

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: к.ф.-м.н. Исаченко Р.В.

15.06.2024, 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 длиной T , $y_i \in \overline{1, K}$ – метка i -ой траектории. Траекторией размерности r и длиной T назовем $X = [X_1, X_2, \dots, X_T]$ такой, что $X_i = [\mathbf{q}_i, \dot{\mathbf{q}}_i]$, $\dot{\mathbf{q}}_i \in \mathbb{R}^r$ – скорость в i -ый момент времени, $\mathbf{q}_i \in \mathbb{R}^r$ – координата в i -ый момент времени.

Найти

Требуется найти метод классификации $p(\hat{y}|X, D) = p(\hat{y}|L, D)p(L|X)$, где $\hat{y} \in \overline{1, K}^N$ – предсказанные метки классов, $p(L|X)$ – метод сопоставления траектории X лагранжиану L . В работе в качестве метода рассматривается точечная оценка $p(L|X) = [L = \hat{L}]$.

Критерий

Модели сравниваются с помощью метрик Accuracy и F1Macro

Постановка задачи восстановления траектории

Дано

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

Найти

Требуется найти функцию $p(\hat{\mathbf{y}}|X, D) = p(\hat{\mathbf{y}}|X, L)p(L|D)$, где D – данная обучающая выборка, $\hat{Y} = \hat{X} = \{\hat{\mathbf{y}}_i = \hat{\ddot{\mathbf{q}}}_i\}_{i=1}^N$ – предсказанная динамика траектории, $L \in Q$, где Q – это семейство рассматриваемых функций

Критерий

В качестве функции потерь берется средняя квадратичная ошибка

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

Лагранжева нейронная сеть

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

Рис.: Схема нейронной сети

Параметризация

$p(L|D)$, $L \in Q = \{\hat{L}: (\mathbf{x}|\mathbf{w}) \rightarrow L_{\mathbf{x},\mathbf{w}}\}$ \hat{L} - полносвязная нейронная сеть с функцией активации SoftPlus, w - параметры нейронной сети.

$p(L|D) = [L = \hat{L}]$, где \hat{L} – ОМП лагранжиана системы, полученная из $\mathcal{L}(\hat{\mathbf{y}}, \mathbf{y}) \rightarrow \min$

Система для нахождения ускорений

$$p(\hat{\mathbf{y}}|X, L) : (\nabla_{\dot{\mathbf{q}}\dot{\mathbf{q}}} L) \ddot{\mathbf{q}} = [\nabla_{\mathbf{q}} L - (\nabla_{\dot{\mathbf{q}}\mathbf{q}} L) \dot{\mathbf{q}}] .$$

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.

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

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

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

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

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

Идея классификатора

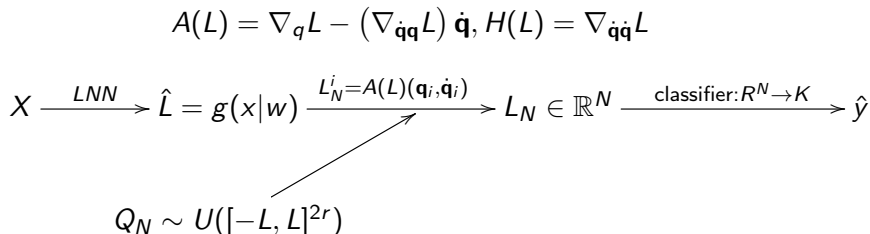


Рис.: Схема нейронной сети

Аппроксимация нормы

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

Эквивалентность нахождения минимума невязки и отклонения ускорений

Норма

$$\begin{aligned} A(L) &= \nabla_q L - (\nabla_{\dot{q}q} L) \dot{q}, H(L) = \nabla_{\dot{q}\dot{q}} L \\ \|L\|_L &= \|(A(L), H(L))\|_2 \\ A(L)(q, \dot{q}) &= H(L)(q, \dot{q})\ddot{q} \end{aligned} \quad (1)$$

Теорема(Терентьев, 2024)

$\|\delta L\|_L = 0 \Leftrightarrow$ п.в. $\delta\ddot{q} = 0$, где $\delta L = L' - L$, $\delta\ddot{q} = \ddot{q}' - \ddot{q}$, где $L', L \in \mathcal{X}$ являются вариациями лагранжиана и ускорения на множестве \mathcal{X} , на котором функциональное уравнение 1 относительно L имеет единственное решение, и $\forall L \in \mathcal{X} : \mathbf{H}(L) \neq 0,.$

