

Детекция зависимостей во временных рядах

I. M. Latypov, E. Vladimirov, V. V. Strizhov

latypov.im@phystech.edu

MIPT

Для многих прикладных задач, связанных с динамическими системами, требуется решать задачу выявления причинно-следственных зависимостей между временными рядами. Выявление этих зависимостей может улучшить качество модели, или рассматриваться как отдельная задача классификации. В данной работе для обнаружения подобных зависимостей предлагается использовать ****модели пространства состояний****: скрытые состояния модели ODE-RNN рассматриваются как представление временного ряда, и показывается что к этому представлению применим метод сходящегося перекрестного отображения (Convergent Cross Mapping). Работа метода проверяется на данных акселерометра и гироскопа при выполнении упражнений.

Ключевые слова: *Neural CDE, CCM, временные ряды*

1 Введение

Работа посвящена задаче поиска причинно-следственных связей между временными рядами. Эта задача актуальна, поскольку на практике у изучаемой динамической системы несколько наблюдаемых величин [3], измерения которых представляют собой временной ряд. Исследование этих рядов является частью задачи исследования системы. Выявление причинно-следственных взаимосвязей между этими временными рядами наблюдаемых величин может рассматриваться как подзадача или как основная задача, например на основе анализа зависимости временных рядов данных гироскопов у танцующей пары можно делать выводы о качестве их взаимодействия.

2 Связанные работы

Кросс Корреляция - метод проверяет корреляцию временных рядов при их сдвигах. Зависимость оценивается на основе максимальной полученной корреляции.

Тест Гренжера [5] - обучается модель для предсказания одного ряда. Обучается другая модель для предсказания второго ряда. Если качество предсказаний на второй модели существенно возрастает, то делается вывод о зависимости временных рядов.

Кластеризация квазипериодических временных рядов для распознавания человеческой деятельности [1] - использование метода главных компонент с новой метрикой, [2] - описывает метод построения описания объекта на основе экспертно определенных генерирующих функций. Метод перекрестного сходящегося отображения [4] будет рассмотрен далее в деталях как основа для построения нашего метода.

Приведем достоинства и недостатки некоторых из них в таблице 1.

3 Математическая постановка

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

$$\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}, \mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$$

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.

Метод	достоинства	недостатки
Тест Грэнджера	Легко применять	Не дает представлений о виде зависимости рядов. Предсказания могут не улучшиться из-за неверной модели.
Кросс Корреляция	Легко применять	Квадратичное от длины ряда время работы. Корреляция не является достаточным условием зависимости
ССМ	Легко применять. Работает лучше методов предложенных выше.	Квадратичное от длины ряда время работы. Исследование \cite{McCracken2014} выделяет другие недостатки.
ССМ + ODE-RNN (предлагаемый)	Может работать с многомерными временными рядами	Нужно обучать Квадратичное от длины ряда время работы

Ставится задача построения отображения $\varphi : \{\curvearrowright \times \curvearrowright\} \rightarrow \mathbb{R}$ 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 $\mathbf{x} = [x_1, \dots, x_N]$ и $\mathbf{y} = [y_1, \dots, y_N]$. Индексирование такое же как в постановке. Идея метода основана не на теореме Такенса [6]. Теорема утверждает, что для временного ряда, представляющего собой наблюдения за динамической системой и удовлетворяющего перечисленным выше условиям, по векторам $[x_i, x_{i+1}, \dots, x_{L-1}]$ можно построить 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 теореме должно быть выполнено $L \geq 2m$, где m - истинная размерность системы. Чаще всего 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 воссоставляется по траекторий эволюций другого. Опишем как это делается. По рядам строится матрица Ганкеля:

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

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+L-1} t -ую строку \mathbf{H}_x , \mathbf{y}_{t+L-1} t -ую строку \mathbf{H}_y . Тогда вектора \mathbf{x}_t рассматриваются как точки в траекторном пространстве \mathbf{M}_x , t - как точки в траекторном пространстве \mathbf{M}_y . В этих пространствах выбирается евклидова метрика. Для восстановления измерения y в момент времени $t \in L, \dots, N$ найдем k ближайших соседей вектора \mathbf{x}_t . Обозначим их по возрастанию расстояния до \mathbf{x}_t :

$$\mathbf{x}_{t_1}, \mathbf{x}_{t_2}, \dots, \mathbf{x}_{t_k}$$

После этого строится прогноз y_t следующим образом:

