Детекция зависимостей во временных рядах

I. M. Latypov, E. Vladimirov, V. V. Strizhov
latypov.im@phystech.edu
MIPT

Для многих прикладных задач [2][3][6] требуется решать задачу выявления зависимостей между временными рядами. Выявление этих зависимостей и их последующее использование направлено на то, чтобы улучшить качество модели. В статье для обнаружения подобных зависимостей предлагается использовать *модели пространства состояний*. Приведен пример использования метода на основе скрытых состояния модели ODE-RNN [9]. Скрытые состояния рассматриваются как представление временного ряда, и после этого к ним примяется метод сходящегося перекрестного отображения (Convergent Cross Mapping) [10].

Ключевые слова: *временные ряды, CCM, ODE-RNN, Neural ODE*

1 Введение

Решается задача поиска причинно-следственных связей между временными рядами. На практике у изучаемой динамической системы несколько наблюдаемых величин, измерения которых представляют собой временной ряд. Исследование этих рядов входит в состав задачи исследования системы. Выявление причинно-следственных взаимосвязей между временными рядами наблюдаемых величин рассматривается – важная часть исследования временных рядов. Например на основе анализа зависимости временных рядов данных гироскопов у танцующей пары можно делать выводы о качестве их взаимодействия.

2 О подходах к решению задачи

Кросс Корреляция – метод проверяет корррреляцию временных рядов при из сдвигах. Зависимость оценивается на основе максимальной полученной корреляции.

Тест Гренжера [5] - на паре временных рядов обучаются две модели: первая обучается предсказвать первый временной ряд только на данных этого ряда. Вторая тоже обучается предсказвыать первый временной ряд, но уже на данных обоих временных рядов. Если качество предсказаний на второй модели существенно возрастает, то делается вывод о зависимости временных рядов.

Кластеризация квазипериодических временных рядов [1] – использование метода главных компонент с новой метрикой.

[7] – описывает метод построения описания объекта на основе экспертно определенных генерирующих функций.

Метод перекрестного сходящегося отображения [8] будет рассмотрен далее в деталях как основа для построения нашего метода.

Общим недостатком для всех методов является квадратичное от длительности наблюдений время работы. Другие достоинства и недостатки некоторых методов можно посмотреть в таблице 1.

Таблица 1 Сравнение методов

Метод	Достоинства	Недостатки
Тест Грэнджера [5]	Анализ рядов совмещается с построением модели предсказания.	Не дает представлений о виде зависимости рядов. Предсказания могут не улучшиться из-за неверной модели.
Кросс Корреляция [12]	Не требуется дополнительная обработка данных	Корреляция не является достаточным условием зависимости.
CCM [10]	Возможна работа с более сложными зависимостями чем в предыдущих методах.	Исследование [8] выделяет недостатки.
CCM + ODE-RNN (предлагаемый)	Применяется к многомерными временными рядам.	Так как метод параметрический, его нужно обучать на исследуемых данных.

3 Математическая постановка

Обозначим $T = \{t_1, ...t_k\}$ - моменты наблюдений. И введем обозначения

$$\mathbf{x} = \{\mathbf{x_1}, \mathbf{x_2}, ...\mathbf{x_k}\}, \mathbf{y} = \{\mathbf{y_1}, \mathbf{y_2}, ...\mathbf{y_k}\}$$

наблюдения за парой многомерных временных рядов. $x_i \in \mathbb{R}^m$, $y_i \in \mathbb{R}^n$, в работе опыты проводятся при $m = n = 3$. Наблюдения x_i, y_i сделаны в момент t_i. Промежутки между наблюдениями одинаковы, то есть частота семплирования рядов постоянна.

Ставится задача построения отображения $\varphi : \{\mathbf{x} \times \mathbf{y}\} \to \mathbb{R}$ по значениям которой будет делаться вывод о зависимости временных рядов.

4 предлагаемый метод

В качестве основы для модели берется метод сходящегося перекрестного отображения, поэтому рассмотрим его подробно. Метод применяется для пары одномерных временных рядов $\mathbf{x} = [x_1, ..., x_N]$ и $\mathbf{y} = [y_1, ..., y_N]$. Индексирование такое же как в **постановке**. Идея метода основана на теореме Таккенса [11]. Теорема утверждает, что для временного ряда, представляющего собой наблюдения за динамической системой и удовлетворяющего перечисленным выше условиям – а именно для рядов с постоянной частотой семплирования, по векторам $[x_i, x_{i+1}, ..., x_{L-1}]$ строится изоморфизм в пространство, в котором развивается динамическая система. То есть такие вектора описывают динамическую систему. Здесь L - размерность отображения. Далее будем называть это размерностью погружения. Согласно этой теореме должно быть выполнено $L \geqslant 2m$, где m - истинная размерность системы. На практике m нам не известно.

