

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ  
МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ (государственный университет)  
ФИЗТЕХ-ШКОЛА ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ  
КАФЕДРА «ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ»

Тихонов Денис Максимович

**Выбор моделей прогнозирования  
квазипериодических временных рядов**

03.04.01 — Прикладные математика и физика

Выпускная квалификационная работа магистра

**Научный руководитель:**

д. ф.-м. н. Стрижов Вадим Викторович

Москва

2022

## Оглавление

|   |    |
|---|----|
| Введение . . . . .  | 4  |
| 1. Постановка задачи выбора модели прогнозирования . . . . .              | 7  |
| 2. Авторегрессионные модели . . . . .                                     | 9  |
| 3. Постановка задачи классификации по модели аппроксимации . . . . .      | 10 |
| 4. Модель фазовой траектории . . . . .                                    | 12 |
| 5. Модель классификации весовых коэффициентов . . . . .                   | 15 |
| 6. Модель классификации точек фазовой траектории . . . . .                | 17 |
| 7. Выбор моделей прогнозирования с помощью сферических гармоник . . . . . | 18 |
| 8. Альтернативные методы выбора моделей прогнозирования . . . . .         | 22 |
| 9. Эксперимент по классификации сферическими гармониками . . . . .        | 23 |
| 10. Эксперимент по выбору модели прогнозирования . . . . .                | 27 |
| Заключение . . . . .  | 31 |
| Список литературы . . . . .   | 32 |

## Аннотация

Решается задача выбора модели прогнозирования квазипериодических временных рядов. Классические методы оценки эффективности прогнозирования не учитывают предысторию и априорное знание о периодичности временных рядов. Для учета этой информации применяется модель фазовой траектории исследуемого временного ряда. Для перехода в фазовое пространство временной ряд векторизуется с помощью метода задержек. Для понижения сложности модели в фазовом пространстве выбирается подпространство. Проекция фазовой траектории на сферу в полученном подпространстве аппроксимируется линейной комбинацией сферических гармоник. Полученная модель оценивает вероятность принадлежности точек фазовой траектории исследуемому временному ряду. Дополнительно свойства модели фазовой траектории были исследованы в задаче классификации. По полученным оценкам вероятности предлагается метод выбора наилучшей модели прогнозирования. Вычислительный эксперимент проведен на измерениях акселерометра мобильного устройства с тремя классами движений человека.

**Ключевые слова:** *выбор моделей, временные ряды, прогнозирование, аппроксимация, классификация, фазовое пространство, сферические гармоники.*

## Введение

Исследуется проблема выбора алгоритмов прогнозирования квазипериодических временных рядов. Квазипериодический временной ряд незначительно меняет характерные частоту и период во времени. Это создает неточность оценки прогнозирования для классических методов, опирающихся в предположении об i.i.d.

Методы оценки качества прогнозирования используются для решения двух основных задач. Первая задача — это оценка предсказательной способности и стабильности модели. Вторая задача — это выбор модели на основе полученных оценок качества из *множества возможных моделей*. *Множество возможных моделей* — это конечное множество моделей из разных семейства и/или с разными настройками параметров.

Значительный объем работ посвящен первой задаче, например [1, 2, 3]. В данной работе решается вторая задача. Методы наиболее точно оценивающие предсказательную способность не обязательно будут иметь лучшую ранжирующую способность для выбора моделей. Эта идея проиллюстрирована на Рис. 1.

На Рис. 1 приведен пример различия оценок. Метода  $E_1$  наиболее точно предсказывает ошибку на данных из будущего. Ранжирование модели по методу  $E_1$  отличается от истинного рейтинга предложенных моделей  $g_1 > g_2 > g_3 > g_4$ . Метод  $E_2$  имеет худшие предсказания ошибки моделей, но лучшее ранжирует моделей. Этот пример показывает, что метод  $E_1$  лучше подходит для оценка предсказательной способности и стабильности модели, а метод  $E_2$

для выбора модели.

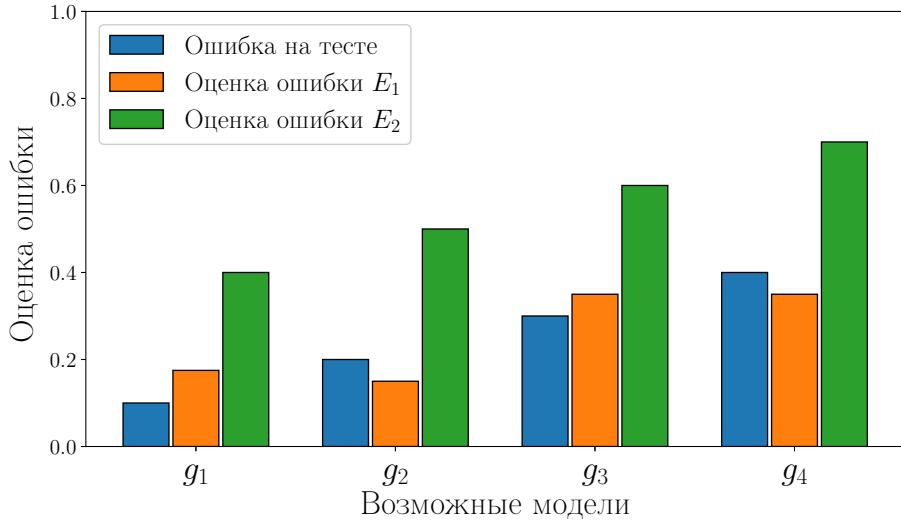


Рис. 1. Выбор модели для прогнозирования временных рядов

Цель данной работы — используя априорное знание о периодичности предложить метод выбора модели. Для этого производится переход в пространство фазовых траекторий. Переход осуществляется методом задержек [8]. Метод задержек используется при анализе нестационарных временных рядов. Например, в методе сингулярного спектрального анализа [7] разложения на компоненты и прогноз основаны на траекторной матрице. Она позволяет перейти от скалярного временного ряда к многомерному представлению. Метод задержек так же получили широкое распространение в анализе нелинейных динамических систем [11, 8]. Избыточная размерность фазового пространства [7, 9, 12] приводит к неустойчивости исследуемых моделей и избыточно сложному описанию временного ряда. Для понижения размерности фазового пространства предлагается

использовать метод главных компонент [5, 10].

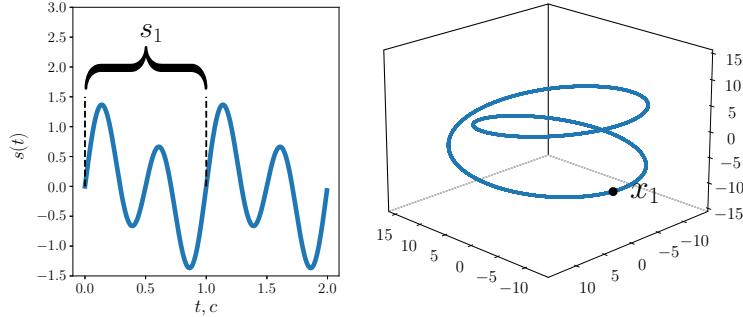


Рис. 2. Слева: сегмент временного ряда, справа: фазовая траектория.

