Восстановление траектории движения руки по видео

Владимиров Эдуард
vladimirov.ea@phystech.edu

Исаченко Роман
isa-ro@yandex.ru

Курдюкова Антонина
kurdiukova.ad@phystech.edu

12 марта 2022 г.

Аннотация

Решается задача прогнозирования временного ряда со сложной структурой. Под сложной структурой понимается наличие зависимостей и варьирующийся период. Требуется найти причинно-следственные связи между рядами и снизить размерность траекторных пространств. В работе показано, что методы канонического корреляционного анализа, такие как метод главных компонент, метод частичных наименьших квадратов и другие, являются частным случаем метода перекрестных отображений Сугихары. Для демонстрации результатов работы используется траектория движения руки, восстановленная по видео, и сигнал акселерометра.

Ключевые слова: временной ряд · фазовая траектория · траекторное подпространство · сходящееся перекрёстное отображение · частичные наименьшие квадраты · канонический корреляционный анализ

1 Введение

В данной работе решается задача прогнозирования временного ряда на основе других рядов. Одна из трудностей задачи заключается в обнаружении связи между рядами и исключении несвязанных рядов из прогностической модели. Решение этой проблемы повышает качество прогноза.

В данной работе применяется метод сходящегося перекрёстного отображения (convergent cross mapping, CCM) [1, 2], который эффективен для рядов, порождённых динамической системой. Он основан на сравнении ближайших соседей в траекторном пространстве ряда $\mathbf{x}$, полученных с помощью ряда $\mathbf{y}$.

При построении прогностической модели используется траекторная матрица (или матрица сдвига), описывающая фазовое пространство временного ряда. Например, в методе анализа спектральных компонент (singular spectrum analysis, SSA) [3, 4, 5] прогноз временного ряда основан на спектральном разложении ковариационной матрицы, полученной по траекторной. В CCM матрицы сдвига используются для проверки наличия липшицева отображения между траекторными пространствами.

Однако размерность траекторного пространства может оказаться чрезмерно высокой, что приводит к неустойчивости прогностической модели. В таком случае необходимо снизить размерность траекторного пространства путём построения проекции фазовой траектории в некоторое подпространство. Для CCM нет конкретного способа выбрать подпространство, в котором аппроксимируется фазовая траектория. В работе [6] эта проблема решается с помощью сферической регрессии. Согласно этому методу, информация об искомом подпространстве извлекается из множества эмпирических направлений $\{\mathbf{x}_i - \mathbf{x}_j \,|\, i < j\}$. В работе [7] используется автоматический выбор пары главных компонент. Идея заключается в сравнении спектральных плотностей главных компонент. Также используется простой перебор по главным компонентам [8].

Метод проекции в латентное пространство (partial least squares, PLS) [9, 10] отбирает наиболее значимые признаки и строит новые как их линейные комбинации. Это позволяет получить простую, точную и устойчивую прогностическую модель. Наряду с PLS используется метод канонического анализа корреляции (CCA) [11]. Он похож на PLS за исключением того, что первый метод максимизирует ковариацию, а последний — корреляцию. Недостатком этих моделей является их низкая точность при оценивании нелинейных зависимостей между данными. Разработаны нелинейные модели PLS[12, 13] и CCA[14, 15]. В данной статье используется модель NNPLS [16], которая преобразует исходные данные с помощью нейронной сети.

В теоретической части работы показано, что методы понижения размерности CCA и PLS являются частными случаями CCM. Для этой цели вводятся различные меры близости между целевой переменной и её аппроксимацией, являющейся линейной комбинацией её ближайших соседей, для сравнения качества согласованности исходного пространства и его подпространства.

В качестве модели для предсказания временного ряда по набору рядов используется алгоритм локально взвешенного глобального линейного отображения (sequential locally weighted global linear map, SMap) [17].

Эксперимент проводится на наборе вручную собранных данных. Он представляет собой совокупность ключевых точек, полученных по видео движения человека, а также показания акселерометра и гироскопа, снятые с руки человека. В эксперименте строится прогноз рядов, использующий обнаруженные связанные компоненты рядов.