Метод рассматривает отображение временных рядов в траекторное подпространство с матрицей Ганкеля временного ряда и оценивает, насколько хорошо траектория эволюции одного ряда восставнавливается по траекторий эволюций другого. Опишем как это

делается. Для рядов строится матрица Ганкеля:

$$\mathbf{H}_x = \begin{pmatrix} x_1 & x_2 & ... & x_{L-1} & x_L \\ x_2 & x_3 & ... & x_L & x_{L+1} \\ ... & ... & ... & ... & ... \\ x_{N-L+1} & x_{N-L+2} & ... & x_{N-1} & x_N \end{pmatrix}$$

L - размерность погружения, используемая для построения отображения в пространство эволюции системы. Так же строится матрица Ганкеля второго ряда $\mathbf{H}_y$. Обозначим через $\mathbf{x}_t$ $t-L$-ую строку H_x, $\mathbf{y}_t$ $t-L$-ую строку H_y. Тогда вектора $\mathbf{x}_t$ рассматриваются как точки в траекторнорм пространстве $\mathbf{M}_x$, $\mathbf{y}_t$ - как точки в траекторнорм пространстве $\mathbf{M}_y$. В этих пространствах выбирается евклидова метрика. Для восстановления измерения y в момент времени $t \in L+1, ..., N$ найдем k ближайших соседей вектора $\mathbf{x}_t$. Обозначим их по возрастанию расстояния до $\mathbf{x}_t$:

$$[\mathbf{x}_{t_1}, \mathbf{x}_{t_2}, ..., \mathbf{x}_{t_k}]$$

. $\mathbf{x}_{t_k}$ – самый дальний их k соседей. После этого строится прогноз y_t следующим образом:

$$y^t = \sum_{i=1}^{k} w_i y_{t_i}$$

Где

$$w_i = \frac{u^i}{\sum_i u_i}, \qquad u_i = \exp\left(- \frac{\|\mathbf{x_t} - \mathbf{x_{t_i}}\|_2}{\|\mathbf{x_t} - \mathbf{x_{t_k}}\|_2} \right)$$

Для обнаружения зависимости рядов рассматривается корреляция между предсказаниями и значениями ряда. На основании величины корреляции делается вывод о зависимости или независимости рядов.

Чтобы развить этот метод рассмотрим параметрическое построение погружений. Для этого обратимся к моделям пространства состояний – моделям дискретного описания динамической системы. При таком подходе совсместно с временным рядом рассматривается дополнительный вектор скрытых состояний, который эволюционирует совместно с наблюдениями за системой.

В самом простом виде уравнения развития скрытых состояний системы можно задать как два уравнения эволюции. Вектор скрытых состояний системы u, вектор наблюдений за системой x и уравнения:

$$\begin{aligned} u_t &= F(u_{t-1}, y_t) \\ z_t &= G(u_t, y_t) \end{aligned} \tag{1}$$

Здесь z – моделируемая величина. Второе уравнение нам не нужно, так что далее рассматриваем только первое, из которого получаются скрытые состояния.

Итоговая модель выглядит просто:

$$\begin{aligned} u_0 &= f(x_0) \\ u_{k+1} &= \psi(u_k, x_k) \end{aligned} \tag{2}$$

Мы выделили два различных способа задания функции ψ – непрерывное и дискретное и рассмотрели их в экспериментах. Дискретное изменение моделирорвалось с помощью

GRU модуля. Для моделирования непрерывных изменений была взята модель ODE-RNN [9],

ODE-RNN строится на последовательном применении - Neural ODE [4] и RNN. В качестве u_i – используются скрытые состояния RNN модуля. В моменты наблюдений скрытое состояние меняется с RNN $u_{t_i} = RNN(\widetilde{u_{t_i}}, x_{t_i})$, затем полученное скрытое состояние эволюционирует к следующему моменту времени с NeuralODE: $u_{t_{i+1}} = = \text{ODESolve}(f_\theta, u_{t_i}, (t_i, t_{i+1}))$. Здесь f_θ – уравнение использующееся в качестве дифференцциального уравнения. ODESolve – вызов Neural ODE. Видно, что модель работает независимо от промежутков между семплированими.

