

ПОИСК И ВОССТАНОВЛЕНИЕ 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 Space State Modelling (SSM) на архитектурах основанных на Neural ODE. Работа метода проверяется на данных акселерометра и гироскопа при беге и других действиях. Также предложен способ оценки качества взаимодействия партнеров на парных танцах.

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

1 Введение

Работа посвящена задаче поиска причинно-следственных связей между временными рядами. Эта задача актуальна, поскольку на практике у изучаемой системы несколько наблюдаемых величин [4], исследование которых является частью задачи исследования системы. Выявление причинно-следственных взаимосвязей между 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.

Существует множество методов для обнаружения связей между временными рядами. Среди них тест Гренжера, метод сходящегося перекрестного отображения (convergent cross mapping, CCM) и другие. Принципы их работы будут рассмотрены в разделе "связанные работы". Далее в разделе математическая постановка приведена строгая постановка задачи и предлагаемый метод решения. В разделе теоритичесие предпосылки приведены теоритические предпосылки на которых основывается предлагаемый метод.

Кратко опишем данные на которых проводятся исследования.

Первый датасет - данные акселерометра и гироскопа, собранные за время выполнения упражнений (бег, ходьба и т.д.). Точки предскавляют собой трехмерные вектора - значения измерений в трех направлениях. Ожидается что на парах наблюдений акселерометр-гироскоп метод будет показывать сильную зависимость.

Второй датасет - данные пар акселерометров (тут должна быть ссылка, но я не нашел датасет: ().

2 Связанные работы

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

Тест Гренжера [2] - 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.

Идея работы CCM [7] основана не теореме Таккенса [8]. Метод рассматривает отображение временных рядов в траекторное подпространство с матрицей Ганкеля временного

1. скорее всего для SSM есть термин на русском языке. Лучше использовать его, раз статья на русском
2. Очень коротко, лучше развить и добавить пару поясняющих предложений
3. Если тут CCM, имеет смысл этот термин тоже ввести в аннотации
4. Проверьте по википедии как пишется фамилия на русском. И нужна ссылка.
5. Кавычки пишутся по-другому. См. учебник по латеху (или хотя бы мою презентацию от 9 марта)
6. тут нужны кавычки
7. датасет => выборка
8. а вообще, не уверен что этот текст нужен здесь. я бы его перенес в эксперимент
9. Обзор литературы
10. вместо дефиса нужно длинное тире (проверить по всему тексту)
11. перед цитированием нужен пробел (проверить по всему тексту)

ряда и оценивает "схожесть" траекторий эволюций этих рядов предсказывая один ряд с помощью другого.

А так же [1] - использование метода главных компонент с новой метрикой, [3] - описывает метод построения описания объекта на основе экспертно определенных генерирующих функций.

Приведем достоинства и недостатки некоторых из них в таблице 1.

Метод	достоинства	недостатки
Тест Гренжера	Легко применять	Не дает представлений о виде зависимости рядов. К тому же предсказания могут не улучшиться из-за неверной модели.
Кросс Корреляция	Легко применять	Приходится использовать весь датасет + квадратичное от длины ряда время работы. Так же хорошо известно, что корреляция не является достаточным условием зависимости
ССМ	Легко применять. Работает лучше методов предложенных выше.	Использование всего датасета, квадратичное от длины ряда время работы. Так же исследование \ выделяет другие недостатки.
CMM + ODE-RNN	Точнее может выявлять зависимости между временными рядами	Нужно обучать на достаточно большом куске данных

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 исследуются при $m = n = 3$. \mathbf{x}_i -наблюдение в момент времени t_i

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

Для построения модели берется State space model (модель в пространстве состояний) - модель дискретного описания динамической системы. При таком подходе вводится вектор скрытых состояний системы u и вектор наблюдений за системой v и задаются уравнения их изменений:

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

пока не разобрался что здесь представляет собой z_t , скоро будет

Получаем идею для модели

12. формулы по возможности лучше делать без переносов
13. обозначение понятное, но не очень корректное, лучше бы $R^m \times R^n$
14. дефис
15. если вектор, то жирный

$$\begin{aligned}\gamma_0 &= f(x_0) \\ \gamma_{k+1} &= \psi(\gamma_k, x_k)\end{aligned}\tag{2}$$

Мы рассмотрели 3 модели:

1. Дискретные функции перехода - простой RNN, GRU.
2. Непрерывные переходы – скрытое состояние между наблюдениями меняется непрерывно с помощью [5], (ссылка на ODE-RNN)
3. Непрерывные переходы с предсказанием - coming soon

4 теоретические предпосылки

Основой для построения метода послужила теорема Таккенса и метод CCM основанный на ней.

