«Московский физико-технический институт (национальный исследовательский университет)»
Физтех-школа прикладной математики и информатики
Кафедра «Интеллектуальные системы»

Курдюкова Антонина Дмитриевна

Снижение размерности фазового пространства в задачах канонического корреляционного анализа

03.03.01 – Прикладные математика и физика

Выпускная квалификационная работа бакалавра

Научный руководитель:
д.ф.-м.н. Стрижов Вадим Викторович

Москва

2022

Содержание

1 Введение **5**

2 Постановка задачи **7**

3 Теоретическая часть **7**

3.1 Теорема Такенса 9

3.2 Метод сходящихся перекрестных отображений 10

3.3 Метод проекций на латентные структуры 10

4 Методы Сугихары **12**

4.1 Simplex projection 12

4.2 Sequential Locally Weighted Global Linear Maps (S-Map) 12

4.3 Multivariate embeddings 13

4.4 Multiview embeddings 13

5 Вычислительный эксперимент **13**

5.1 Акселерометр + гироскоп, ходьба 14

5.2 Несвязанные сигналы, акселерометр + акселерометр 15

5.3 Медленная ходьба, акселерометр + гироскоп 15

Аннотация

Данная работа посвящена задаче снижения размерности фазового пространства методами канонического корреляционного анализа. Исследуется связь между методом канонического корреляционного анализа и методом сходящихся перекрестных отображений Сугихары. Вид прогностических моделей представляется в виде условия принадлежности двух аттракторов к общей динамической системе. Аттракторы восстанавливаются в исходном и целевом фазовых пространствах. В работе рассмотрены метод проекций на латентные структуры, метод канонического кореляционного анализа, их нелинейные модификации, метод сходящихс] перекрестных отображений, seq2seq, Neural ODE. Сформулирован вариант теоремы о вложениях Такенса для проверки удовлетворенгия методов прогноза условиям Сугихары. Решается прикладная задача в восстановления траектории движения руки человека по сигналу акселерометра. Рассматривается видеоряд ходьбы человека с акселерометром на руке.

Ключевые слова: *снижение размерности, фазовое пространство, аттрактор, метод сходящихся перекрестных отображений, теорема Такенса о вложениях*

1 Введение

Решается задача прогнозирования сигналов походки человека. Такие сигналы обладают сложной структурой. Под сложной структурой понимаются зависимости и изменяющийся период. Рассматриваются два связанных фазовых пространства. Одно из них является исходным, другое - целевым. Например, фазовое пространство сигналов акселерометра и гироскопа одного мобильного устройства; пространства сигналов двух акселерометров в правой и левой руке человека; пространство траектории движения руки, восстановленной по видеоряду движения человека, и пространство сигнала акселерометра на этой руке.

Целью работы является построение более простой модели, работающей не хуже уже существующих моделей прогнозирования временных рядов.

Для улучшения качества прогноза, а также для упрощения прогностической модели предлагается учесть зависимости между временными рядами, а также перейти в пространство меньшей размерности. Снижение размерности позволяет учитывать внутреннее низкоразмерное проедставление временных рядов в прогностической модели.

Для определения наличия связи между временными рядами используется метод сходящегося перекрестного отображения (convergent cross mapping, CCM) [**?**, **?**]. Метод CCM проверяет, насколько близки точки фазового пространства временного ряда s_1, соответствующие ближайшим соседям ряда s_2.Под близостью понимается существование взаимно однозначного соответствия, которое отображает окрестность фазовой траектории s_1 в окрестность фазовой траектории s_2.

Для снижения размерности траекторного пространства используются метод проекций в латентное пространство (partial least squares PLS) [**?**, **?**], его нелинейная модификация [?], seq2seq[..], NeuralODE[..]. Снижение размерности позволяет сделать прогностическую модель более устойчивой, изучить связь между главными компонентами временных рядов, а также найти траекторное подпространство, в котором удастся обнаружить связь между временными рядами.

