

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

## Найти

Требуется найти функцию  $p(\hat{\mathbf{y}}|X, D) = p(\hat{\mathbf{y}}|X, L)p(L|D)$ , 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

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

Рис.: Схема нейронной сети

## Параметризация

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

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

# Теоретические результаты о эквивалентности нахождения минимума невязки и отклонения ускорений

## Lemma

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

$$A(L) = \nabla_q L - (\nabla_{\dot{q}q} L) \dot{q}, H(L) = \nabla_{\dot{q}\dot{q}} L$$
$$\|L\|_L = \|(A(L), H(L))\|_2$$

## Lemma

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

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

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

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

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



## Идея классификатора

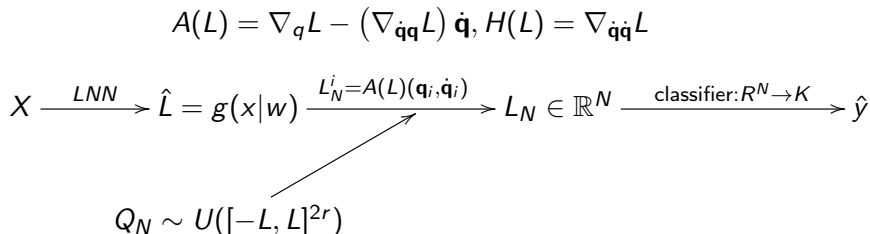


Рис.: Схема нейронной сети

## Аппроксимация нормы

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

# Результат регуляризации

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

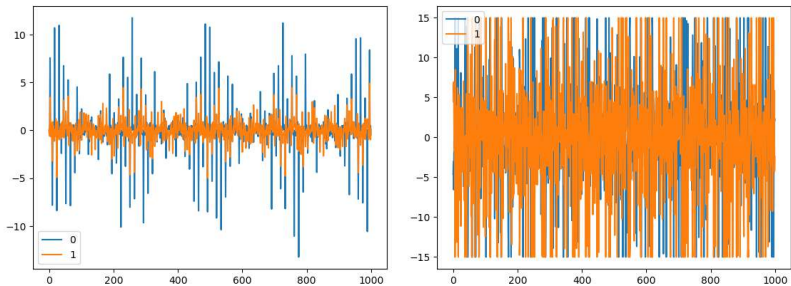
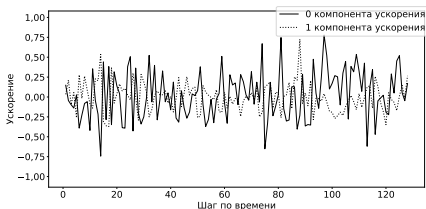
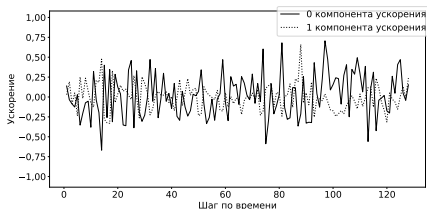


Рис.: График остатков для траектории двойного маятника, слева с регуляризацией, справа без