2 Постановка задачи

Пусть значения исходного временного ряда $\mathbf{x}(t)$ доступны в моменты времени $t = 1, 2, \ldots, n$. Предполагается, что на значения $\mathbf{x}(t)$ оказывает влияние набор внешних факторов $\mathbf{y}_1(t), \ldots, \mathbf{y}_m(t)$

В момент прогноза n необходимо определить будущие значения исходного ряда $\mathbf{x}(t)$ в моменты времени $n+1, \ldots, n+p$, учитывая влияние внешних факторов $\mathbf{y}_1(t), \ldots, \mathbf{y}_m(t)$. При этом считаем, что значения внешних факторов являются доступными в моменты

времени:
$$\mathbf{y}_1(n+1),\dots,\mathbf{y}_1(n+p),\dots,\mathbf{y}_m(n+1),\dots,\mathbf{y}_m(n+p)$$
.

Для вычисления будущих значений временного ряда требуется определить функциональную зависимость, отражающую связь между прошлыми значениями $\mathbf{x}$ и будущими, а также принимающую во внимание влияние внешних факторов $\mathbf{y}_1,\dots,\mathbf{y}_m$:

$$\mathbf{x}(t) = F(\mathbf{x}(t-1),\dots,\mathbf{y}_1(t),\, \mathbf{y}_1(t-1),\dots,\mathbf{y}_m(t),\, \mathbf{y}_m(t-1),...) + \varepsilon_t. \tag{1}$$

Зависимость (1) называется моделью прогнозирования с учётом внешних факторов. Требуется создать такую модель, для которой среднее квадратичное отклонение истинного значения от прогнозируемого стремится к минимальному для заданного p.

$$\hat{E} = \frac{1}{p}\sum_{i=n+1}^{n+p}\varepsilon_i^2 \to \min_F. \tag{2}$$

2.1 Метод CCM

Зададим траекторную матрицу временного ряда $\mathbf{x} = [x_1,\dots,x_n]$ следующим образом:

$$\mathbf{H}_\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_\tau \\ x_2 & x_3 & \dots & x_{\tau+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_N & x_{N+1} & \dots & x_n \end{bmatrix},$$

где N — число задержек, $\tau = n - N + 1$.

Обозначим i-ый столбец матрицы $\mathbf{H}_\mathbf{x}$ как $\mathbf{x}_i$. Матрица $\mathbf{H}_\mathbf{x}$ имеет вид:

$$\mathbf{H}_\mathbf{x} = [\mathbf{x}_1,\dots,\mathbf{x}_\tau], \qquad \mathbf{x}_i = [x_i, x_{i+1},\dots,x_{i+N-1}]^\mathsf{T}$$

Заметим, что все векторы $\mathbf{x}_t$ принадлежат N-мерному траекторному пространству $\mathbb{H}_\mathbf{x} \subseteq \mathbb{R}^N$ ряда $\mathbf{x}$ и образуют фазовую траекторию $\mathbf{x}(t) \in \mathbb{R}^N$.

Обнаружение зависимости между рядами $\mathbf{x}$ и $\mathbf{z}$ осуществляется следующим образом. Возьмём элемент $\mathbf{x}_0$ из траекторного пространства $\mathbb{H}_\mathbf{x}$ и найдём k ближайших соседей в этом же пространстве. Обозначим их временные индексы (от ближнего к дальнему) как $t_1,\dots,t_k$.

Так как оба ряда определены на одной временной оси, то по значению ряда $\mathbf{x}$ в момент времени $t_0 \in \{1,\dots,n\}$ можно однозначно получить значение ряда $\mathbf{z}$ в тот же момент времени, и наоборот. Введём отображение из $\mathbb{H}_\mathbf{x}$ в $\mathbb{H}_\mathbf{z}$ следующим образом:

$$\varphi : \mathbf{x}_0 \mapsto \widehat{\mathbf{z}_0} = \sum_{i=1}^{k} w_i \mathbf{z}_{t_i}, \qquad w_i = \frac{u_i}{\sum\limits_{j=1}^{k} u_j}, \qquad u_i = \exp(-||\mathbf{x}_0 - \mathbf{x}_{t_i}||)$$

.