В работе исследуется связь между методами корреляционного анализа и методом сходящегося перекрестного отображения. Требуется построить прогностическую модель, связывающую метод сходящегогся перекрестного отображения и методы канонического корреляционного анализа. Для CCM нет способа выбора собственного подпространства, в котором аппроксимируется многообразие компакта и работает прогностическая модель. На текущий момент выбор собственного пространства осуществляется перебором по главным компонентам, например в [?]. Работа Исаченко [..] по PLS дает возможность обобщить методы выбора подпространства с PLS на CCM.

Определение 1 *Динамическая система – множество элементов, для которого задана функциональная зависимость между временем и положением в фазовом пространстве каждого элемента системы.*

Динамическая система представляет собой такую математическую модель некоего объекта, процесса или явления, в которой пренебрегают «флуктуациями и всеми другими статистическими явлениям».

Определение 2 *Многообразие – хаусдорфово топологическое пространство со счётной базой, каждая точка которого обладает окрестностью, гомеоморфной евклидову пространству* $\mathbb{R}^n$

2 Постановка задачи

Дан временной ряд $s_1 = \{s_i^1\}_{i=1}^{N_1}$. Значения временного ряда заданы в моменты времени $1, \ldots, N_1$. Требуется построить прогноз ряда на следующие m значений $N_1 + 1, \ldots, N_1 + m$.

При построении прогностической модели $\mathcal{F}$ нужно учесть влияние ряда $s_2 = \{s_i^2\}_{i=1}^{N_2}$ на ряд s_1. Значения ряда s_2 в моменты времени $N_1 + 1, \ldots, N_1 + m$ известны, то есть $N_2 > N_1 + m$.

Для построения прогноза ряда s_1 на один шаг по времени вперед будем учитывать L предыдущих значений этого ряда и все предшествующие текущему моменту времени значения ряда s_2. Тогда прогностическая модель имеет вид:

$$\widehat{s}\,_{t+1}^{1} = \mathcal{F}(\widehat{\mathbf{w}}, s_t^1, \ldots, s_{t-L+1}^1, s_1^2, \ldots, s_t^2), \tag{1}$$

$$\widehat{\mathbf{w}} = \arg\min_{\mathbf{w}} L(\mathbf{w}, s_1, \widehat{s_1}),$$

где L — функция потерь.

3 Теоретическая часть

Пусть $s_1 = \{s_i^1\}_{i=1}^{N_1}$ и $s_2 = \{s_i^2\}_{i=1}^{N_2}$ — заданные временные ряды. Опишем, как строится фазовое пространство $\mathbf{X}$ временного ряда. Строится ганкелева матрица для ряда s_1:

$$\mathbf{H_1} = \begin{bmatrix} s_1 & \ldots & s_{n_1} \\ s_2 & \ldots & s_{n_1+1} \\ \vdots & \ddots & \vdots \\ s_{k_1} & \ldots & s_{N_1} \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} \mathbf{s}_1^1, \mathbf{s}_2^1, \ldots, \mathbf{s}_k^1 \end{bmatrix}, \quad k_1 = N_1 - n_1 + 1,$$

где n – ширина окна. Аналогично для временного ряда s_2. Тогда вектора $\mathbf{s}^1_1, \mathbf{s}^1_2, \ldots, \mathbf{s}^1_k$ образуют фазовую траекторию, или, иными словами, аттрактор $\mathbf{M}_1$ временного ряда s_1. На эти же вектора натянуто фазовое пространство $\mathbf{X_1}$ размерности n_1 временного ряда s_1. Формально, под терминами *фазовое пространство* и *аттрактор* будем понимать следующее:

Определение 3 *Фазовое пространство* $\mathbf{X}$ *динамической системы – совокупность всех допустимых состояний динамической системы.*

Определение 4 *Аттрактор* $\mathbf{M}$ *– компактное подмножество фазового пространства динамической системы, все траектории из некоторой окрестности которого стремятся к нему при времени, стремящемся к бесконечности.*