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

$$w_i = \frac{u^i}{\sum_i u_i}, \quad u_i = \exp\left(-\frac{\|\mathbf{x}_t - \mathbf{x}_{t_i}\|_2}{\|\mathbf{x}_t - \mathbf{x}_{t_k}\|_2}\right)$$

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 совместно с временным рядом рассматривается

дополнительный вектор скрытых состояний, который эволюционирует совместно с наблюдениями за системой.

В самом простом виде уравнения развития скрытых состояний системы могут выглядеть следующим образом: вектор скрытых состояний системы u , вектор наблюдений за системой x

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

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

Итоговая модель выглядит крайне просто:

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

Мы выделили два принципиально различных способа задания функции ψ – непрерывное и дискретное и рассмотрели их в экспериментах. Дискретное изменение моделировалось с помощью GRU модуля. Для моделирования непрерывных изменений была взята модель ODE-RNN [?].

возможно понадобится краткое описание ODE-RNN

5 анализ свойств предложенного метода

Предлагаю вообще убрать этот раздел

6 Эксперименты

Кратко опишем выборку 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

Сначала рассмотрим работу метода на паре одномерных временных рядов. В качестве функции ψ используется ODE-RNN. Рассмотрим траектории временных рядов в скрытых состояниях. Для этого в методе CCM берется матрица ганкеля и из него выделяются главные компоненты. Подробнее в работе [8]. получившиеся траектории можно увидеть на картинках 2. Для наглядности траектории были спроецированы на сферу.

В нашем методе матрица скрытых состояний получается не ганкелевой, поэтому к ней перед построением траекторий применяется метод многомерной гусеницы (анализ сингулярного спектра) [7]

То есть в скрытых состояниях появляются одинаковые периодические структуры, поэтому можем применить метод CCM скрытым состояниям. Результат применения можно