Утверждается, что ряды $\mathbf{x}$ и $\mathbf{z}$ связаны, если отображение φ является липшицевым.

$$\rho_{\mathbb{H}_\mathbf{z}}(\varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j)) \leqslant C\rho_{\mathbb{H}_\mathbf{x}}(\mathbf{x}_i, \mathbf{x}_j) \qquad \mathbf{x}_i, \mathbf{x}_j \in \mathbb{H}_\mathbf{x}.$$

Проверим наличие этого отображения следующим образом. Введём меру близости векторов в окрестностях $U_k(\mathbf{x}_{t_0})$ и $U_k(\mathbf{z}_{t_0})$:

$$L(\mathbf{x}, \mathbf{z}) = \frac{R(U_k(\mathbf{x}_{t_0}))}{R(U_k(\mathbf{z}_{t_0}))}, \qquad R(U_k(\mathbf{x}_{t_0})) = \frac{1}{k}\sum_{i=1}^{k} \rho_{\mathbb{H}_\mathbf{x}}(\mathbf{x}_{t_0}, \mathbf{x}_{t_j}) \tag{3}$$

Если $L(x, z)$ больше некоторого порога $C(n)$, то ряд $\mathbf{z}$ зависит от ряда $\mathbf{x}$

2.2 Метод PLS

Метод частичных наименьших квадратов восстанавливает связь между наборами данных $\mathbf{X}$ и $\mathbf{Y}$. Матрицы объектов $\mathbf{X}$ и целевая матрица $\mathbf{Y}$ проецируются на латентное пространство $\mathbb{R}^l$ меньшей размерности следующим образом:

$$\underset{n\times m}{\mathbf{X}} = \underset{n\times l}{\mathbf{T}} \cdot \underset{l\times m}{\mathbf{P}^\mathsf{T}} + \underset{n\times m}{\mathbf{E}}$$

$$\underset{n\times k}{\mathbf{Y}} = \underset{n\times l}{\mathbf{U}} \cdot \underset{l\times k}{\mathbf{Q}^\mathsf{T}} + \underset{n\times k}{\mathbf{F}},$$

где $\mathbf{T}$ и $\mathbf{U}$ — матрицы описания объектов и исходов в латентном пространстве, $\mathbf{P}$ и $\mathbf{Q}$ — матрицы перехода из латентного пространства в исходное, $\mathbf{E}$ и $\mathbf{F}$ — матрицы остатков.

Функция преобразования исходных данных имеет вид:

$$f(\mathbf{X}) = \mathbf{X}\mathbf{W}_\mathbf{x} \qquad g(\mathbf{Y}) = \mathbf{Y}\mathbf{W}_\mathbf{y},$$

где матрицы весов $\mathbf{W}_\mathbf{x} \in \mathbb{R}^{m\times l}$ и $\mathbf{W}_\mathbf{y} \in \mathbb{R}^{k\times l}$ находятся путём максимизации выборочной ковариации:

$$(\mathbf{W}_\mathbf{x}, \mathbf{W}_\mathbf{y}) = \underset{\mathbf{W}_\mathbf{x}, \mathbf{W}_\mathbf{y}}{argmax}\, \mathrm{Cov}(\mathbf{X}\mathbf{W}_\mathbf{x}, \mathbf{Y}\mathbf{W}_\mathbf{y})$$

2.3 Метод CCA

Канонический корреляционный анализ находит две матрицы перехода в латентные пространства для $\mathbf{X}$ и $\mathbf{Y}$ соответственно, так чтобы коэффициент корреляции между проекциями был максимальным.

$$(\mathbf{w}_{\mathbf{x}_1}, \mathbf{w}_{\mathbf{y}_1}) = \underset{\mathbf{w}_\mathbf{x}, \mathbf{w}_\mathbf{y}}{argmax}\, \mathrm{Corr}(\mathbf{X}\mathbf{w}_\mathbf{x}, \mathbf{Y}\mathbf{w}_\mathbf{y}) \tag{4}$$

Первые столбцы матриц весов находятся путём решения данной задачи оптимизации (4). Затем ищутся векторы, максимизирующие ту же корреляцию, но с ограничением, что они не коррелируют с первой парой векторов. Процедура продолжается l шагов, где l — размерность латентного пространства.

2.4 Алгоритм предсказания

Пусть дана обучающая выборка

$$\mathfrak{D} = \{(\mathbf{x}_i, y_i) \mid i = 1, \dots, d\} = (\mathbf{X}, \mathbf{y}),$$

где $\mathbf{x}_i$ — элемент траекторного пространства, а y_i — будущее значение временного ряда. В одномерном случае:

$$\mathbf{x}_i = [x_i, x_{i+1}, \dots, x_{i+N-1}]^\mathsf{T}, \quad y_i = x_{i+N}.$$

Для многомерного ряда его предсказанием служит k-ая компонента, а в качестве $\mathbf{x}_i$ выступает его значение в момент времени i. Поэтому алгоритм придётся запускать D раз, где D — размерность ряда.