5 Вычислительные эксперименты

Кратко опишем выбрку на которой проводились эксперименты. Данные представляют собой записи показаний акселерометра и гироскопа при ходьбе и при беге на протяжении 30 – 40 секунд. За это время делается примерно 15 движений. Частота семплирования равна 200 Гц. Датчики находятся на концах рук, так что можем рассматривать пары временных рядов и применять к ним метод. Можно посмотреть пример временного ряда гироскопа на картинке 1.

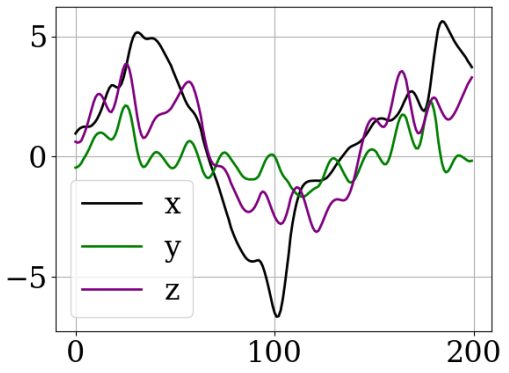

Рис. 1 Ряд значений показаний гироскопа на отрезке в 250 семплов

Сначала посмотрим на работу метода на паре одномерных временных рядов. В качестве функции ψ используется ODE-RNN. Отрисуем траектории скрытых состояний временных рядов в скрытых состояниях, которые получаются методом CCM. Для этого из погружений рядов – матрицы Ганкеля выделяются главные компоненты. Подробнее в работе [14]. Для наглядности траектории были спроецированы на сферу.

В предложенном методе матрица скрытых состояний получается не ганкелевой, поэтому к ней перед построением траекторий применяется метод многомерной гусеницы (анализ сингулярного спектра) [13] Результаты можно увидеть на рис. 3 и 4.

То есть в скрытых состояниях появляются одинаковые периодические структуры. Применим метод CCM к скрытым состояниям. Для сравнения приведен результат применения CCM к самим временным рядам. Результаты можно увидеть на рис. 5.

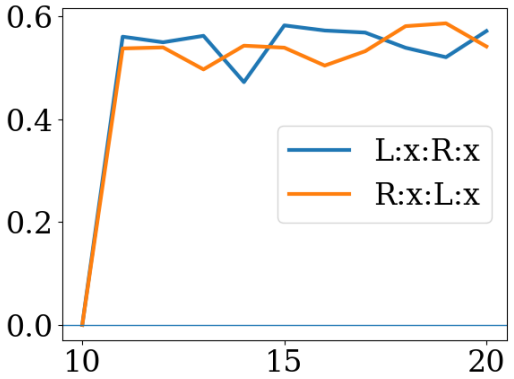

Рис. 2 ODE-RNN + CCM для правого и левого гироскопа на трехмерных данных. Оси: ox – количество опорных точек метода CCM, oy – корреляция, получаемая при применении метода

Видно что на рассматриваемых данных предложенный метод так же выделяет зависимость. Как отмечалось ранее - наблюдения на датасете с акселерометром - трехмерные, и предложенный метод создается чтобы работать именно с такими данными. Посмотреть траектории скрытых состояний и результат применения метода к данным можно на рис. 2 и 6.

6 анализ свойств предложенного метода

Из рисунка 4 видно, что траектории предложенного метода совпадают не так хорошо как у CCM. Это происходит из-за того, что модель параметрическая. ODE-RNN обучаются на разных данных, и поэтому достигаются разные локальные минимумы.

Следующий важный гиперпараметр при подготовке модели - количество этох для тренировки. При экспериментах модель обучалась 9 эпох. Если обучение длится мало, то корреляция снижается, а траектории сильно отличаются. Если же модель переобучается, то траектории становятся хаотичными и корреляция соответственно значительно уменьшается.

Так же важна размерность погружения. Если взять слишком большую размерность, то корреляция начинает уменьшаться. На это предположительно есть две причины. Первая связана с проклятием размерности – евклидова метрика плохо работает на большиз размерностях. Вторая – модель слишком сложная для рассматриваемых данных, из-за чего некоторое подмножество весов плохо обучается.