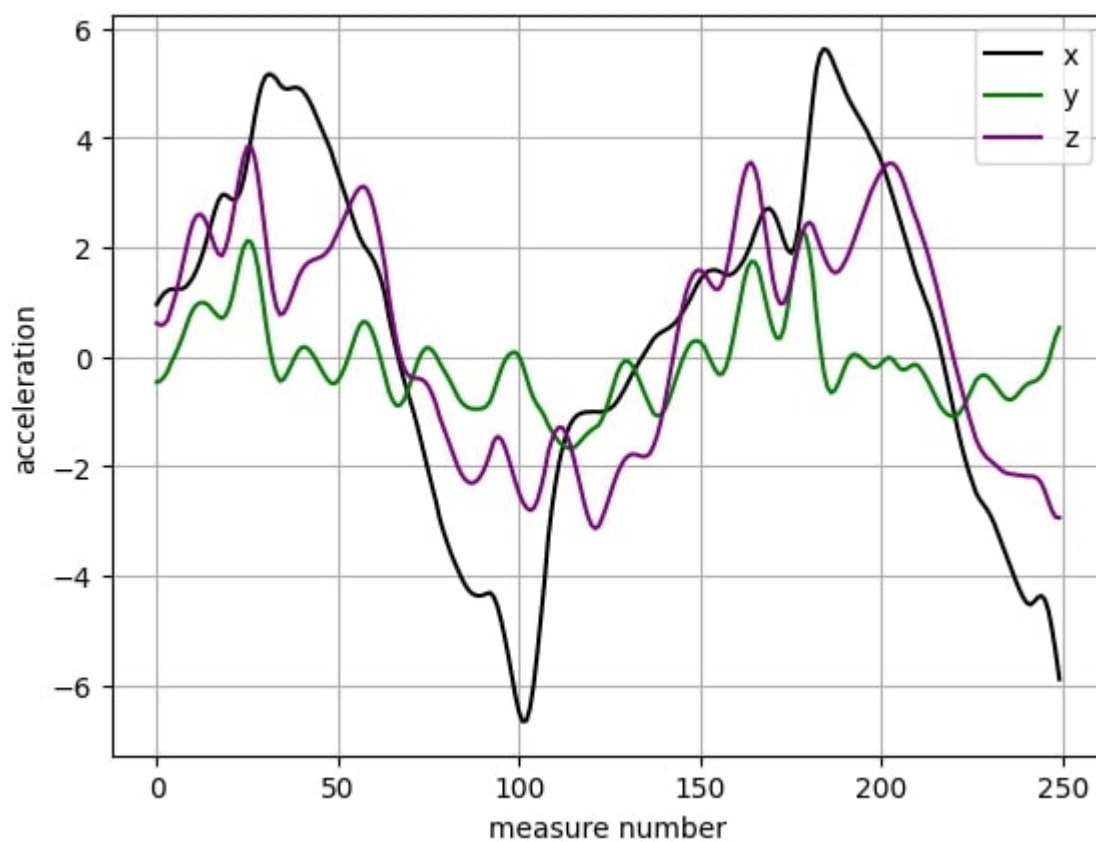


Рис. 1 Ряд значений показаний гироскопа на отрезке в 250 семплов

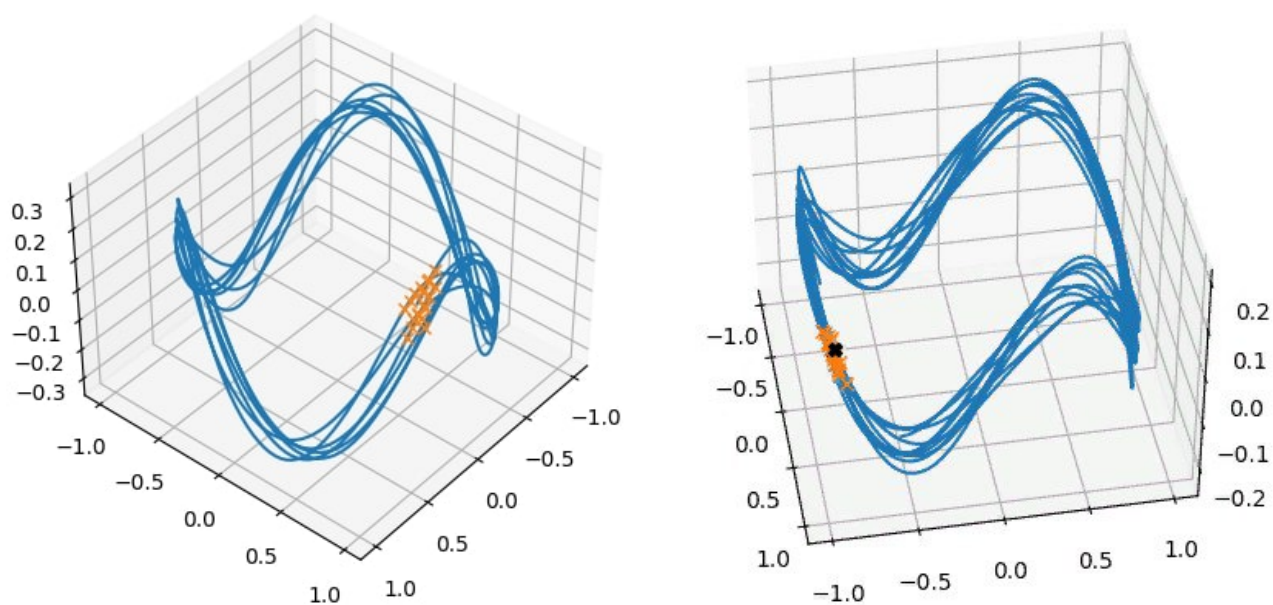


Рис. 2 слева направо - траектория скрытых состояний при использовании ССМ левого/правого гироскопа