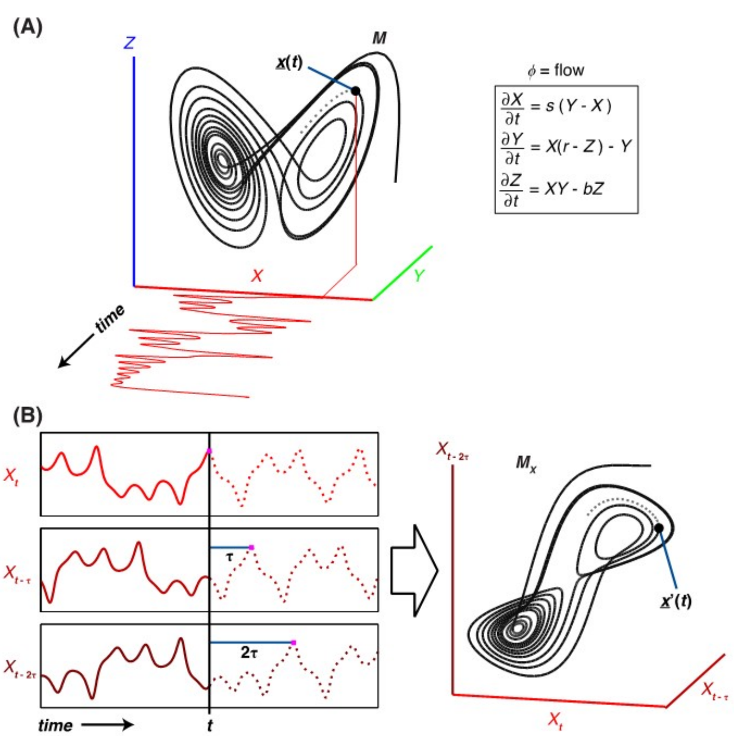


Figure 3. Empirical dynamic modeling **(A)** Example Lorenz system. The attractor manifold ***M*** is the set of states that the system progresses through time. Projection of the system state from ***M*** to the coordinate axis *X* generates a time series. **(B)** Lags of the time series *X* are used as coordinate axes to construct the shadow manifold $\boldsymbol{M}_X$ which is diffeomorphic (maps 1:1) to the original manifold ***M***. The visual similarity between $\mathbf{M}_X$ and ***M*** is apparent.

Рис. 1

Динамическая система может быть описана системой дифференциальных уравнений, где каждая переменная может зависеть от состояния и изменения других оставшихся переменных. Евклидово пространство данных переменных образует пространство состояний системы. Многообразие состояний системы в этом пространстве образует аттрактор $\mathbf{M}$. Проекция многообразия $\mathbf{M}$ на координатные оси дает временной ряд соответствущей наблюдаемой. С другой стороны, по временному ряду наблбдаемой можно восстановить многообразие аттрактора в фазовом пространстве.

3.1 Теорема Такенса

Теорему Такенса можно сформулировать следующим образом:

Теорема 1 *Пусть* $\mathbf{M}$ *— компактное многообразие размерности* d*,* ϕ *— гладкое векторное поле,* X *— гладкая функция, заданная на* $\mathbf{M}$*.*
Тогда отображение $\mathbf{\Phi}_{(\phi,X)(\underline{m})} : \mathbf{M} \to \mathbb{R}^{2d+1}$*, которое задается следующим образом:*

$$\mathbf{\Phi}_{(\phi,X)(\underline{m})} = \langle X(\underline{m}),\, X(\phi(\underline{m})),\, X(\phi^2(\underline{m})),\, \dots,\, X(\phi^{2d}(\underline{m}))\rangle,$$

является вложением, где ϕ *— поток, заданный на* $\mathbf{M}$*.*

Теорема показывает, что скрытое представление $\mathbf{M}_X$ исходного многообразия $\mathbf{M}$ восстанавливается по одной лишь его проекции, то есть по временному ряду X_t. Теорема проиллюстрирована на рисунке 1 (B). Изображен один временной ряд X_t и две его копии $X_{t-\tau}$ и $X_{t-2\tau}$, сдвинутые на τ и 2τ. Тогда в координатном пространстве $(X_t, Xt-\tau), X_{t-2\tau}$ временные ряды представляют собой скрытое представление $\mathbf{M}_X$ многообразия $\mathbf{M}_X$. Представление $\mathbf{M}_X$ сохраняет важные математические свойства исходной динамической системы, например, топологию исходного многообразия. Более того, метод представляет взаимно однозначное соответствие между $\mathbf{M}$ и $\mathbf{M}_X$.

