Аппроксимация фазовой траектории квазипериодических сигналов методом сферической регрессии

Карина Усманова

Московский физико-технический институт
Физтех-школа прикладной математики и информатики
Факультет управления и прикладной математики
Кафедра «Интеллектуальные системы»

Научный руководитель: д.ф.-м.н. В.В.Стрижов

Москва 2020

Снижение размерности траекторного пространства

Задача

Решается задача поиска связей между временными рядами.

Проблема

Размерность траекторного пространства временного ряда может быть избыточна. Это усложняет описание ряда и приводит к неустойчивости прогностических моделей.

Требуется

Понизить размерность траекторного пространства временного ряда. В полученном пространстве меньшей размерности построить аппроксимацию исходного временного ряда.

Предлагается

Использовать метод сферической регрессии для снижения размерности траекторного пространства.

Литература

- Li B., Wang S. **On directional regression for dimension reduction** // Taylor & Francis. 2007.
- Li B., Zha H., Chiaromonte F. **Contour regression: a general approach to dimension reduction** // The Annals of Statistics. 2005.
- Katrutsa A. and Strijov V. **Stress test procedure for feature selection algorithms** // Chemometricsand Intelligent Laboratory Systems. 2015.
- Li J., Cheng K., Wang S., Morstatter F., Trevino R. P., Tang J., Liu H. **Feature selection: A dataperspective** // ACM Computing Surveys. 2018.

Траекторное пространство

Траекторная матрица ряда $\mathbf{s} = [s_1, s_2, \dots, s_N]$:

$$\mathbf{H}_\mathbf{s} = \begin{bmatrix} s_1 & s_2 & \dots & s_{p-1} & s_p \\ s_2 & s_3 & \dots & s_p & s_{p+1} \\ \vdots & \vdots & & \vdots & \vdots \\ s_{N-p+1} & s_{N-p+2} & \dots & s_{N-1} & s_N \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_m \end{bmatrix}, \quad m = N-p+1,$$

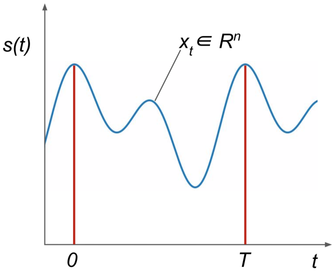
где p – ширина окна, $\mathbf{s}_i \in \mathbb{H}_s \subseteq \mathbb{R}^p$.

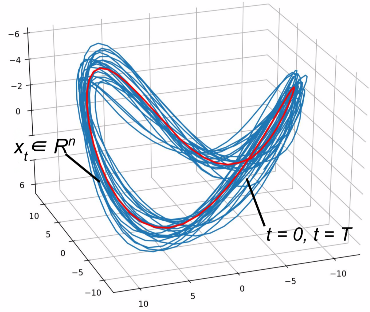
Фазовая траектория

Последовательность $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\}$ образует фазовую траекторию

$$\mathbf{x}_p(t) \in \mathbb{R}^p.$$

На рисунке представлен временной ряд и проекция его фазовой траектории в трехмерное пространство. $\mathbf{x}_t = \mathbf{x}(t)$ – точка на фазовой траектории в момент времени t.





Сферическая регрессия

- Задана выборка $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\}$ из распределения $(\mathbf{X}, \mathbf{y}), \mathbf{X} \in \mathbb{R}^{p \times n}, \mathbf{y} \in \mathbb{R}^n$.
- Метод сферической регрессии строит проекцию $\mathbf{X}$ в некоторое пространство $\mathbb{S} \subseteq \mathbb{R}^{q \times n}, q < p$.
- Информация, необходимая для построения пространства $\mathbb{S}$, извлекается из множества эмпирических направлений

$$Q = \{\mathbf{x}_i - \mathbf{x}_j \,|\, i < j\}$$

  и их квадратичных моментов.
- Интуитивно, направления $(\mathbf{x}_i - \mathbf{x}_j)$ из Q, принадлежащие $\mathbb{S}$ должны сильнее зависеть от $\mathbf{y}$, чем направления из $\mathbb{S}^{\perp}$.