На Рис. 2 показан сегмент синтетического временного ряда и его фазовая траектория в пространство размерности три. Фазовая траектория получена с помощью метода задержек и метода главных компонент над матрице задержек.

В выбранном фазовом подпространстве малой размерности фазовая траектория проецируется на  $p$ -мерную единичную сферу. Полученную на поверхности сферы функцию предлагается аппроксимировать линейной комбинацией сферических гармоник. Полученные коэффициенты в дальнейшем используются как признаковое описание в задаче классификации. Значения модели интерпретируются как вероятность принадлежности точки фазового пространства выбранному временному ряду. На основе оценок вероятностей принадлежности предлагается метод выбора модели.

## 1. Постановка задачи выбора модели прогнозирования

Выбор модели — это процесс использования доступных обучающих данных для выбора модели прогнозирования  $g_{\text{selected}}$  среди множества  $m_g$  возможных моделей  $G = \{g_1, \dots, g_{m_g}\}$ . На Рис. 3 показана схема выбора модели. Метод  $E$  оценивает модели  $g \in G$ , используя имеющиеся данные и функцию потерь  $L$ .

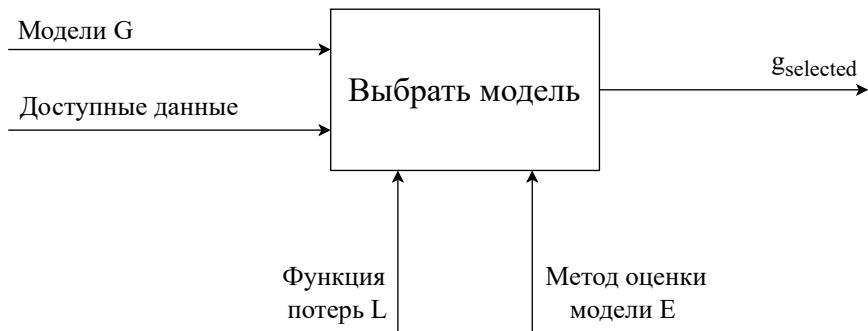


Рис. 3. Схема выбора модели

Цель — найти и выбрать  $g^* \in G$ , которая является моделью с наилучшей предсказательной способностью в будущем. Задача выбора модели:

$$g_{\text{selected}} \in \arg \min_{g \in G} E(g, \mathcal{D}, L), \quad (1)$$

где  $L$  — функция потерь оценивающая качество модели, а  $E$  — метод метод выбора модели,  $\mathcal{D}$  — имеющийся набор данных о временном ряде.

Метод  $E$  ранжирует модели на имеющихся данных. По полученному рейтингу выбирается  $g_{\text{selected}}$ . В идеальном случае  $g_{\text{selected}}$  является моделью  $g^*$  с наилучшей предсказательной способностью

в будущем. На практике это выполняется не всегда, поскольку ранжирование метода  $E$  оценочное.

Цель этой работы — предложить метод оценки  $E$  для поиска  $g^*$  с априорным знанием о периодичности временного ряда. Эта задача также представима в виде анализа ранжирующей способности метода  $E$ , в котором  $g_{\text{selected}}$  — это модель с наивысшим рейтингом.

## 2. Авторегрессионные модели

В качестве прогностических моделей были выбраны авторегрессионные линейные модели. Для оценки таких моделей существуют подходы, такие как информационные критерии, позволяющие ранжировать модели и находить парето оптимальные. **Модель AR-авторегрессии** порядка  $m$  записывается следующим образом:

$$s_t = w_0 + \sum_{i=1}^m w_i s_{t-i},$$

где  $w_0$  — константа,  $w_1, \dots, w_m$  — параметры модели.

Оптимальные параметры  $\hat{\mathbf{w}}$  модели авторегрессии определяются минимизацией функции ошибки, например среднеквадратичной ошибки (MSE):

$$\hat{\mathbf{w}} = \arg \min_{w \in \mathbb{R}^m} \sum_{i=1}^n \|s_i - \hat{s}_i\|_2^2.$$

**Модель авторегрессионного скользящего среднего** за-  
действует модель скользящего среднего порядка  $q$ :

$$x_t = \underbrace{w_0 + \sum_{i=1}^m w_i s_{t-i}}_{AR(m)} + \underbrace{\sum_{i=1}^q \theta_i \varepsilon_{t-i}}_{MA(q)},$$

где  $w_0$  — константа,  $w_1, \dots, w_m, \theta_1, \dots, \theta_q$  — параметры модели,  
 $\varepsilon_{t-1}, \dots, \varepsilon_{t-q}$  — ошибки.

### 3. Постановка задачи классификации по модели аппроксимации

Пусть заданы  $\mathbf{S}$  — множество временных рядов,  $\mathbf{Z}$  — множество номеров классов. Требуется по конечной выборке  $\mathbf{S}^b = \{(\mathbf{s}_1, y_1), \dots, (\mathbf{s}_b, y_b)\}$ , где  $\mathbf{s} = [s_1, \dots, s_N]$  — временной ряд,  $b$  — число объектов в выборке, построить отображение

$$y* : \mathbf{S} \rightarrow \mathbf{Z}. \quad (2)$$

По имеющемуся временному ряду  $\mathbf{s}$  строится траекторная матрица

$$\mathbf{H}_s^n = \begin{bmatrix} s_1 & s_2 & \dots & s_{n-1} & s_n \\ s_2 & s_3 & \dots & s_n & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{N-n+1} & s_{N-n+2} & \dots & s_{N-1} & s_N \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_m \end{bmatrix}, \quad m = N - n + 1,$$

где  $N$  — длина временного ряда,  $n$  — ширина окна, не меньшая, чем предполагаемый период,  $\mathbf{s}_t = [s_t, s_{t+1}, \dots, s_{t+n-1}] \in \mathbb{H}_s \subseteq \mathbb{R}^n$  — векторы, образующие фазовую траекторию ряда  $\mathbf{s}$ .

Размерность траекторного пространства  $\mathbb{H}_s^n$  избыточна. Предлагается снижать размерность  $\mathbb{H}_s^n \rightarrow \mathbb{H}_x^p$  с помощью метода главных компонент при  $p \ll n$ :

$$\mathbf{H}_x^p = \mathbf{H}_s^n \mathbf{U} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_m \end{bmatrix}, \quad \mathbf{x}_t \in \mathbb{R}^p, \quad t \in [1, m] \quad (3)$$

где  $\mathbf{U}$  — матрица преобразования алгоритма метода главных компонент с количеством компонент равным  $p$ , соответствующим наибольшим собственным значениям.

В полученном подпространстве фазовая траектория переводится из декартовых в сферические координаты  $\mathbb{H}_x^p \rightarrow \mathbb{S}_r^p$ :

$$\phi : \mathbf{x} \rightarrow \mathbf{r} = [r, \alpha_{p-1}, \dots, \alpha_1], \quad \mathbf{a} = [\alpha_{p-1}, \dots, \alpha_1], \quad \mathbf{S}_r^p = \phi(\mathbf{H}_x^p).$$

По полученным представлениям точек в пространстве  $\mathbb{S}^p$  строится модель фазовой траектории как линейная комбинация сферических гармоник.