Пусть M - компактное многообразие размерности m

Динамическая система 1. динамической системой φ с дискретным временем на M назовем диффеоморфизм $\varphi : M \rightarrow M$.

Эволюция системы начинается с $x_0 \in M$ и следующие состояния определяются рекурсивно $x_{t+1} = \varphi(x_t)$.

Пусть φ - динамическая система. Определим гладкую функцию $y : M \rightarrow \mathbb{R}$ - наблюдения за динамической системой. Хочется из набора наблюдений получить информацию об эволюции системы φ в M .

Для этого нам понадобится следующая теорема

Таккенс [8] 1. Пусть M - компактное многообразие размерности m . Для пар (φ, y) , где $\varphi : M \rightarrow M$ гладкий диффеоморфизм и $y : M \rightarrow \mathbb{R}$ гладкая функция, общим свойством является то, что $\Phi : M \rightarrow \mathbb{R}^{2m+1}$, определяемое как

$$\Phi_{(\varphi, y)}(x) = (y(x), y \circ \varphi(x), \dots, y \circ \varphi^{2m}(x))$$

является вложением $A(\varphi) \rightarrow \mathbb{R}^{2m+1}$. Под гладкостью понимается дважды непрерывная дифференцируемость.

$A(\varphi)$ - аттрактор динамической системы.

То есть для изучения свойств траектории динамической системы *phi* можно использовать траектории, которые получаются при этих погружениях.

Пусть даны две динамические системы: φ с наблюдениями x на многообразии M_x и ψ с наблюдениями y на многообразии M_y . Нужно выяснить, есть ли причинно-следственная связь между системами. Поскольку мы не имеем представления об устройстве многообразий, рассмотрим их вложения $\Phi(\varphi, x)$ и $\Phi(\psi, y)$.

(TODO расписать здесь идею CCM)

Мы хотим проверить, можно ли вместо таких погружений рассматривать систему *psi*, которая будет развиваться в зависимости от наблюдений за φ . Более формально:

Есть динамическая система φ , которая развивается в многообразии M размерности m . $y : M \rightarrow \mathbb{R}$ - непрерывный наблюдатель за системой φ . Введем (динамическую систему ?) $\psi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$. Её вектора обозначим γ так, что

$$\begin{aligned}\gamma_0 &= f(y_0) \\ \gamma_{k+1} &= \psi(\gamma_k, y_k)\end{aligned}\tag{3}$$

16. эта часть текста непонятная. Если это черновик - ок. А так - нужно определиться, либо вы эти модели описываете (тогда нужен более-менее подробный текст), либо ссылаетесь в эксперименте

17. служит или является

18. Предлагаю рабочее название секции "Анализ свойств предложенного метода"

19. Хочется -> Требуется

20. Предложение обычно не начинают с переменной. Нужно перефразировать

(Пока это поток мыслей)

Если динамическая система φ задана с непрерывным временем, то вместо функции ψ рассмотрим интеграл

$$\begin{aligned}\gamma(t_0) &= f(y(t_0)) \\ \gamma(t) &= \int_{t_0}^t \psi(\gamma(\tau), y(\tau)) d\tau\end{aligned}\quad (4)$$

Ввиду невозможности непрерывного измерения y при интегрировании будем использовать интерполяцию или $y(\tau) = g(\gamma(t))$ и уточнять y по мере интегрирования.

5 Предлагаемый метод

пока просто накидал идей

Для выявления зависимости рядов мы предлагаем использовать скрытые состояния нашей архитектуры как эмбединги в CCM.

Предлагаемая архитектура является доработанным ODE-RNN. ODE-RNN обеспечивает непрерывную эволюцию скрытого состояния, но во время применения RNN эта непрерывность нарушается. С целью исправления этого недостатка в интегратор? подается сам RNN модуль, но тогда возникает проблема - неоткуда брать вход в RNN модуль. Эту проблему предлагается решить подачей на вход RNN модулю его же предсказания. Таким образом во время интегрирования модуль будет учиться порождать последовательность на которой обучается.