Эквивалентность исходной оптимизационной задачи и задачи с ограничениями

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

Теорема(Терентьев, 2024)

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

Аппроксимация нормы

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

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

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

Теорема(Терентьев, 2024)

Исходная норма $\|\cdot\|_L$ с любой наперед заданной точностью ϵ приближается l_2 -нормой, при стремлении числа сэмплов N и меры пространства из которого берут сэмплы Ω к бесконечности

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

Эквивалентность приближенной задачи исходной

Теорема(Терентьев, 2024)

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

Исследования проводились на наборе данных Physical Activity Monitoring(PAMAP2) Набор данных содержит записи с трех наборов гироскопов и акселерометров: закрепленных на запястье преобладающей руки, закрепленных на груди, закрепленных на локте преобладающей руки. Число испытуемых: $M = 9$. Число видов активностей(классов): $K = 24$, каждая активность длилась 2-4 минуты , частота сэмплирования 100 Гц.

Остатки траектории в зависимости от регуляризации

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

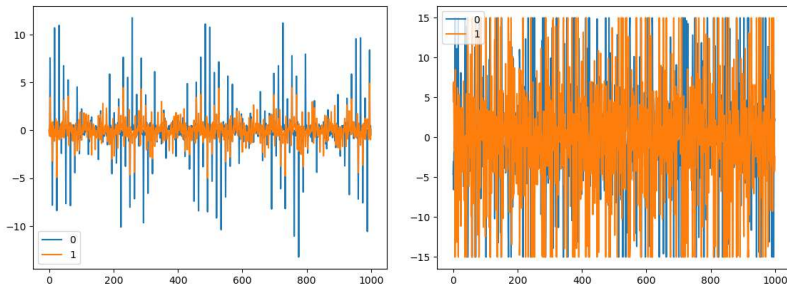
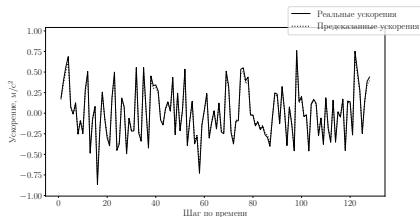
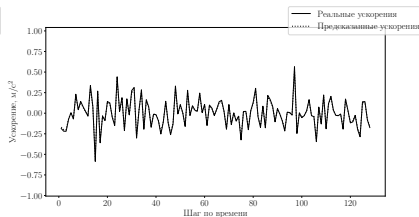


Рис.: График остатков для траектории двойного маятника, слева с регуляризацией, справа без