Предлагается представить  $y^*$  (2) как суперпозицию отображений

$$t \mapsto \mathbf{s} \mapsto \mathbb{H}_s^n \rightarrow \mathbb{H}_x^p \rightarrow \mathbb{S}_r^p \rightarrow \mathbb{W}^{p-1} \rightarrow \mathbf{Z},$$

где  $\mathbb{W}^{p-1}$  — пространство весов модели аппроксимации,  $\mathbb{H}_s^n$  — фазовое пространство, полученное методом задержек,  $\mathbb{H}_x^p$  — фазовое подпространство в декартовых координатах,  $\mathbb{S}_r^p$  — фазовое подпространство в сферических координатах.

## 4. Модель фазовой траектории

В полученном подпространстве  $\mathbb{S}^p$  строится модель фазовой траектории

$$f_{\text{sp}} : \mathbb{R}^{|\mathbf{w}_{\text{sp}}|} \times \mathbb{S}_r^p \rightarrow \mathbb{R}. \quad (4)$$

Нормированные значения функции  $f_{\text{sp}}$  интерпретируются как вероятность принадлежности точки фазового пространства  $\mathbb{S}$  к фазовой траектории  $\mathbf{S}$  временного ряда  $\mathbf{s}$

$$\pi(\mathbf{a}) \approx f_{\text{sh}}(\mathbf{w}_{\text{sp}}, \mathbf{a}) = \sum_{l_{p-1}=0}^{N_{\text{approx}}} \sum_{l_{p-2}=0}^{l_{p-1}} \cdots \sum_{l_1=-l_2}^{l_2} w_{l_{p-1}, \dots, l_1} Y_{l_{p-1}, \dots, l_1}(\mathbf{a}), \quad (5)$$

где  $\pi(\mathbf{a})$  — функция проекций, определенная ниже,  $l_{p-1}, \dots, l_1$  — индексы, определяющие сферические гармоники и удовлетворяющие условию  $l_{p-1} \geq l_{p-2} \dots l_2 \geq |l_1|$ ,  $N_{\text{approx}}$  — максимальное значения старшего индекса,  $Y_{l_{p-1}, \dots, l_1}(\mathbf{a})$  — вещественные сферические гармоники, определенные ниже,  $w_{l_{p-1}, \dots, l_1}$  — весовые коэффициенты  $w_{l_{p-1}, \dots, l_1} \in \mathbb{W}^{p-1}$ .  $N_{\text{approx}}$  связан с точностью модели, чем больше значение, тем точнее аппроксимация и больше переобучение.

В качестве базисных функций на поверхности  $(p - 1)$ -мерной сферы используются сферические гармоники:

$$Y_{l_{p-1}, \dots, l_1}(\alpha_{p-1}, \dots, \alpha_1) = \left[ \prod_{k=2}^{p-1} \overline{P}_{l_k}^{l_{k-1}}(\alpha_k) \right] \frac{1}{\sqrt{2\pi}} \exp(\pm i|l_1|\alpha_1), \quad (6)$$

За функцию  $\overline{P}_{l_k}^{l_{k-1}}(\alpha_k)$  обозначается

$$\overline{P}_{l_k}^{l_{k-1}}(\alpha) = c_{l_k}^{l_{k-1}} \cdot (\sin \alpha)^{\frac{-(k-2)}{2}} P_{l_k + \frac{(k-2)}{2}}^{-\left(l_{k-1} + \frac{(k-2)}{2}\right)}(\cos \alpha),$$

$$c_{l_k}^{l_{k-1}} = \sqrt{\frac{2l_{k-1} + k - 1}{2} \frac{(l_k + l_{k-1} + k - 2)!}{(l_k - l_{k-1})!}}$$

где  $P_\mu^{-\eta}(x)$  — полиномы Лежандра,  $c_{l_k}^{l_{k-1}}$  — нормировочный коэффициент. Подробнее о выводе формул и введенных обозначениях в

Сферические гармониками определены в комплексном пространстве. В исследуемом случае нет необходимости в реальной и комплексной части одновременно. Используются только вещественные сферические гармоники. Это упрощает реализацию и сохраняет свойства ортонормированности. Так вещественные сферические гармоники представимы в следующем виде

$$Y_{l_{p-1}, \dots, l_1}(\alpha_{p-1}, \dots, \alpha_1) = \begin{cases} \operatorname{Re}(\mathcal{Y}_{l_{p-1}, \dots, l_1}(\alpha_{p-1}, \dots, \alpha_1)), & \text{если } l_1 \geq 0 \\ \operatorname{Im}(\mathcal{Y}_{l_{p-1}, \dots, l_1}(\alpha_{p-1}, \dots, \alpha_1)), & \text{иначе.} \end{cases} \quad (7)$$

Для расчета  $w_{l_{p-1}, \dots, l_1}$  весовых коэффициентов в (5) необходимо определить проекцию фазовой траектории  $\pi(\mathbf{a})$  на поверхности сферы  $\mathbb{S}^{p-1}$ . Для этого поверхность сферы разбивается на области, как показано слева на Рис. 4.

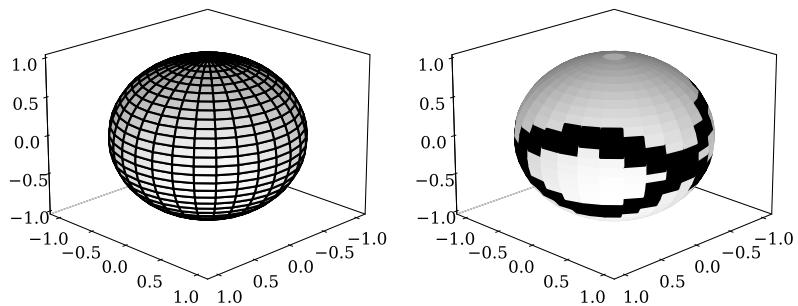


Рис. 4. Слева: области на сфере, справа: полученная проекция.

По полученному разбиению  $\mathbb{A}^{p-1}$  на сфере и по точкам фазовой траектории строится функция проекции  $\pi(\mathbf{a}), \mathbf{a} \in \mathbb{A}^{p-1}$

$$\pi(\mathbf{a}) = \begin{cases} 1, & \text{если } \mathbf{a} \in \mathbb{A}^{p-1} \cap \mathbf{S}_z^p, \\ 0, & \text{если } \mathbf{a} \in \mathbb{A}^{p-1} \setminus \mathbf{S}_z^p. \end{cases} \quad (8)$$

Пример проекции справа на Рис. 4.