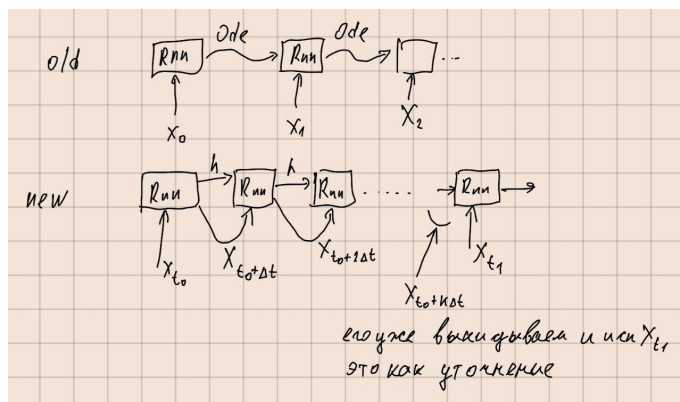


Рис. 1 A simple caption

Для выяснения зависимости рядов delay embedding ам сопоставляется скрытое состояние после пропускания его через сетку. После этого к полученным погружениям применяется CCM.

6 эксперименты

Сначала рассмотрим работу метода на паре одномерных временных рядов. В качестве модели метода используется ODE-RNN[6]. Рассмотрим получаемую корреляцию на x координатах гироскопа и акселерометра. Также рядом приведем результат применения методов к случайным данным.

Для сравнения был взят метод CCM.