7 Заключение

Предложили метод для выявления взаимной зависимости временных рядов. В отличие от других популярных методов его можно применить к многомерным данным. К тому же метод является параметрическим. Это и достаток метода – после обучения выявляются более глубокие признаки данных, и его недостаток – необходимо подбирать параметры и обучать модель.

Опыты показали, что метод работает с парой измерений датчиков на теле так же как и CCM. Кроме того и одномерные, и многомерные данные погружаются в структуры, похожие на погружения при методе CCM. То есть информация о периодичности рядов, их частоте и другие данные сохраняются при погружении.

Границы применимости метода пока что открытый вопрос.

8 Список литературы

9 *

Список литературы

[1] V. V. Strijov A. V. Grabovoy. Quasi-periodic time series clustering for human activity recognition. *http://strijov.com/papers/Grabovoy2019QuasiPeriodicTimeSeries.pdf*, 2018.

[2] Giuseppe Averta, Federica Barontini, Vincenzo Catrambone, Sami Haddadin, Giacomo Handjaras, Jeremia P O Held, Tingli Hu, Eike Jakubowitz, Christoph M Kanzler, Johannes Kühn, Olivier Lambercy, Andrea Leo, Alina Obermeier, Emiliano Ricciardi, Anne Schwarz, Gaetano Valenza, Antonio Bicchi, and Matteo Bianchi. U-Limb: A multi-modal, multi-center database on arm motion control in healthy and post-stroke conditions. *GigaScience*, 10(6), 06 2021. giab043.

[3] Jonathan Camargo, Aditya Ramanathan, Will Flanagan, and Aaron Young. A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions. *Journal of Biomechanics*, 119:110320, 2021.

[4] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

[5] C.W.J. Granger. Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control*, 2:329–352, 1980.

[6] Tomoyuki Higuchi. Applications of quasi-periodic oscillation models to seasonal small count time series. *Computational Statistics & Data Analysis*, 30(3):281–301, 1999.

[7] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. "activity recognition using cell phone accelerometers,". *Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data*, 2010.

[8] James M. McCracken and Robert S. Weigel. Convergent cross-mapping and pairwise asymmetric inference. *Physical Review E*, 90(6), dec 2014.

[9] Yulia Rubanova, Ricky T. Q. Chen, and David Duvenaud. Latent odes for irregularly-sampled time series, 2019.

[10] George Sugihara, Robert May, Hao Ye, Chih hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. Detecting causality in complex ecosystems. *Science*, 2012.

[11] Floris Takkens. Detecting strange attractors in turbulence. *Springer Lecture Notes in Mathematics vol 898, pp 366–81*, 1981.

[12] Jae-Chern Yoo and Tae Hee Han. Fast normalized cross-correlation. *Circuits, systems and signal processing*, 28:819–843, 2009.

[13] Под редакцией Д.Л.Данилова и А.А.Жиглявского. Главные компоненты временных рядов: метод "Гусеница". *Санкт-Петербургский университет*, 1997.

[14] В.В. Стрижов К.Р. Усманова. МОДЕЛИ ОБНАРУЖЕНИЯ ЗАВИСИМОСТЕЙВО ВРЕМЕННЫХ РЯДАХ В ЗАДАЧАХ ПОСТРОЕНИЯПРОГНОСТИЧЕСКИХ МОДЕЛЕЙ. *«Системы и средства информатики»*, 2018.