Составляется система относительно весовых коэффициентов

$$w_{l_{p-1}, \dots, l_1}$$

$$\begin{pmatrix} Y_{l_{p-1}, \dots, l_1}(\mathbf{a}_1) & \dots & Y_{0, \dots, 0}(\mathbf{a}_1) \\ \vdots & \ddots & \vdots \\ Y_{l_{p-1}, \dots, l_1}(\mathbf{a}_d) & \dots & Y_{0, \dots, 0}(\mathbf{a}_d) \end{pmatrix} \begin{pmatrix} w_{l_{p-1}, \dots, l_1} \\ \vdots \\ w_{0, \dots, 0} \end{pmatrix} = \begin{pmatrix} \pi(\mathbf{a}_1) \\ \vdots \\ \pi(\mathbf{a}_d) \end{pmatrix}, \quad (9)$$

где  $Y_{l_{p-1}, \dots, l_1}(\mathbf{a}_i)$  — значение сферической гармоники в точке  $\mathbf{a}_i$ ,  $d$  — количество точек в сетке на сфере Рис. 4. В более короткой записи:

$$\boldsymbol{\Pi} = \mathbf{w}_{\text{sp}}^\top \mathbf{Y}. \quad (10)$$

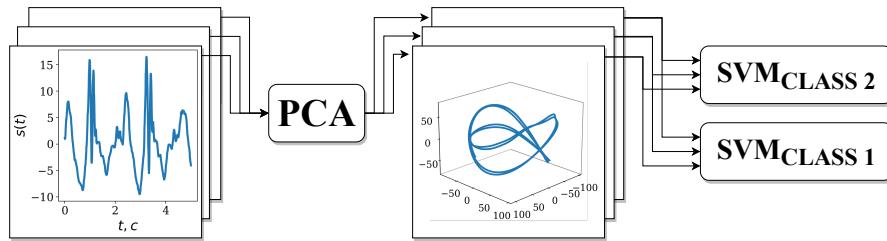
Эта постановка сводит решения к МНК, что ускоряет процедуру расчета весовых коэффициентов. В силу экспоненциального роста числа сферических гармоник при повышении размерности фазового пространства, вводятся регуляризаторы

$$\hat{\mathbf{w}}_{\text{sp}} = \operatorname{argmin}_{\mathbf{w}_{\text{sp}}} \|\mathbf{Y}\mathbf{w}_{\text{sp}} - \boldsymbol{\Pi}\|^2 + \lambda \|\mathbf{w}_{\text{sp}}\|^2. \quad (11)$$

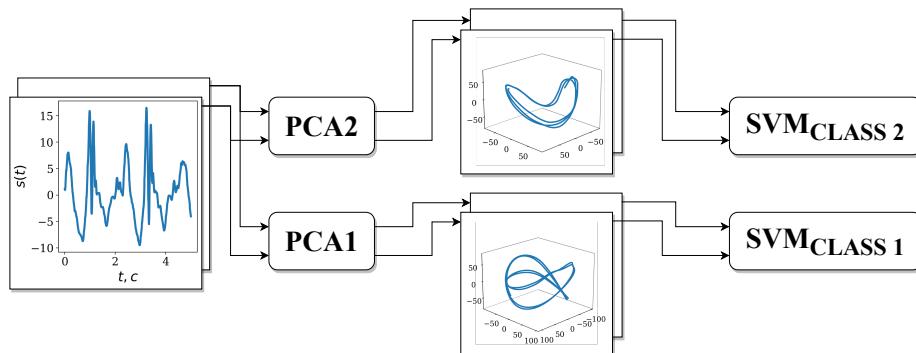
Решение представимо аналитически:

$$\hat{\mathbf{w}}_{\text{sp}} = (\lambda \mathbf{I} - \mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Y}^\top \boldsymbol{\Pi}. \quad (12)$$

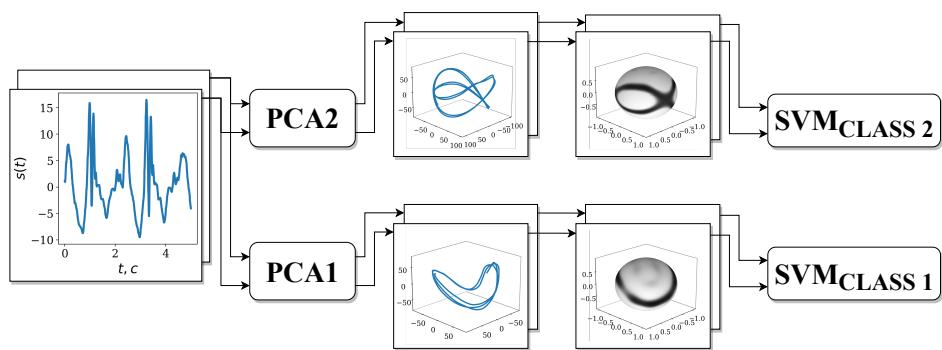
## 5. Модель классификации весовых коэффициентов



(a) Классификации с совместной матрицей PCA для всех классов



(б) Классификации с отдельной матрицей PCA для каждого класса



(в) Предложенный метод классификации

Рис. 5. Сравнение различных подходов к классификации фазовой траектории

Полученные весовые коэффициенты (12) для каждого  $\mathbf{s} \in \mathbf{S}^b$  используются в качестве признакового описания в задаче классификации. Для классификации используется метод опорных векторов с мягким зазором

$$g(\mathbf{w}_{\text{sh}}) = \text{sign}(\mathbf{w}_{\text{svm}}^\top \mathbf{w}_{\text{sh}}). \quad (13)$$

Весовые коэффициенты  $\mathbf{w}_{\text{svm}}$  (13) оптимизируются

$$\begin{cases} \frac{1}{2} \|\mathbf{w}_{\text{svm}}\|_2^2 + C \sum_{i=1}^{\hat{m}} \xi_i \rightarrow \min_{w_{\text{svm}}, \xi} \\ y_i \cdot \mathbf{w}_{\text{svm}}^\top \mathbf{w}_{\text{sh},i} \geq 1 - \xi_i, \quad i = 1, \dots, b \\ \xi_i \geq 0, \quad i = 1, \dots, b, \end{cases} \quad (14)$$

где  $C$  — параметр настройки метода, который позволяет регулировать отношение между максимизацией ширины разделяющей полосы и минимизацией суммарной ошибки,  $\xi_i$  — величина, характеризующая ошибку на объектах  $w_{\text{svm}} i$ .

## 6. Модель классификации точек фазовой траектории

В качестве альтернативы используется ближайшая по архитектуре модель описанная в [6]. В этой работе классифицируются точки фазовой траектории  $\mathbf{H}_x^p$  по стратегии One-Vs-Rest. Решается задача

$$g(\mathbf{x}) = \text{sign}(\mathbf{w}_{\text{svm}}^\top \mathbf{x}), \quad \mathbf{x} \in \mathbf{H}_x^p. \quad (15)$$

Сравниваются два варианта классификации. В первом матрица преобразования метода главных компонент (3) строится для объединения всех фазовых траекторий полученных методом задержек и далее классифицируется для каждого класса отдельным методом опорных векторов как показано на Рис. 5а. Во втором варианте матрица преобразования метода главных компонент строится для каждого класса отдельно и далее классифицируется аналогично первому варианту как показано на Рис. 5б.

## 7. Выбор моделей прогнозирования с помощью сферических гармоник

По временному ряду  $\mathbf{s}$  строится траекторная матрица

$$\mathbf{H}_s^n = \begin{bmatrix} s_1 & s_2 & \dots & s_{n-1} & s_n \\ s_2 & s_3 & \dots & s_n & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{N-n+1} & s_{N-n+2} & \dots & s_{N-1} & s_N \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_m \end{bmatrix},$$

где  $m = N - n + 1$  — число строк,  $N$  — длина временного ряда,  $n$  — ширина окна, не меньшая, чем предполагаемый период,  $\mathbf{s}_t = [s_t, s_{t+1}, \dots, s_{t+n-1}] \in \mathbb{H}_s \subseteq \mathbb{R}^n$  — векторы, образующие фазовую траекторию ряда  $\mathbf{s}$ .