# Вычислительный эксперимент



(a) Исходный ряд



(b) Предсказанный ряд

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

# Кривые обучения

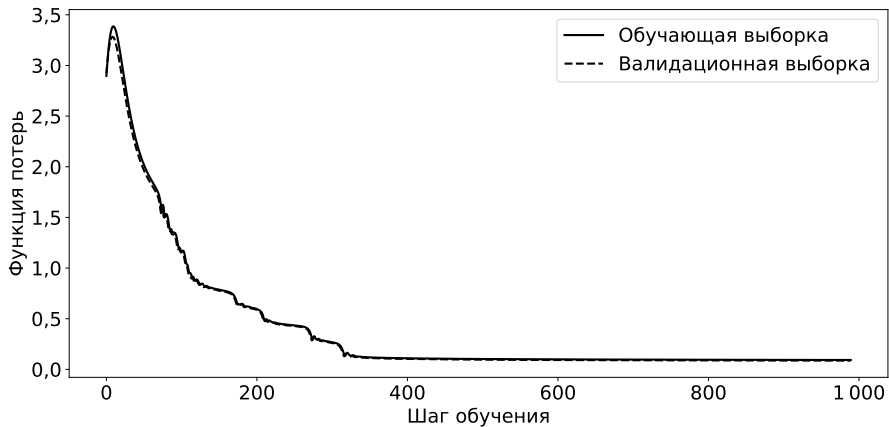


Рис.: График обучения модели

# Вычислительный эксперимент

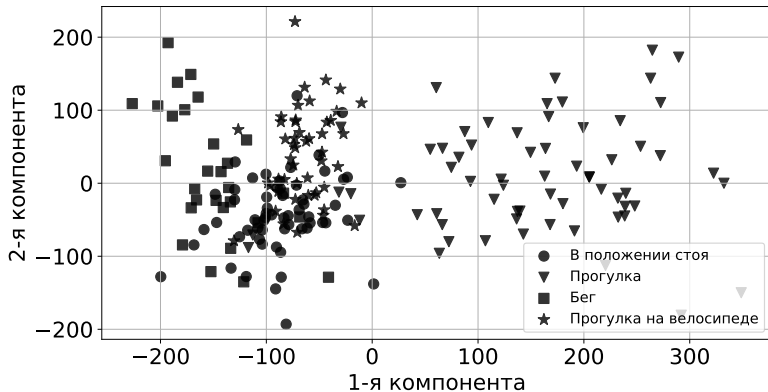


Рис.: Распределения данных в 2D

# Вычислительный эксперимент

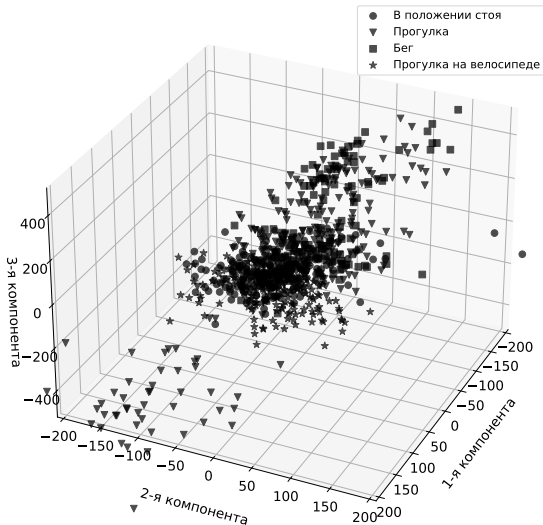


Рис.: Распределения данных в 3D

# Вычислительный эксперимент

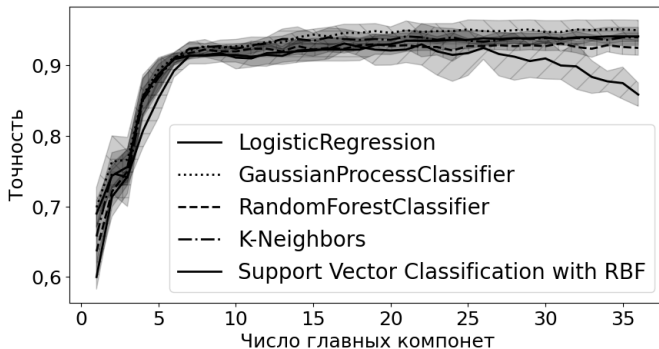


Рис.: Точность классификации выбранных метод в зависимости от количества главных компонент

# Вычислительный эксперимент

Классификатор	Метрика		
	Accuracy	Balanced-accuracy	F1 Macro
Логистическая регрессия	$0,927 \pm 0,014$	$0,924 \pm 0,13$	$0,927 \pm 0,14$
<b>Гауссовский процесс</b>	<b><math>0,946 \pm 0,010</math></b>	<b><math>0,941 \pm 0,011</math></b>	<b><math>0,946 \pm 0,010</math></b>
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, идея которого заключается в классификации не самих рядов, а систем, порождающих данные ряды.
2. Предложен метод оценки лагранжиана системы на основе LNN-сетей и регуляризация устраняющая нефизические колебания системы.
3. Показано, что в предложенном пространстве лагранжианы линейно разделимы.
4. Исследованы методы метрической классификации в данном пространстве и показано, что линейные классификаторы дают лучшие метрики классификации.