Восстановление динамики системы



(a) Исходный ряд



(b) Предсказанный ряд

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

Кривые обучения

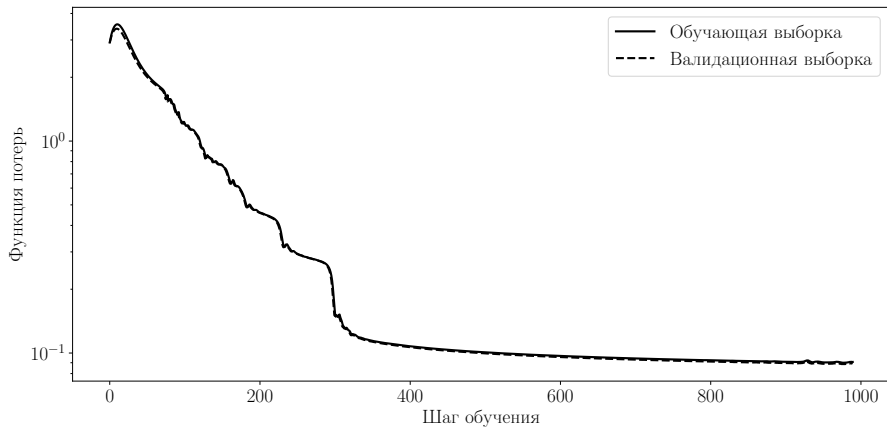


Рис.: График обучения модели

Распределение классов в пространстве Лагранжианов

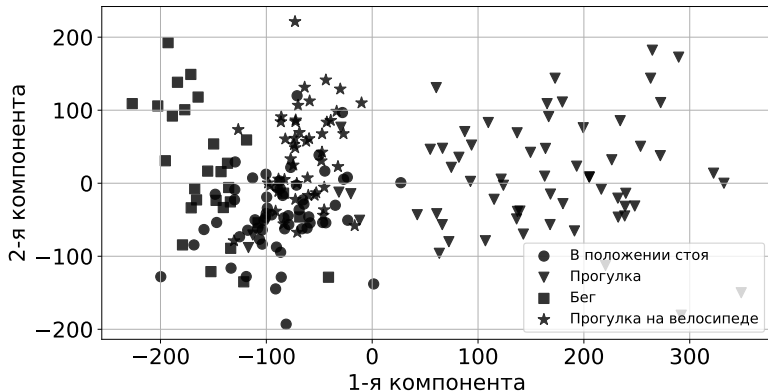


Рис.: Проекция пространства лагранжианов на плоскость

Распределение классов в пространстве Лагранжианов

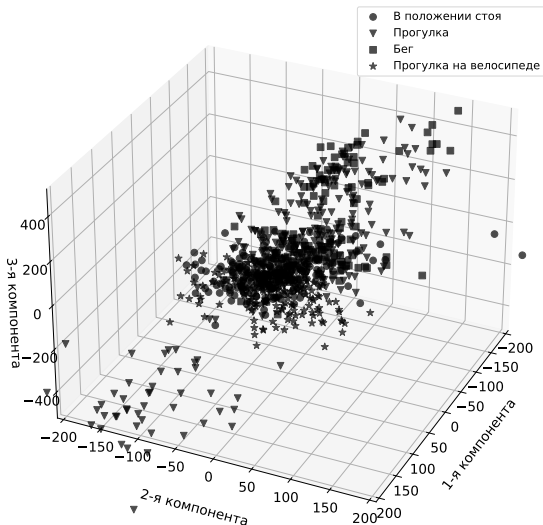


Рис.: Проекция пространства лагранжианов в трехмерное пространство

Результат работы классификаторов

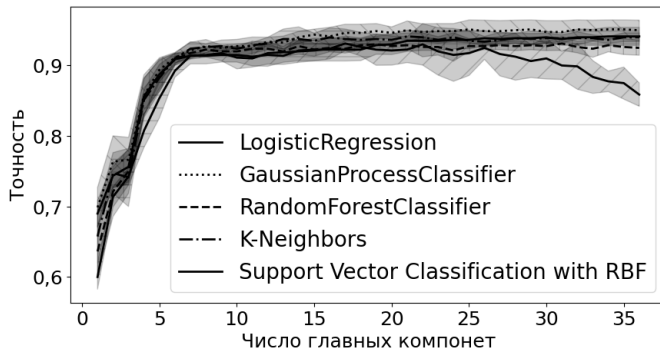


Рис.: Точность классификации выбранных метод в зависимости от количества главных компонент

Результат работы классификаторов

Классификатор	Метрика		
	Accuracy	Balanced-accuracy	F1 Macro
Логистическая регрессия	$0,927 \pm 0,014$	$0,924 \pm 0,13$	$0,927 \pm 0,14$
Гауссовский процесс	$0,946 \pm 0,010$	$0,941 \pm 0,011$	$0,946 \pm 0,010$
Случайный лес	$0,932 \pm 0,007$	$0,928 \pm 0,008$	$0,933 \pm 0,008$
К-ближайших соседей	$0,939 \pm 0,009$	$0,935 \pm 0,010$	$0,940 \pm 0,008$
SVC с гауссовским ядром	$0,933 \pm 0,012$	$0,927 \pm 0,013$	$0,933 \pm 0,011$

Таблица: Результат классификаторов на предложенной векторизации данных

Выносятся на защиту

1. Предложен метод физико-информированного подхода к классификации многомерных временных рядов, на основе классификации систем их порождающие.
2. Предложен метод оценки лагранжиана системы на основе LNN-сетей и регуляризация устраняющая нефизические колебания системы.
3. Доказано, что, если исходные лагранжианы отделимы, то и их оценки тоже отделимы.
4. Доказана эквивалентность задачи нахождения минимума невязки и отклонения ускорений, и показаны достаточные условия для этого.
5. Доказано, что выбранная проекция пространства лагранжианов приближает норму с любой наперед заданной точностью.
6. Исследованы методы метрической классификации в данном пространстве и показано, что линейные классификаторы дают лучшие метрики классификации.