По полученной матрице  $\mathbf{H}_s^n$  строится модель аппроксимации сферическими гармониками, как описано в разделе "Модель фазовой траектории".

Для интерпретации значений модели как вероятности принадлежности вносятся два изменения. К модели применяется функция сигмоиды  $\sigma(x) = \frac{1}{1+e^{-x}}$ :

$$f_{\text{sh}}(\mathbf{w}_{\text{sp}}, \mathbf{a}) = \sigma \left( \sum_{l_{p-1}=0}^{N_{\text{approx}}} \cdots \sum_{l_1=-l_2}^{l_2} w_{l_{p-1}, \dots, l_1} Y_{l_{p-1}, \dots, l_1}(\mathbf{a}) \right),$$

где  $l_{p-1}, \dots, l_1$  — индексы, определяющие сферические гармоники,  $N_{\text{approx}}$  — максимальное значения старшего индекса,  $Y_{l_{p-1}, \dots, l_1}(\mathbf{a})$  — сферические гармоники,  $w_{l_{p-1}, \dots, l_1}$  — весовые коэффициенты.

Оптимационная задача в обозначениях (10) :

$$\hat{\mathbf{w}}_{\text{sp}} = \arg \max_{\mathbf{w}_{\text{sp}}} \sum_{i=0}^d \pi(\mathbf{a}_i) \cdot \ln(\sigma(\mathbf{w}_{\text{sp}}^\top \mathbf{Y}_i)) + (1 - \pi(\mathbf{a}_i)) \cdot \ln(1 - \sigma(\mathbf{w}_{\text{sp}}^\top \mathbf{Y}_i)) \quad (16)$$

где  $\mathbf{Y}_i$  — строка матрицы (10),  $d$  — количество областей разбиения на сфере.

С помощью выбранной модели прогнозирования  $g$  получаются  $m_s$  последовательных прогнозов будущих значений  $\hat{\mathbf{s}} = [\hat{s}_1, \dots, \hat{s}_{m_s}]$ . По спрогнозированным значениям строится траекторная матрица:

$$\hat{\mathbf{H}}_s^n = \begin{bmatrix} s_{N-n+1} & s_{N-n+2} & \dots & s_{N-1} & \hat{s}_1 \\ s_{N-n+2} & s_{N-n+3} & \dots & \hat{s}_1 & \hat{s}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{s}_{m_s-n+1} & \hat{s}_{m_s-n+2} & \dots & \hat{s}_{m_s-1} & \hat{s}_{m_s} \end{bmatrix}$$

К полученной матрице  $\hat{\mathbf{H}}_s^n$  применяется матрица преобразований метода главных компонент, полученная ранее при понижении размерности пространства траекторной матрицы на реальных данных, и переход в сферические координаты

$$\hat{\mathbf{H}}_x^p = \hat{\mathbf{H}}_s^n \mathbf{U}, \quad \mathbf{S}_r^p = \phi(\mathbf{H}_x^p).$$

Далее для каждой точки фазовой траектории с помощью функции аппроксимации  $f_{\text{sh}}$  ставится в соответствие значение

$$\hat{p} = \hat{p}(\hat{s}|t).$$

Оно интерпретируется как вероятность принадлежности точки фазовой траектории к области возможных значений по имеющейся предыстории ряда  $\mathbf{s}$ .

Так схема получения оценок вероятности выглядит следующим образом:

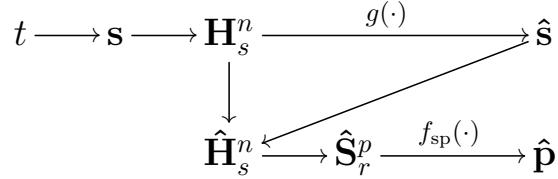


Рис. 6. Схема модели предсказания будущих значений временного ряда и оценки их соответствия предыстории.

По временному ряду  $\mathbf{s}$  строится матрица задержек  $\mathbf{H}_s^n$ . На матрице  $\mathbf{H}_s^n$  настраиваются параметры моделей прогнозирования  $g \in G$  и модель аппроксимации  $f_{sp}(\cdot)$ . По прогнозам модели  $g$  строится задержек над прогнозами  $\hat{\mathbf{H}}_s^n$ . Такой подход позволяет получить оценку соответствия фазовой траектории  $\hat{\mathbf{p}} = [\hat{p}_1, \dots, \hat{p}_{m_s}]$  для каждой спрогнозированной точки  $\hat{\mathbf{s}} = [\hat{s}_1, \dots, \hat{s}_{m_s}]$  для каждой модели  $g \in G$ . Причем для всех моделей функция аппроксимации  $f_{sh}(\cdot)$  и матрица метода главных компонент  $\mathbf{U}$  одни и те же.

$$\begin{pmatrix} \{\hat{\mathbf{s}}_1, \hat{\mathbf{p}}_1\} \\ \vdots \\ \{\hat{\mathbf{s}}_{m_g}, \hat{\mathbf{p}}_{m_g}\} \end{pmatrix} \xrightarrow{\text{Decision Rule}} \hat{\mathbf{s}}_{\text{selected}} \rightarrow g_{\text{selected}}$$

Рис. 7. Схема выбора и/или агрегирования моделей согласно предсказаниями и их оценкам их соответствия предыстории.

По полученному множеству  $\hat{\mathbf{p}} = [\hat{p}_1, \dots, \hat{p}_{m_s}]$  для каждой модели считается логарифм правдоподобия.

$$L(\hat{s}|t) = \sum_{i=1}^{m_s} \ln(\hat{p}_i(\hat{s}|t))$$

Так для каждой модели  $g \in G$  ставится в соответствие логарифм правдоподобия. Выбирается та модель, у которой логарифм правдоподобия больше. Для соответствия начальной постановке — минимальное значение отрицательного логарифма правдоподобия.

$$g_{\text{selected}} = \arg \min_{g \in G} -L(\hat{s}|t, g, s)$$

Подобное ранжирование интерпретируется как соответствие прогнозов уже имеющейся предыстории.

## 8. Альтернативные методы выбора моделей прогнозирования

Для сравнения применяются три вида информационных критериев учитывающих количество параметров моделей и их точность на выборке для обучения. Обозначения в формулах ниже  $RSS$  — остаточная сумма квадратов,  $n$  — число объектов в выборке,  $k$  — число параметров модели.

**Информационный критерий Акаике** — это критерий выбора, оценивающий модели с разным числом параметров:

$$AIC = n \cdot \ln \frac{RSS}{n-2} + 2 \cdot k$$

**Байесовский информационный критерий**, аналогичный AIC, но с более строгим штрафом за количество параметров модели:

$$BIC = n \cdot \ln RSS + k \cdot \ln n$$

**Mallows's  $C_p$**  похожий подход с учетом дисперсии точности среди конечного множества доступных линейных моделей  $g \in G$ :

$$C_p = \frac{1}{n}(RSS + 2d\hat{\sigma}^2)$$

где  $d$  — мощность множества  $G$ ,  $\hat{\sigma}^2 = \frac{\min_{g \in G} RSS}{n-k-1}$ . Считается, что наименьший  $RSS$  будет у модели с наибольшим количеством параметров.