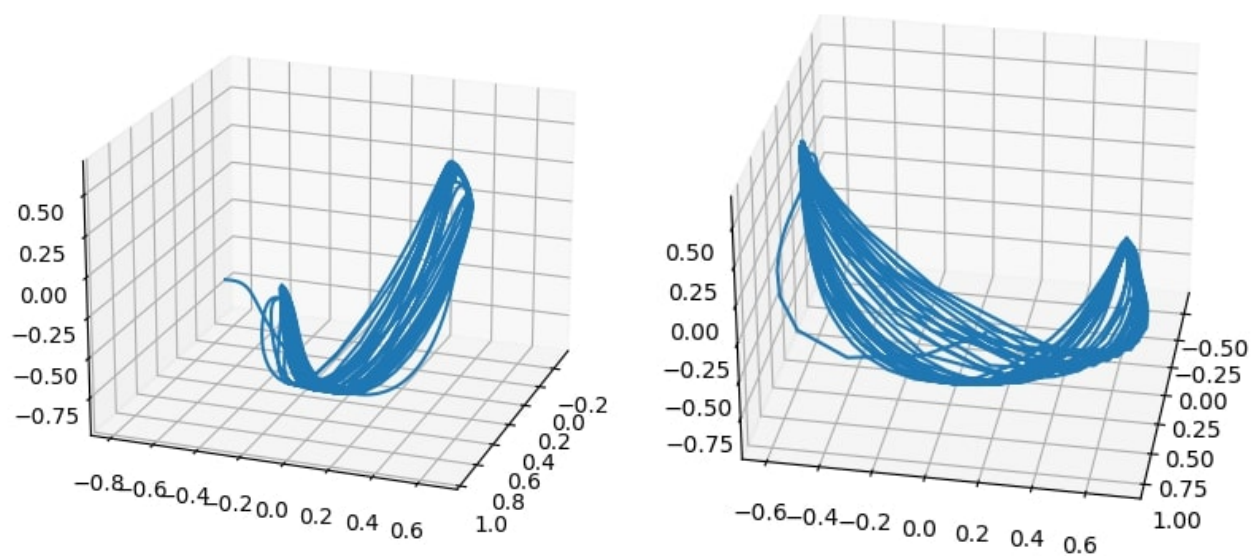


Рис. 3 слева направо - траектория скрытых состояний при использовании ODE-RNN + гусеница левого/правого гироскопа