10 картинки

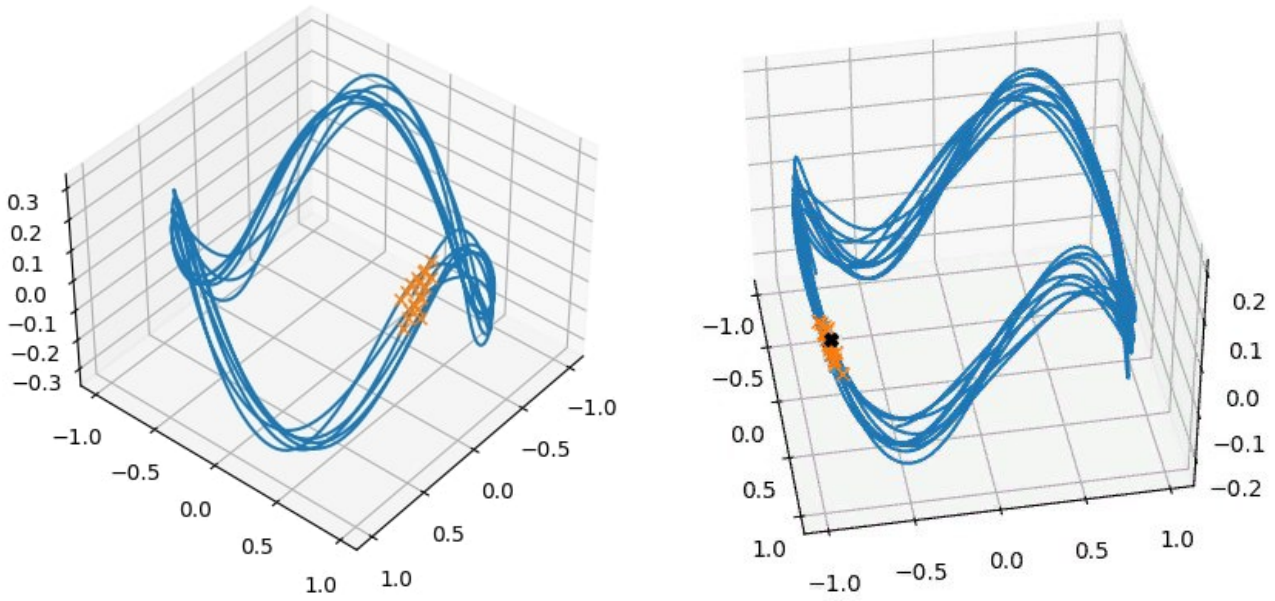

Рис. 3 слева направо - траектория скрытых состояний при использовании CCM левого/правого гироскопа на одномерных данных

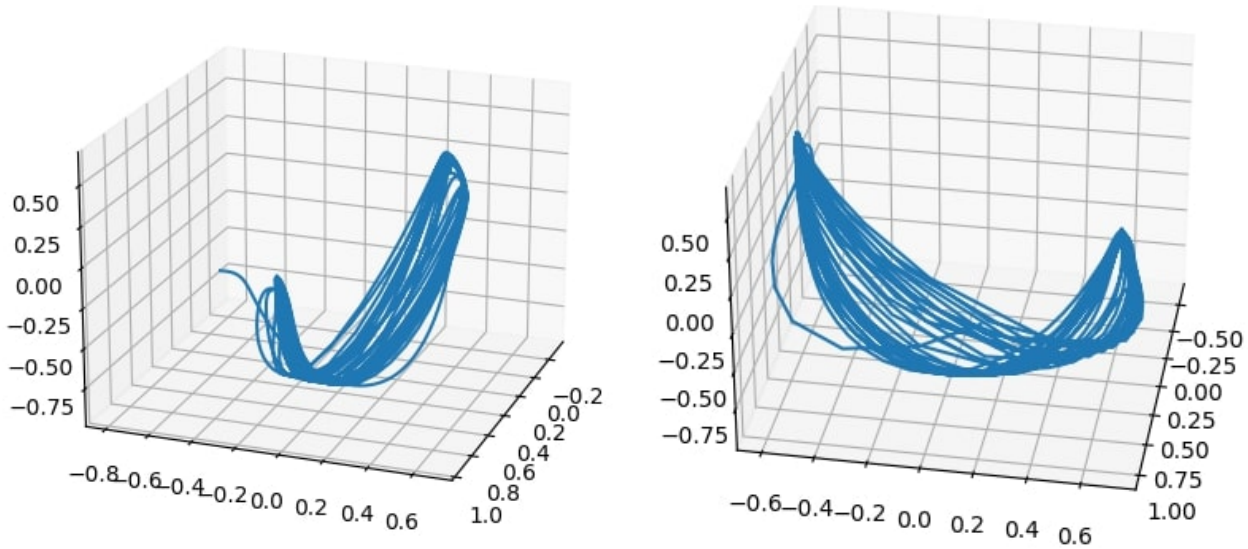

Рис. 4 слева направо - траектория скрытых состояний при использовании ODE-RNN + гусеница левого/правого гироскопа на одномерных данных