## 9. Эксперимент по классификации сферическими гармониками

Задан набор данных, измеренный датчиками акселерометра телефона. Он собирается с частотой дискретизации 50 Гц. Двадцать четыре участника разного пола, возраста, веса и роста выполнили три типа повторяемых физических активностей: движение по лестнице, ходьба, бег. Более подробное описание набора данных в [4].

Участники имеют разные типы походки и, как следствие, разные фазовые траектории. Поэтому были отобраны три человека со схожими показателями роста и веса.

Алгоритм классификации схож с вариантом на Рис. 5б. Основное отличие в аппроксимация фазовой траектории линейной комбинацией сферических гармоник как показано на Рис. 5в.

Размер предыстории выбирается равным  $n = 150$ , что советует трем секундам. Период движений в данных не превосходит двух с половиной секунд.

Значение старшего индекса, используемого для построения модели сферических (5) гармоник, выбирается равным  $N_{\text{approx}} = 10$  для размерности фазового пространства  $\mathbf{H}_x^p$  равного трем,  $N_{\text{approx}} = 6$  для размерности четыре и  $N_{\text{approx}} = 3$  для размерности пять. Для пространств большей размерности эксперимент не проводился из-за малого объема данных и вычислительной неэффективности подсчета сферических гармоник для размерности пространства равной шести и более.

Таблица 1. F-score классификации на трех пользователях с одинаковым типом фазовой траектории

| Метод           | Ходьба |      |      | Бег  |      |      | Лестница |      |      |
|-----------------|--------|------|------|------|------|------|----------|------|------|
| Размерность $p$ | 3      | 4    | 5    | 3    | 4    | 5    | 3        | 4    | 5    |
| Пользователь 1  | 0.98   | 0.99 | 0.95 | 0.97 | 0.99 | 0.8  | 0.86     | 0.84 | 0.92 |
| Пользователь 2  | 0.99   | 0.97 | 0.57 | 0.98 | 0.98 | 0.91 | 0.95     | 0.68 | 0.86 |
| Пользователь 3  | 0.96   | 0.93 | 0.58 | 0.89 | 0.64 | 0.82 | 0.85     | 0.65 | 0.86 |

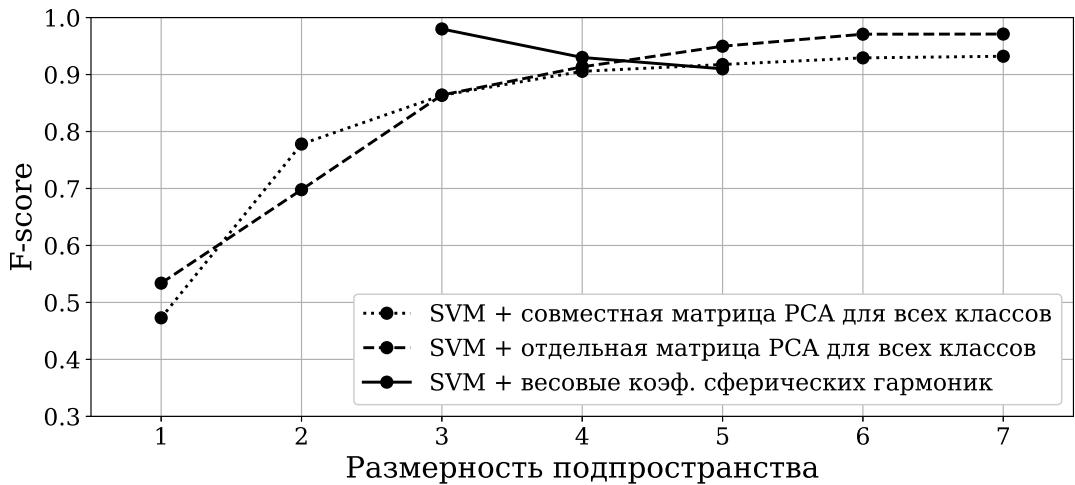


Рис. 8. F-score на ходьбе, беге, спуску и подъему по лестнице для одного пользователя

На Рис. 8 показано качество классификации F-score движений одного человека. Видно, что для эффективной классификации достаточно трехмерных сферических гармоник.

При применении методов, обученных на данных первого пользователя, для классификации временных рядов записанных с второго и третьего пользователей качество падает в зависимости от типа движения. Так ходьба и бег классифицируются лучше, чем движение по лестнице. В таблице 1 представлена качество класси-

фикации для нескольких пользователей. При этом из-за увеличения размерности пространства падает общее качество классификации. Средние значения представлены на Рис. 9 и в таблице 2.

Таблица 2. Среднее качество классификации на трех пользователях с одинаковым типом фазовой траектории

| Метод  | F-score |      |      |
|--|---------|------|------|
| Размерность $p$                                | 3       | 4    | 5    |
| SVM + весовые коэф. сферических гармоник       | 0.90    | 0.82 | 0.81 |
| SVM + совместная матрица PCA для всех классов  | 0.79    | 0.80 | 0.89 |
| SVM + отдельная матрица PCA для каждого класса | 0.78    | 0.81 | 0.81 |