3.2 Метод сходящихся перекрестных отображений

Метод сходящихся перекрестных отображений (convergent cross mapping, CMM) используется для исследования временных рядов на нанличие причинно–следственной связи. Корелляция не подразумевает причинно–следственную связь между рядами. Метод основан на теореме Такенса о вложениях. В общем случае многообразие аттрактора динамической системы может быть восстановлено по одной наблюдаемой $\mathbf{X}$.

Согласно методу временной ряд s_1 может быть восстановлен по ряду s_2 только если временной ряд s_2 связан с рядом s_1. Временные ряды считаются связанными, если окрестность фазовой траектории $\mathbf{x}$ временного ряда s_1 взаимно однозначно отображается в окрестность фазовой траектории $\mathbf{y}$ ряда s_2. Иными словами,

Определение 5 *Аттракторы* $\mathbf{M}_1$ *и* $\mathbf{M}_2$ *наблюдаемых* $\mathbf{X}$ *и* $\mathbf{Y}$, *если* $\mathbf{X}$ *и* $\mathbf{Y}$ *принадлежат одной динамической системе.*

3.3 Метод проекций на латентные структуры

Метод проекций на латентные структуры PLS [?, ?] используют для нахождения фундаментальных зависимостей между двумя матрицами $\mathbf{X}$ и $\mathbf{Y}$. Отбираются наиболее значимые прихнаки. Новые признаки являются их линейными комбинациями. Осуществляется переход в фазовое пространство меньшей размерности. Метод PLS позволяет найти фазовое подпространство, в котором наблюдается связь между главными компонентами исходных временных рядов. Это позволяет исследовать наличие связи между временными рядами.

Пусть $\mathbf{X} \in \mathbb{R}^{m \times n}$ и $\mathbf{Y} \in \mathbb{R}^{m \times r}$ — матрицы двух фазовых пространств, построенных по временному ряду $\mathbf{s}_1$ и $\mathbf{s}_2$ соответственно. Требуется построить прогноз временного ряда $\mathbf{s}_2$ с учетом связи с временным рядом $\mathbf{s}_1$. Предполагается линейная зависимость между строками $\mathbf{X}$ и $\mathbf{Y}$:

$$\mathbf{Y}_i = \mathbf{X}_i \cdot \mathbf{\Theta} + \boldsymbol{\varepsilon} \quad \mathbf{Y}_i \in \mathbb{R}^r,\ \mathbf{X}_i \in \mathbb{R}^n,\ i = 1, \ldots, m, \tag{2}$$

где $\mathbf{\Theta}$ — матрица весов линейной зависимости, ε — вектор ошибок.

Ошибка вычисляется по формуле:

$$S(\mathbf{\Theta}, \mathbf{X}, \mathbf{Y}) = \|\mathbf{Y} - \mathbf{X} \cdot \mathbf{\Theta}\|_2^2 \tag{3}$$

Алгоритм PLS находит матрицы $\mathbf{T}, \mathbf{U}, \mathbf{P}, \mathbf{Q}$, с помощью которых осуществляется переход в латентное пространство согласно формулам:

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P} + \mathbf{F} \tag{4}$$

$$\mathbf{Y} = \mathbf{U} \cdot \mathbf{Q} + \mathbf{E}$$

Матрицы $\mathbf{T}, \mathbf{U}$ наилучшим образом описывают $\mathbf{X}$ и $\mathbf{Y}$. Их столбцы ортогональны. Матрицами $\mathbf{P}$ и $\mathbf{Q}$ определяется переход из латентного пространства в исходное. Матрицы $\mathbf{X}$ и $\mathbf{Y}$ — матрицы невязок.