Пусть алгоритм предсказывает значение ряда в момент времени t_0. Вначале каждому $\mathbf{x}_i \in X \ \{x_{t_0}\}$ поставим в соответствие:

$$\mathbf{w}_i = \exp\left(-\frac{\theta \cdot ||\mathbf{x}_i - \mathbf{x}_{t_0}||}{\frac{1}{d-1}\sum\limits_{j=1, j\neq t_0}^{d} ||\mathbf{x}_j - \mathbf{x}_{t_0}||}\right),$$

где θ — коэффициент локализации. Таким образом, близкие к $\mathbf{x}_{t_0}$ соседи имеют больший вес, чем дальние. Затем элементы $\mathbf{x}_i$ и y_i умножаются на вычисленные веса $\mathbf{w}_i$

Для прогнозирования ряда используется авторегрессионная модель порядка $t_0 - 1$:

$$X_{t_0} = \mu + \psi_1 X_{t_0-1} + \ldots + \psi_{t_0-1} X_1 + u_{t_0}, \qquad u_t \sim WN(0, \sigma^2), \quad \psi_p \neq 0.$$

Обозначим полученный прогноз $\widehat{X_{t_0}}$. Оптимальность прогноза будем понимать в смысле среднего квадратичного отклонения:

$$\min_{\psi_1, \dots, \psi_{t_0-1}} E(\widehat{X_{t_0}} - y_0)^2$$

3 Вычислительный эксперимент

Целью эксперимента является сравнение различных стратегий уменьшения размерности целевого пространства. Важной его частью является изучение результатов алгоритма прогнозирования временного ряда, применённого к элементам пространства фазовых траекторий и траекторного подпространства меньшей размерности.

3.1 Описание данных и работы модели

Первоначальные данные представляют собой набор видеороликов, на которых выполняются различные движения руками (циклические и хаотические), а также показания акселерометра и гироскопа частотой в 100 Герц, закреплённых на одной из рук. Далее по видеоряду с помощью фреймворка alphapose [links] получаются координаты конечностей, а именно 68 ключевых точек. Затем полученные многомерные ряды приводятся к одной временной шкале с помощью дублирования значений менее длинного ряда.

Перед началом эксперимента зафиксируем следующие переменные: N — размерность траекторного пространства, k — число ближайших соседей, рассматриваемых в CCM, n_{tr} — размер обучающей выборки, E_{vid} — число признаков, полученных из видеоряда, которые будут использованы в алгоритме. Для каждой пары компонент, взятых из разных временных рядов, применим метод CCM. В результате получим матрицу корреляций, в которой на i-ой строке и j-ом столбце стоит коэффициент корреляции Пирсона между i-ой компонентной "приборного"временного ряда и j-ой компонентой ряда, восстановленного по видео. После этого для каждого целевого признака выбираем E_{vid} видео-признаков, обладающих максимальной корреляцией в соответствующей строке. Затем на основе этих признаков обучается алгоритм.