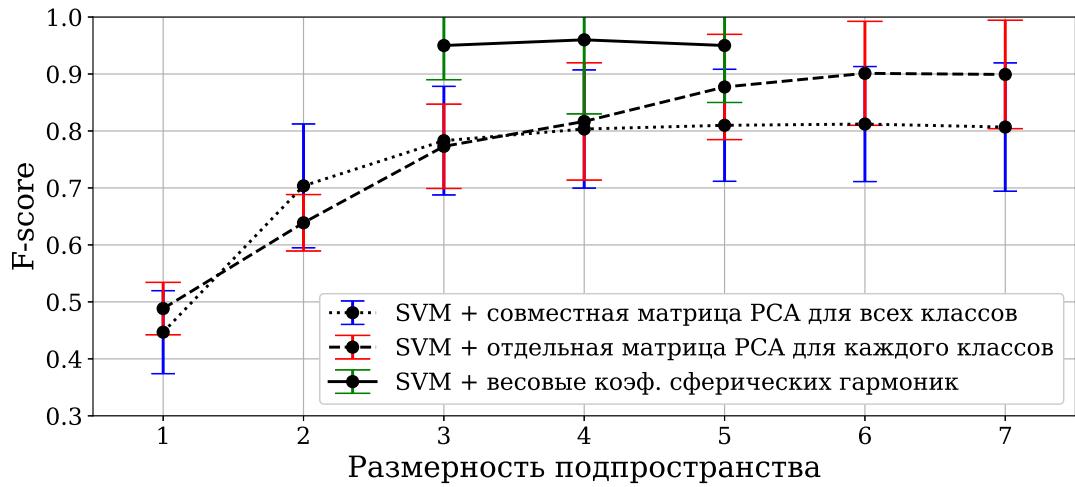
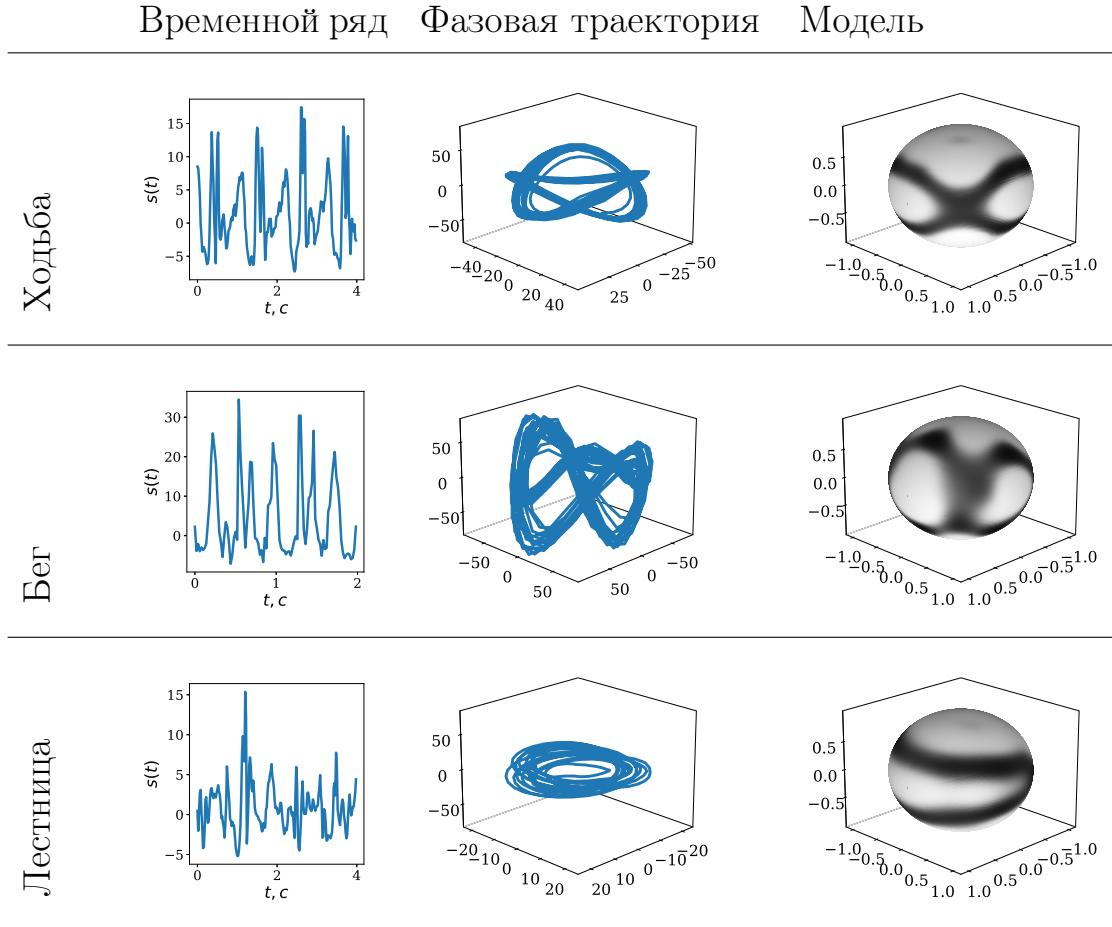


Рис. 9. Средний и СКО F-score на ходьбе, беге, спуску и подъему по лестнице

Таблица 3. Таблица сравнений разных классов движений



В таблице 3 показаны результаты работы алгоритма. Слева фрагмент временного ряда записанного с акселерометра. По центру полученные фазовые траектории в трехмерном пространстве. Справа модели сферических гармоник. Видно, что разные классы временных рядов имеют разные траектории и разные модели аппроксимации.

## 10. Эксперимент по выбору модели прогнозирования

**Данные акселерометра велопрогулки.** Сигнал был записан с датчика акселерометра во время поездки на велосипеде. Частота дискретизации 50Гц. Для оценки модели применялись оценка RMSE на отложенной выборке, информационные критерии BIC, AIC, Mallow's  $C_p$ . Временной ряд был разделен: 50% — для обучения, 25% — для RMSE методом HoldOut, 25% — для оценки методов выбора модели .

Таблица 4. Качества ранжирования моделей на профилю велопрогулки

| G     | MSE тест        | MSE отложенная   | $\hat{p}$     | $C_p$            | AIC             | BIC             |
|-------|-----------------|------------------|---------------|------------------|-----------------|-----------------|
| $g_0$ | 9.724991        | 12.343679        | 0.7439        | <b>21.188266</b> | <b>1.132565</b> | <b>1.664724</b> |
| $g_1$ | 9.736825        | 11.796950        | 0.7312        | 21.262064        | 1.136510        | 1.721885        |
| $g_2$ | 9.523704        | 11.495201        | 0.7595        | 21.570077        | 1.152974        | 1.791565        |
| $g_3$ | <b>9.195918</b> | <b>11.147608</b> | <b>0.8426</b> | 21.834579        | 1.167112        | 1.858919        |
| $g_4$ | <b>9.159391</b> | <b>11.180613</b> | <b>0.8492</b> | 22.493779        | 1.202348        | 1.947371        |
| $g_5$ | 11.213219       | 12.465094        | 0.7808        | 24.320765        | 1.300005        | 2.098244        |
| $g_6$ | 10.610619       | 12.381153        | 0.7958        | 24.836478        | 1.327571        | 2.179026        |
| $g_7$ | 11.351968       | 12.014565        | 0.8183        | 25.050591        | 1.339016        | 2.243687        |
| $g_8$ | 12.638967       | 11.746108        | 0.7952        | 25.379346        | 1.356588        | 2.314475        |
| $g_9$ | 13.561436       | 12.335483        | 0.8161        | 24.528546        | 1.311111        | 2.322214        |

Как видно из таблицы 4 предложенный методу лучше ранжирует модели и выделяет более точную модель в будущем. Близким по точности методом является оценка качества на отложенной выборке. Информационные критерии показали.

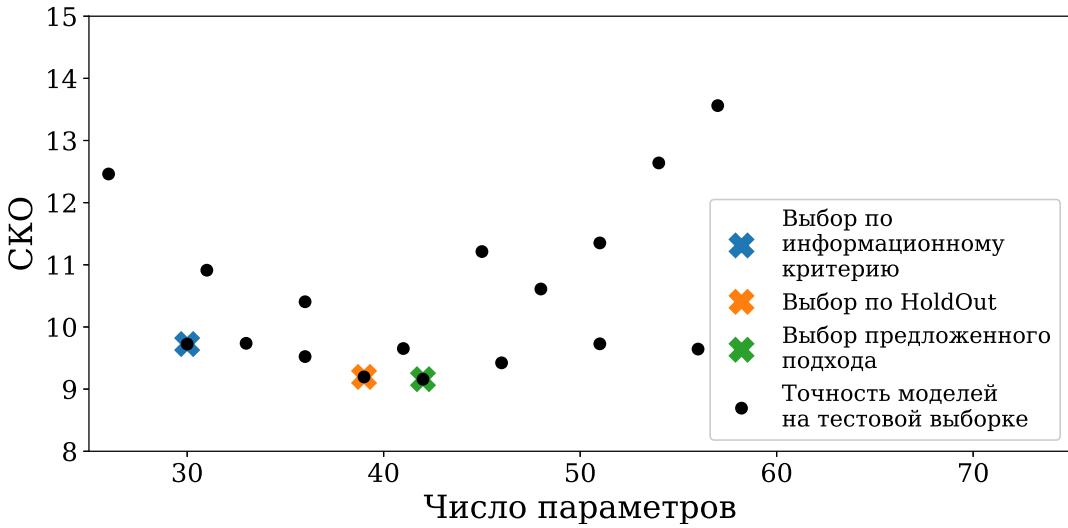


Рис. 10. СКО моделей на тестовой выборке

На из Рис. 10 видно, что предложенный метод и информационные критерии выбирают оптимальный по Парето модели. Предложенный метод выбирает более точные модели. Информационные критерии выбирает более простые модели.