Алгоритм PLS также позволяет определить матрицу $\mathbf{W}$, с помощью которой рассчитывается матрица весов $\mathbf{\Theta}$:

$$\mathbf{\Theta} = \mathbf{W}(\mathbf{P}^{\mathsf{T}}\mathbf{W})^{-1}\mathbf{Q}^{\mathsf{T}} \tag{5}$$

$$
\begin{tikzcd}
 & S \arrow[dl, bend right] \arrow[dr, bend left] & \\
\mathbf{X} \in \mathbb{R}^{m\times n} \arrow[rr, "\mathcal{F}"] \arrow[dr, bend right] & & \mathbf{Y} \in \mathbb{R}^{m\times r} \arrow[dl, bend left] \\
 & \mathbf{T}, \mathbf{U} \in \mathbb{R}^{m\times \ell} \arrow[ul, bend left, "\mathbf{P}"] \arrow[ur, bend right, "\mathbf{Q}"'] &
\end{tikzcd}
\tag{6}
$$

На коммутативной диаграмме S — динамическая система, порождающая часть, $\mathbf{X}$ — фазовое пространство, $\mathbf{Y}$ — наблюдаемое пространство, $\mathcal{F}$ — гомоморфизм, $\mathbf{T}, \mathbf{U}$ — латентно-согласованные пространства, не можем измеритьнапрямую.

4 Методы Сугихары

4.1 Simplex projection

Метод симплексной проекции, описанный в [?], позволяет строить краткосрочный прогноз траекторий хаотических динамических систем. В работе исследуются различия между детерминированным хаосом системы и ошибкой измерения данных и шумом. Оценивается размерность фазового пространства аттрактора хаотического временного ряда. За рамками статьи остаются временные ряды конечной длины.

4.2 Sequential Locally Weighted Global Linear Maps (S-Map)

Метод, описанный в [?], рассматривает временной ряд как результат эволюции динамической системы во времени. Описаны некоторые проблемы, касающиеся прогнозирования при обнаружении нелинейностей и хаоса. Рассматривается нелинейное прогнозирование в задаче классификации. Предлагается метод характеризации нелинейности с помощью S-map и метод анализа нескольких короткосрочный временных рядов сложного аттрактора.

4.3 Multivariate embeddings

В работе [?] рассмтаривается влияние нелинейных процессов на эпизодические взаимосвязи наблюдаемых переменных динамической системы. Подход заключается в построении ряда алгоритмов, от глобального линейного до локального нелинейного, для прогнозирования данных на основе вложений с запаздывающими координатами [?]. В качестве алгоритма прогнозирования используется метод S-map [?].

4.4 Multiview embeddings

Подход, описанный в [?], применяется к сложным взаимосвязанным системам. Основная идея заключается в реконструировании аттрактора многомерного временного ряда с разных точек зрения и их обЪединение в единую модель. Эффективен для коротких и зашумленных временных рядов.

5 Вычислительный эксперимент

Вычислительный эксперимент проводился на данных [?]. Целью вычислительного эксперимента является исследование качества предсказания метода PLS для зависимых и независимых временных рядов. Временые ряды проверяются на зависимость с помощью метода PLS.

Вычислительный эксперимент поддтверждает теоретический факт о том, что метод PLS является частным случаем метода прекрестных отображений Сугихары.

Результаты эксперимента позволяют ответь на вопрос: достаточно ли из двух сигналов, акселеорметра и гироскопа, какого-либо одного из них. Этого можно достичь в случае хорошего качества восстановления одного сигнала по второму. В данном эксперименте по сигналу акселеорметра восстанавливается сигнал гироскопа с помощью алгоритма PLS.