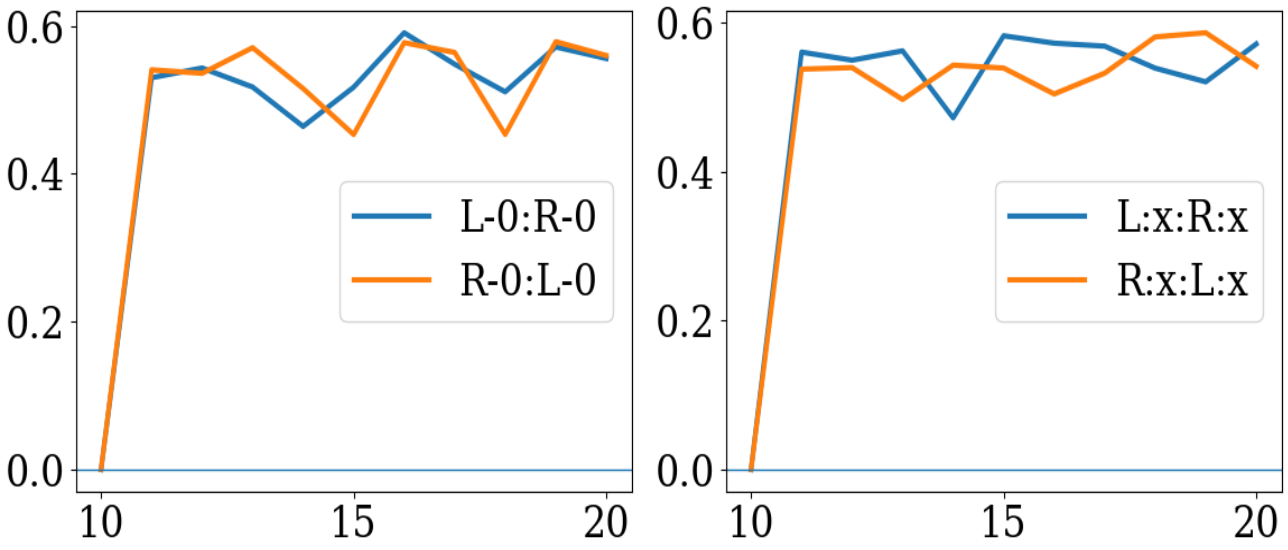

Рис. 5 слева направо: ODE-RNN + CCM / CCM для правого и левого гироскопа на одномерных данных. Оси: ox – количество опорных точек метода CCM, oy – корреляция, получаемая при применении метода

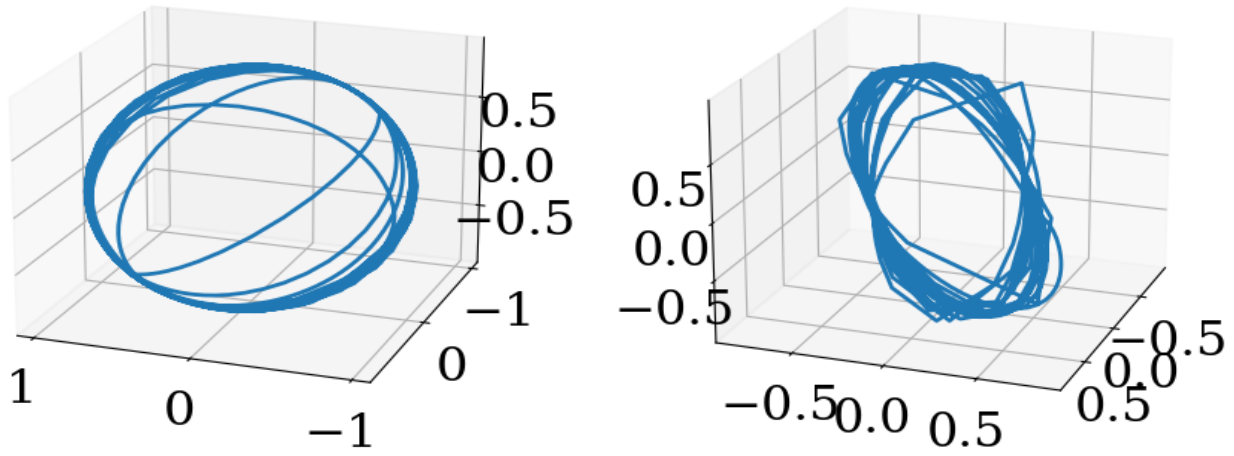

Рис. 6 слева направо - траектория скрытых состояний при использовании ODE-RNN + гусеница левого/правого гироскопа на трехмерных данных