Как видно из Рис. 11 выбранные модели согласно информационным критериям имеют менее вероятную фазовую траекторию. Предложенный метод выбирает более вероятную фазовую траекторию так, чтобы больше точке лежало в области дисперсии.

**Датасет физической активности.** Эксперимент проведен на данных из [4]. Данные — это показатели акселерометра с мобильных устройств десяти пользователей по профилю ходьбы и бега на

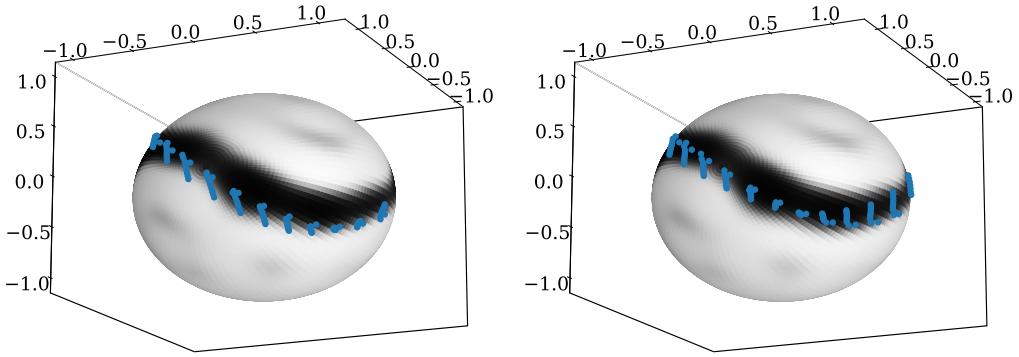


Рис. 11. Слева: выбранная модель по информационным критериям, справа: выбранная модель по предложенному методу и HoldOut.

двух разных маршрутах . Для корректной оценки ошибки среднеквадратичная ошибка MSE была заменена на коэффициент детерминации  $R^2$ . Так амплитуда временного ряда не влияет на сравнительный анализ сравнительный анализ. Для каждого временного ряда считается  $\Delta R^2 = R_{\text{лучшая}}^2 - R_{\text{выбранная}}^2$ .  $R_{\text{лучшая}}^2$  — лучшее значение прогнозов на тестовой выборке.  $R_{\text{выбранная}}^2$  — значение прогнозов выбранной модели на тестовой выборке. Для каждого метода выбора модели считается Средний и СКО  $\Delta R^2$ .

Таблица 5. Качества ранжирования на данных физической активности

|                      | HoldOut | $\hat{p}$ | AIC  | BIC  | Mallow's $C_p$ |
|----------------------|---------|-----------|------|------|----------------|
| Точность выбора      | 0.31    | 0.34      | 0.26 | 0.22 | 0.26           |
| Средний $\Delta R^2$ | 0.11    | 0.09      | 0.11 | 0.10 | 0.11           |
| СКО $\Delta R^2$     | 0.19    | 0.15      | 0.17 | 0.16 | 0.17           |

Как видно из таблицы 5 предложенный метод в среднем лучше выбирает модели. Чем больше точность и чем меньше средний  $\Delta R^2$  тем лучше. Предложенный метод в среднем лучше выбирает модели. Это происходит в силу хорошей настройки модели аппроксимации. На временных рядах с базовыми настройками гиперпараметров модели аппроксимации фазовой траектории качество выбора модели хуже, чем у метода HoldOut и информационных критериев. Это связано с наличием в временных рядах разладки. Так изменится характера движения, то изменится вид фазовой траектории. Метод является чувствительным к изменению с быстрой ходьбы на прогулочную.

## Заключение

Был предложен метод построения фазового пространства. В сферических координатах была построена модель аппроксимации, как линейная комбинация сферических гармоник. Модель аппроксимации сохраняет геометрическую структуру фазовой траектории временного ряда и интерпретируется как вероятность принадлежности точки фазовой траектории к предыстории.

Вычислительные эксперименты показали, что параметры данной модели имеют лучшую разделяющую способность при меньшей размерности подпространства, чем ближайший метод. При меньшем количестве настраиваемый параметров качество лучше. Предложенный метод выбора модели прогнозирования показал лучшее качество ранжирования в сравнении с оценкой на отложенной выборке и информационными критериями BIC, AIC, Mallow's  $C_p$ .

Однако для большей эффективности метод требует сложной настройки и больший размер обучающей выборки, чем альтернативные методы. При невозможности настроить гиперпараметры модели качество хуже. Предложенный метод оказывается излишне чувствительным к изменению характера походки. Так у разных пользователей могут быть разные типы фазовых траекторий для одного типа движения. Подобные особенности будут рассмотрены в будущих работах.

## Список литературы

1. Christoph Bergmeir and José M. Benítez. On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191:192–213, May 2012.
2. Christoph Bergmeir, Rob J. Hyndman, and Bonsoo Koo. A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics and Data Analysis*, 120:70–83, April 2018.
3. Vitor Cerqueira, Luís Torgo, and Carlos Soares. Model selection for time series forecasting: Empirical analysis of different estimators. 04 2021.
4. Malekzadeh et al. Protecting sensory data against sensitive inferences. In *Proceedings of the 1st Workshop in Distributed Systems*, 2018.
5. Kenneth Ezukwoke and Samaneh Zareian. Kernel methods for principal component analysis (pca) a comparative study of classical and kernel pca. 12 2019.
6. Jordan Frank, Shie Mannor, and Doina Precup. Activity and gait recognition with time-delay embeddings. *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI*, 2010.
7. Zhigljavsky A. A. Golyandina N., Nekrutkin V. Analysis of time series structure: Ssa and related techniques. *Chapman and Hall*, 2002.
8. D. Lai and G. Chen. Dynamical systems identification from time-series data: A hankel matrix approach. *Mathematical and Computer*

- Modelling*, 24(3):1–10, 1996.
9. Anastasia Motrenko and Vadim Strijov. Extracting fundamental periods to segment biomedical signals. *IEEE journal of biomedical and health informatics*, 20, 08 2015.
  10. Bernhard Scholkopf, Alex Smola, and Klaus-Robert Muller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10:1299–1319, 07 1998.
  11. Floris Takens. Detecting strange attractors in turbulence. pages 366–381, 1981.
  12. K. Usmanova, Yu Zhuravlev, K. Rudakov, and Vadim Strijov. Approximation of quasiperiodic signal phase trajectory using directional regression. *Computational Mathematics and Cybernetics*, 44:196–202, 10 2020.