5.1 Акселерометр + гироскоп, ходьба

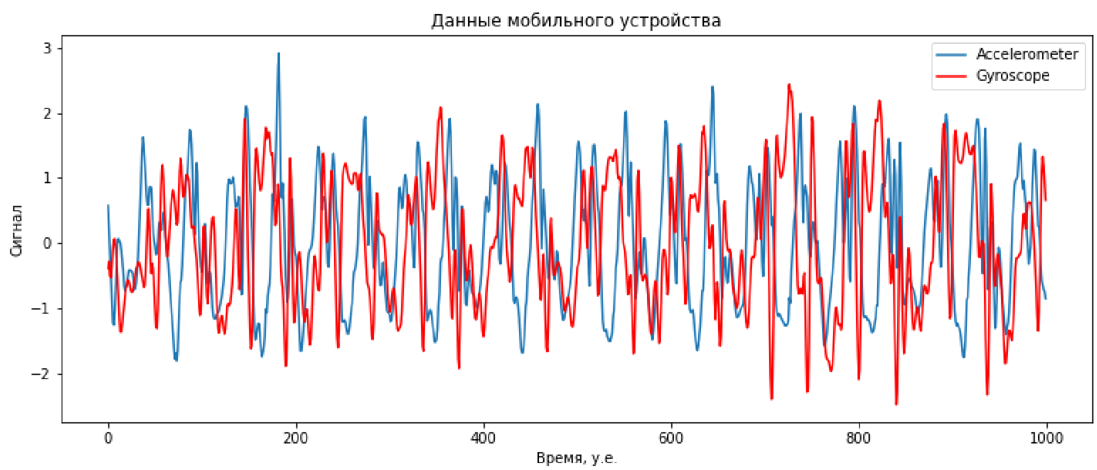
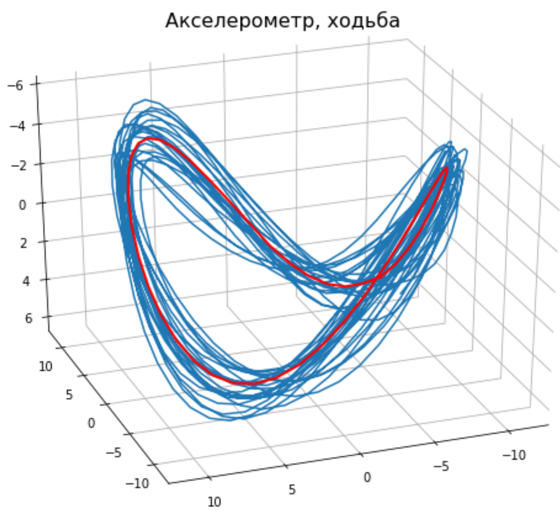
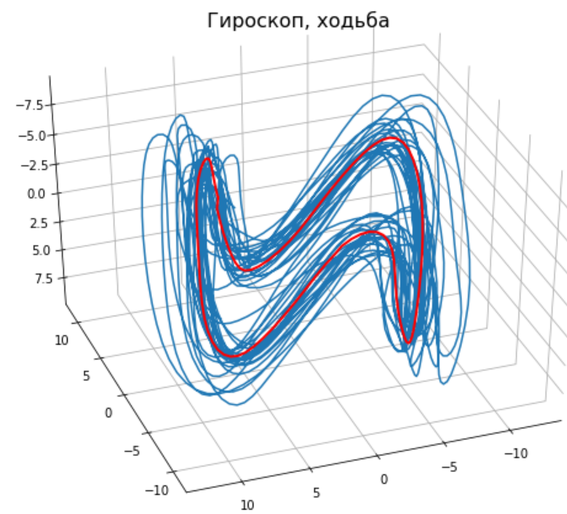