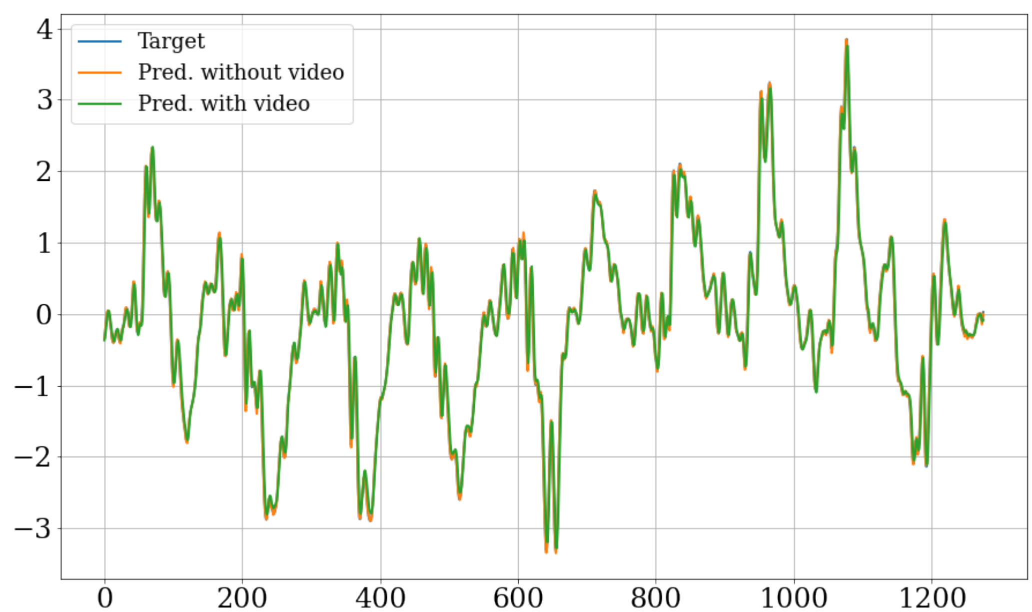

Рис. 1: Синим цветом изображены предсказываемые значения целевой переменной, оранжевым — предсказания, основанные на предыдущих значениях искомой величины, зелёным — прогнозы, использующие данные из видео-ряда

График 1 показывает, что использование дополнительных данных, извлечённых из видео, может быть избыточным. На нём видно, как целевая переменная безукоризненно прогнозируется без использования информации из видео-ряда. Связано это с тем, что описываемая ею зависимость имеет циклический характер, поскольку полученные данные соответствуют циклическому движению руки.

3.2 Сравнение методов снижения размерности

TODO

Список литературы

[1] G. Sugihara, B. Grenfell, and R. M. May. Distinguishing error from chaos in ecological time series. Phil. Trans. Roy. Soc. London B, 330(1257):235–51, 1990.

[2] George Sugihara and Robert M May. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature, 344(6268):734–741, 1990.

[3] Nina Golyandina and D Stepanov. Ssa-based approaches to analysis and forecast of multidimensional time series. In proceedings of the 5th St. Petersburg workshop on simulation, volume 293, page 298. St. Petersburg State University St. Petersburg, Russia, 2005.

[4] Nina Golyandina, Vladimir Nekrutkin, and Anatoly A Zhigljavsky. Analysis of time series structure: SSA and related techniques. CRC press, 2001.

[5] Anatoly Zhigljavsky. Singular spectrum analysis for time series: Introduction to this special issue. Statistics and its Interface, 3(3):255–258, 2010.

[6] Карина Равилевна Усманова, ЮИ Журавлёв, КВ Рудаков, and ВВ Стрижов. Аппроксимация фазовой траектории квазипериодических сигналов методом сферической регрессии. Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика, (4):40–46, 2020.

[7] Th Alexandrov and N Golyandina. Automatic extraction and forecast of time series cyclic components within the framework of ssa. In Proceedings of the 5th St. Petersburg Workshop on Simulation, pages 45–50. St. Petersburg State University St. Petersburg, 2005.

[8] Карина Равилевна Усманова and Вадим Викторович Стрижов. Модели обнаружения зависимостей во временных рядах в задачах построения прогностических моделей. Системы и средства информатики, 29(2):12–30, 2019.

[9] Roman Rosipal. Nonlinear partial least squares an overview. Chemoinformatics and advanced machine learning perspectives: complex computational methods and collaborative techniques, pages 169–189, 2011.

[10] Roman Rosipal and Nicole Kramer. Overview and recent advances in partial least squares. In International Statistical and Optimization Perspectives Workshop Subspace, Latent Structure and Feature Selection, pages 34–51. Springer, 2005.

[11] David R Hardoon, Sandor Szedmak, and John Shawe-Taylor. Canonical correlation analysis: An overview with application to learning methods. Neural computation, 16 (12):2639–2664, 2004.

[12] S Joe Qin and Thomas J McAvoy. Nonlinear pls modeling using neural networks. Computers & Chemical Engineering, 16(4):379–391, 1992.

[13] Hugo Hiden, Ben McKay, Mark Willis, and M Tham. Non-linear partial least squares using genetic programming. Genetic programming, pages 128–133, 1998.

[14] Pei Ling Lai and Colin Fyfe. Kernel and nonlinear canonical correlation analysis. International Journal of Neural Systems, 10(05):365–377, 2000.

[15] Galen Andrew, Raman Arora, Jeff Bilmes, and Karen Livescu. Deep canonical correlation analysis. In International conference on machine learning, pages 1247–1255. PMLR, 2013.

[16] Elif Bulut and Erol Egrioglu. A new partial least square method based on elman neural network. American Journal of Intelligent Systems, 4(4):154–158, 2014.

[17] George Sugihara. Nonlinear forecasting for the classification of natural time series. Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences, 348(1688):477–495, 1994.