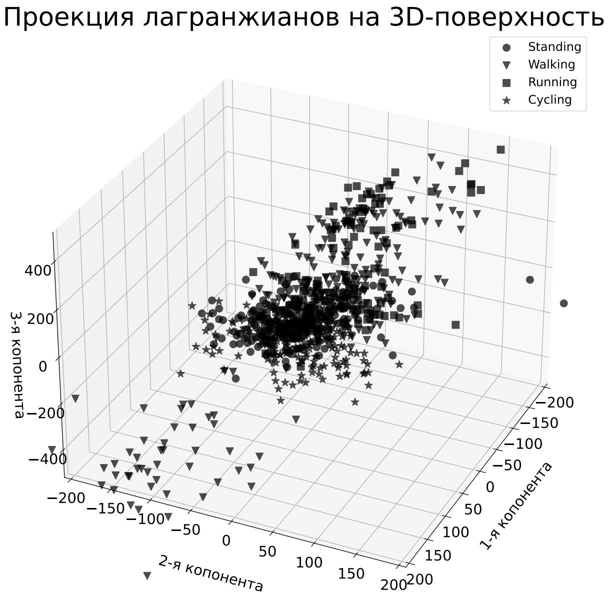


Рис.: Распределения данных в 2D

Вычислительный эксперимент



Вычислительный эксперимент

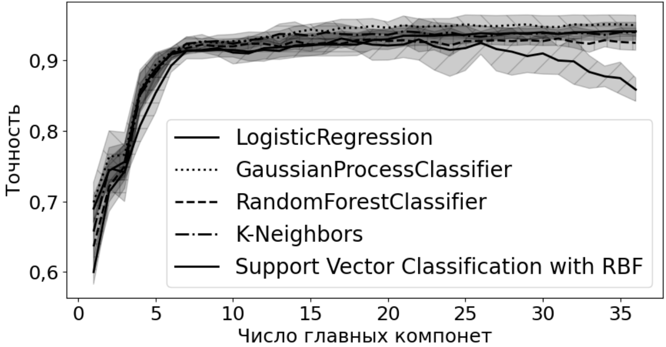


Рис.: Точность классификации выбранных метод в зависимости от количества главных компонент

Вывод

Представлен метод классификации траекторий по параметризованным лагранжианам, аппроксимирующим ланграпжианы динамических систем, для которых траектории получены. Получена модифицированная LNN-сеть дающая вычислительно устойчивую аппроксимацию лагранжиана, которая в дальнейшем исследуется как функция в нормированном пространстве. Исследован способ проекции векторов данного пространство на конечномерное пространство, такая, что разность норм в исходном пространстве и спроецированном не превышает заранее заданную величину. Для уменьшения размерности применен метод главных компонент. вычислительный эксперимент показал, что для данной выборки достаточно оставить $N = 10$ главных компонент и при увеличение числа которых метод не улучшается. Для классификации полученных объектов взят SVC с гауссовским ядром, дающий точность на валидационной выборке $Accuracy = 0.95 \pm 0.02$