Рис. 2: Данные акселерометра и гироскопа одного мобильного устройтсва

Среднеквадратичная ошибка MSE $= 0.48$



(a) Акселерометр



(b) Гироскоп

Рис. 3: Траектории в фазовом пространстве

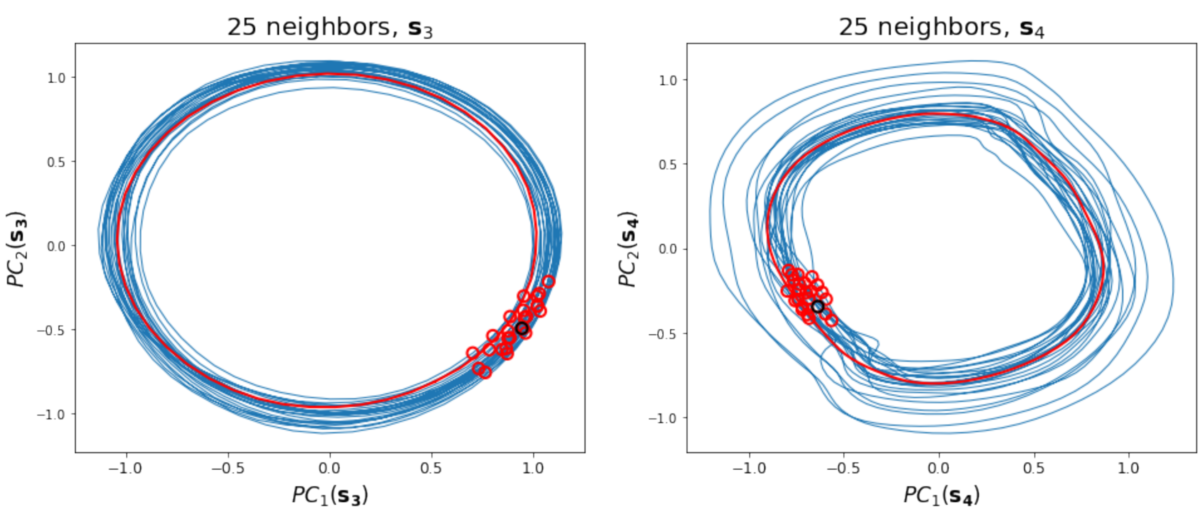


Рис. 4: Метод CCM для проверки наличия связи, PLS

5.2 Несвязанные сигналы, акселерометр + акселерометр

Среднеквадратичная ошибка MSE $= 0.86$

5.3 Медленная ходьба, акселерометр + гироскоп

Рис. 5: Предсказание для зависимых сигналов, PLS