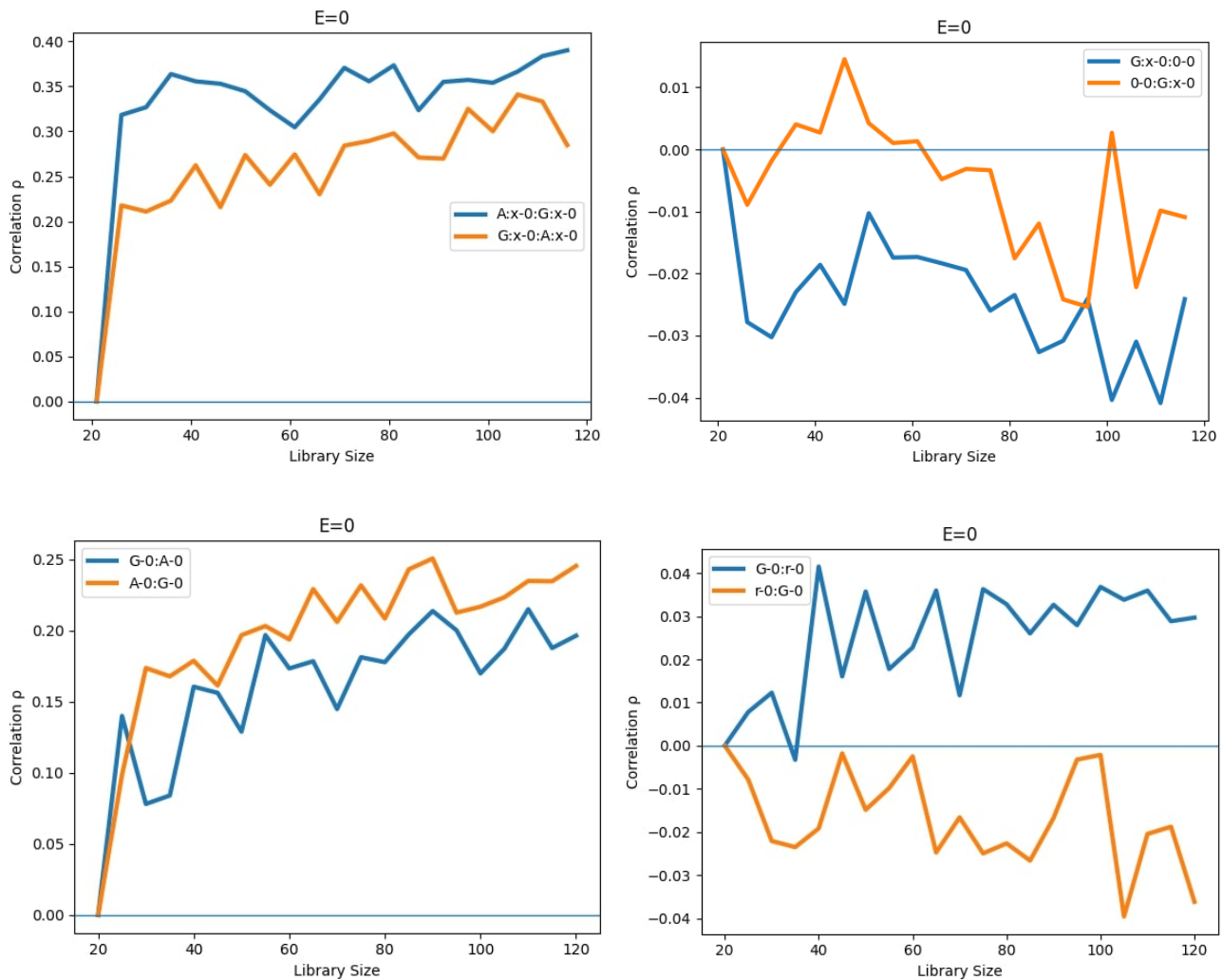


Рис. 4 слева направо, сверху вниз: CCM акселерометр - гироскоп, CCM акселерометр - random, ODE-RNN акселерометр - гироскоп, ODE-RNN акселерометр - random

увидеть на графиках . Для сравнения приведен результат применения методов к паре гироскоп - случайная последовательность.

 нужно будет поменять label в графиках

Видно что предлагаемый метод может выделять зависимость и независимость на рассмотренных данных так же как CCM.

Но это не единственная возможность метода. Как отмечалось ранее - наблюдения на датасете с акселерометром - трехмерные.

 красивые картинки про эксперименты с трехмерными данными. Не совсем понимаю что поставить

7 Заключение

Предложили метод для выявления взаимной зависимости временных рядов. Опыты показали что он работает на паре измерений двух датчиков на теле. В отличие от своих аналогов он позволяет работать с многомерными временными рядами. Границы применимости метода — пока 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] 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.
- [3] machinelearning.ru.
- [4] James M. McCracken and Robert S. Weigel. Convergent cross-mapping and pairwise asymmetric inference. *Physical Review E*, 90(6), dec 2014.
- [5] A. Seth. Granger causality. *Scholarpedia*, 2(7):1667, 2007. revision #127333.
- [6] Floris Takken. Detecting strange attractors in turbulence. *Springer Lecture Notes in Mathematics vol 898*, pp 366–81, 1981.
- [7] Под редакцией Д.Л.Данилова и А.А.Жиглявского. Главные компоненты временных рядов: метод "Гусеница". *Санкт-Петербургский университет*, 1997.
- [8] В.В. Стрижов К.Р. Усманова. МОДЕЛИ ОБНАРУЖЕНИЯ ЗАВИСИМОСТЕЙ ВО ВРЕМЕННЫХ РЯДАХ В ЗАДАЧАХ ПОСТРОЕНИЯ ПРОГНОСТИЧЕСКИХ МОДЕЛЕЙ. «Системы и средства информатики», 2018.