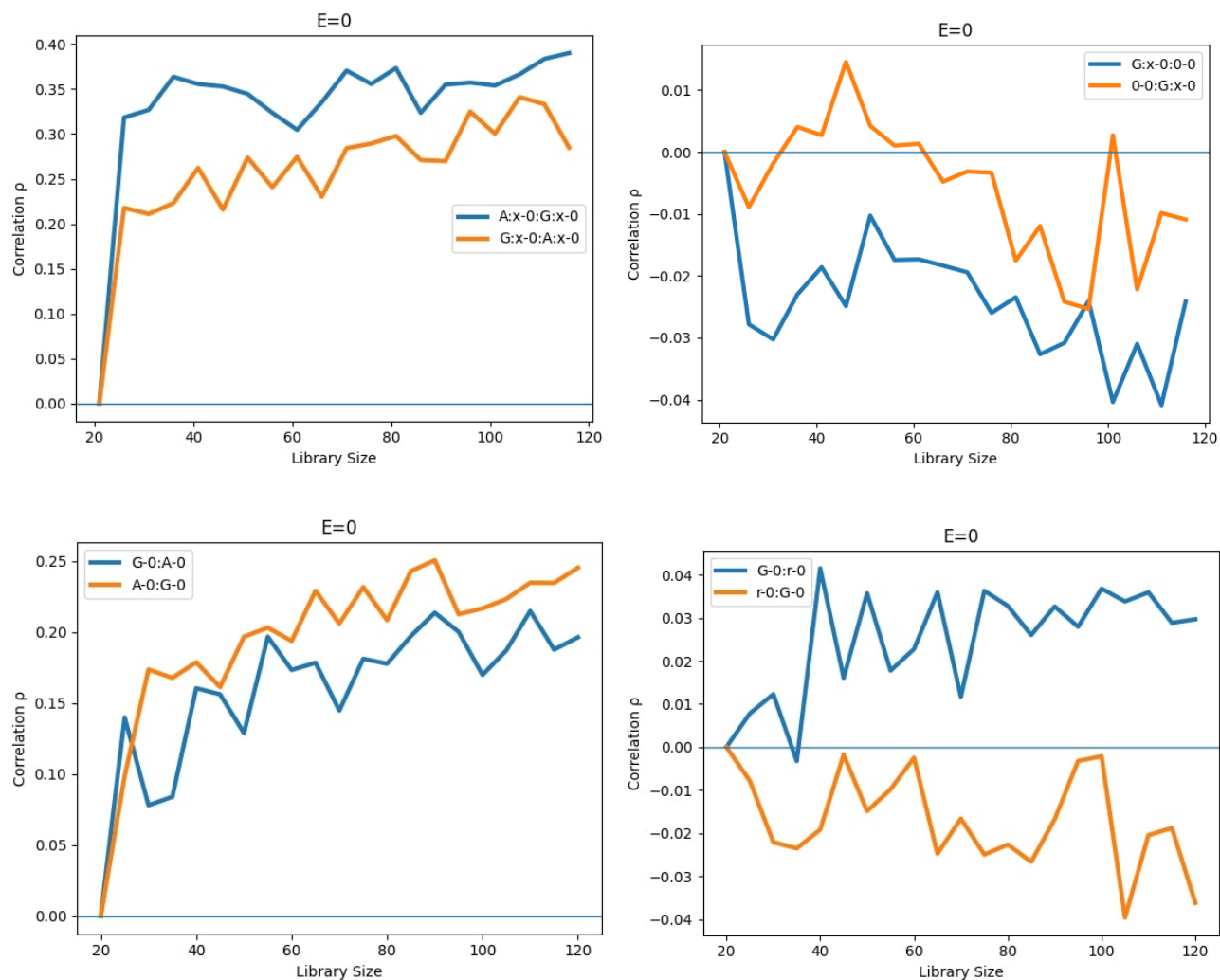


Рис. 2 слева направо, сверху вниз: CCM акселерометр - гироскоп, CCM акселерометр - random, ODE-RNN акселерометр - гироскоп, ODE-RNN акселерометр - random

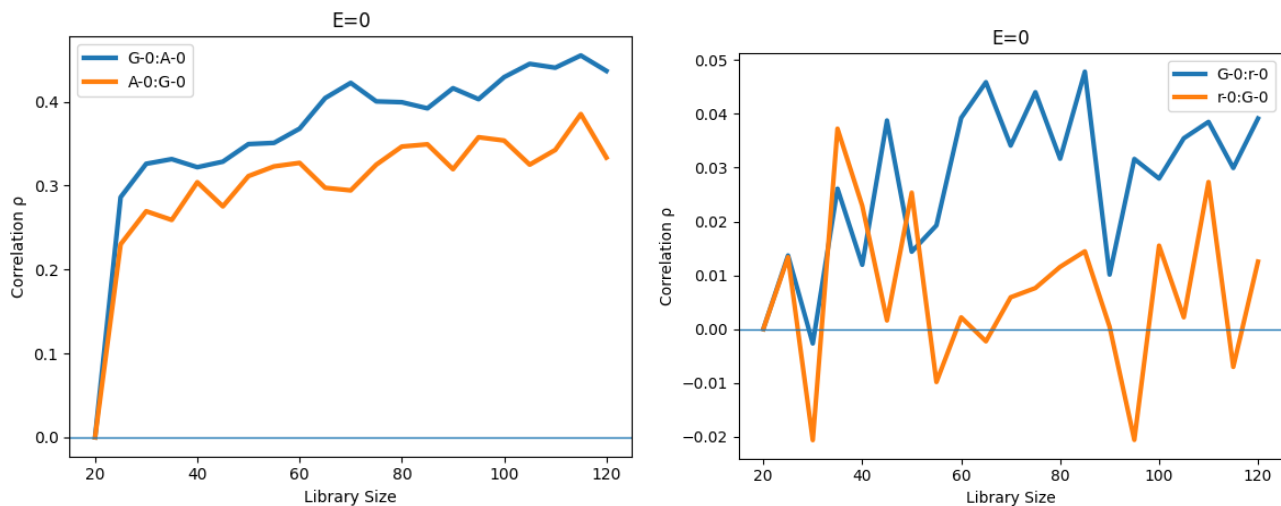


Рис. 3 слева направо: ODE-RNN акселерометр - гироскоп, ODE-RNN акселерометр - random

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

Но это не единственная возможность метода. Как отмечалось ранее - наблюдения на датасете с акселерометром - трехмерные. Но предлагаемый метод отображает многомерное наблюдение в вектор погружений, после чего можно применять ССМ с полученным погружением.

Как можно видеть, модель справляется с трехмерными данными. К тому же сохраняется сходимость ССМ

7 Заключение

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>.
- [2] Clive W.J. Granger. Granger causality.
- [3] 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*.
- [4] machinelearning.ru.
- [5] Jesse Bettencourt David Duvenaud Ricky T. Q. Chen, Yulia Rubanova. Neural ordinary differential equations. <https://dblp.uni-trier.de/>.
- [6] Yulia Rubanova, Ricky T. Q. Chen, and David Duvenaud. Latent odes for irregularly-sampled time series, 2019.

-
- [7] George Sugihara. Detecting causality in complex ecosystems. <https://www.science.org/doi/pdf/10.1126/science.1227079>.
- [8] Floris Takens. Detecting strange attractors in turbulence. *Springer Lecture Notes in Mathematics vol 898*, pp 366–81.