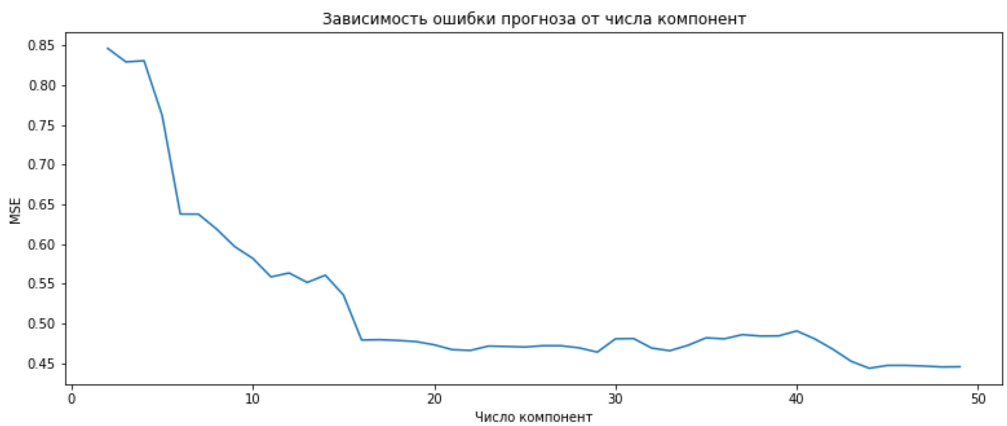


Рис. 6: Зависимость ошибки прогноза от числа компонент, зависимые ряды

Рис. 7: Предсказание для несвязанных сигналов, PLS

Рис. 8: Зависимость ошибки прогноза от числа компонент, несвязанные ряды

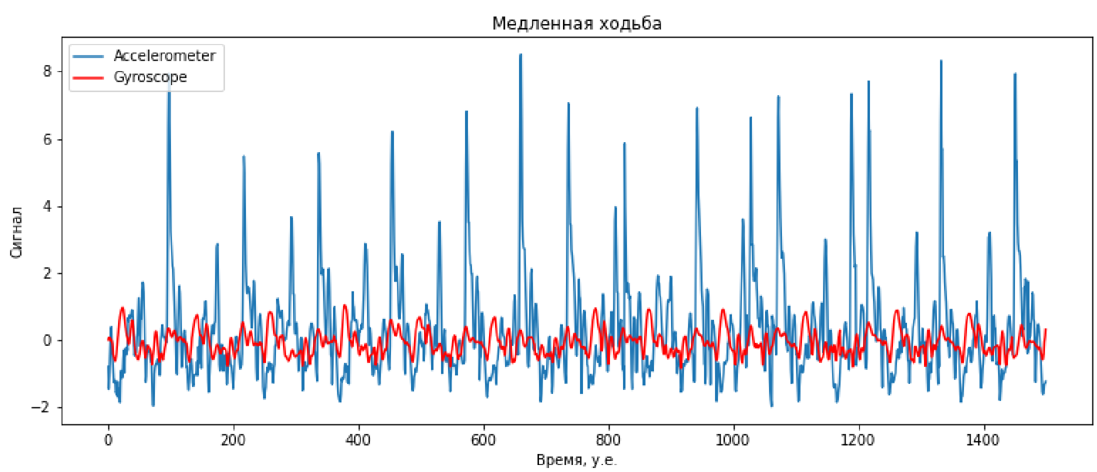
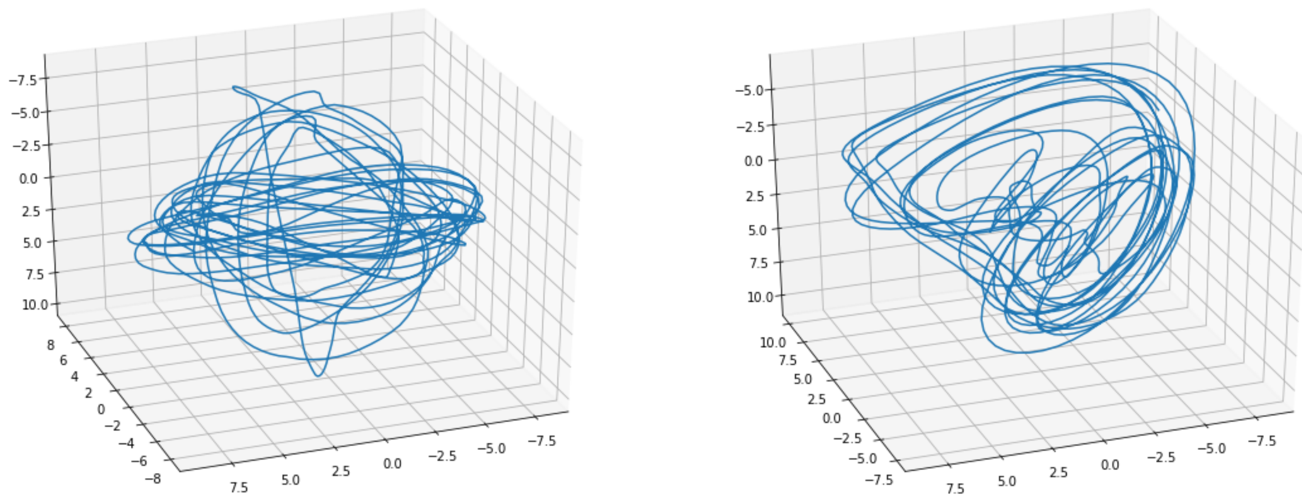


Рис. 9: Данные акселерометра и гироскопа одного мобильного устройтсва



(a) Акселерометр (b) Гироскоп

Рис. 10: Траектории в фазовом пространстве

Рис. 11: Предсказание сигнала гироскопа, PLS