Собственное подпространство в сферических координатах

- Построим отображение траектории $\mathbf{x}_p(t)$ из декартовых координат в сферические:

$$\varphi \; : \mathbf{x}_p(t) \; \to \; \mathbf{z}_p(t) = [\alpha_1(t), \alpha_2(t), \dots \alpha_{p-1}(t), r(t)]^{\mathsf{T}}.$$

- Сферическая регрессия строит отображение:

$$g \; : \; \underbrace{[\alpha_1, \alpha_2, \dots \alpha_{p-1}]^{\mathsf{T}}}_{\boldsymbol{\alpha}} \; \to \; \underbrace{[\beta_1, \beta_2, \dots \beta_{q-1}]^{\mathsf{T}}}_{\boldsymbol{\beta}}, \quad q \leq p,$$

такое, что переменная r может быть восстановлена по углам $[\beta_1, \dots \beta_{q-1}]^{\mathsf{T}}$:

$$f \; : \; \boldsymbol{\beta} \; \mapsto \; \hat{r}.$$

Восстановление траектории по собственному подпространству

$$f \ : \ \boldsymbol{\beta} \ \mapsto \ \hat{r}$$

$$\hat{r} = f(\hat{\mathbf{w}}, \beta_1, \beta_2, \dots \beta_{q-1}), \quad \hat{\mathbf{w}} = \arg\min_{\mathbf{w}} (r - \mathbf{w}^{\mathsf{T}}\boldsymbol{\beta})^2.$$

Восстановленная траектория:

$$\hat{\mathbf{x}}_p(t) = \varphi^{-1}\Big(\hat{\mathbf{z}}_p(t)\Big) = \varphi^{-1}\Big([\boldsymbol{\alpha}, \hat{r}(t)]^{\mathsf{T}}\Big). \qquad (1)$$

Ошибка аппроксимации:

$$\text{MSE}(\mathbf{x}_p, \hat{\mathbf{x}}_p) = \frac{1}{N}\sum_{t=1}^{N} ||\mathbf{x}_p(t) - \hat{\mathbf{x}}_p(t)||^2. \qquad (2)$$

Собственное подпространство ряда $\mathbf{s}$ определяется как подпространство траекторного пространства $\mathbb{H}_{\mathbf{s}}$, имеющее минимальную размерность среди тех, в которых модель f строит адекватную аппроксимацию (1) в смысле квадратичной ошибки (2).

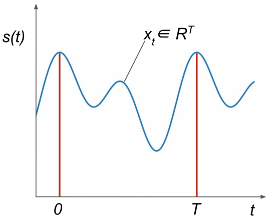
- Ожидаемое значение траектории:

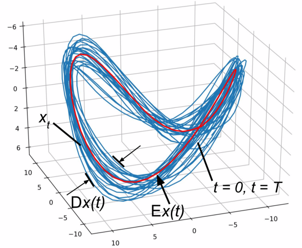
$$\mathsf{E}\mathbf{x}(t) = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}(t+kT),$$

где T – период, K – число целых периодов внутри $[1, N]$.

- Диаметр траектории

$$\mathsf{D}\mathbf{x}(t) = \frac{2}{T}\sum_{t=1}^{T}\sqrt{\frac{1}{K}\sum_{k=1}^{K}\Big(\mathbf{x}(t+kT) - \mathbf{x}^*(t)\Big)^2}.$$





Ожидаемое значение и диаметр траектории

Применение

- Компактное описание траектории
- Проверка наличия связи между временными рядами
- Анализ самопересечений траектории
- Альтернативное определение собственного подпространства

Эмпирическое определение собственного подпространства

Задана проекция $\mathbf{x}_q$ фазовой траектории $\mathbf{x}_p$ в некоторое траекторное подпространство $\mathbb{S} \subset \mathbb{R}^q$. Будем называть $\mathbb{S}$ собственным, если ожидаемое значение траектории $\mathsf{E}\mathbf{x}_q$ не имеет пересечений в пределах диаметра траектории $\mathsf{D}\mathbf{x}_q$.

Для временных рядов $\mathbf{s}_1 = [s_1^1, \ldots, s_N^1]$ и $\mathbf{s}_2 = [s_1^2, \ldots, s_N^2]$ требуется установить наличие связи между ними.

Решение

Считаем, что ряд $\mathbf{s}_2$ зависит от ряда $\mathbf{s}_1$, если существует липшицево отображение $\varphi : \mathbb{H}_{\mathbf{s}_1} \to \mathbb{H}_{\mathbf{s}_2}$:

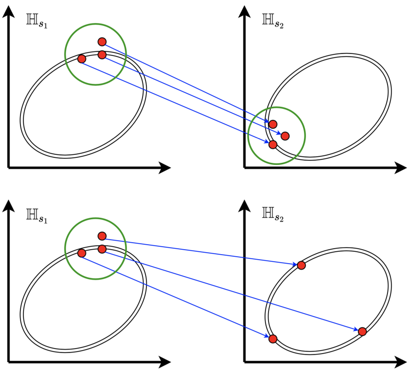
$$\rho_{\mathbb{H}_{\mathbf{s}_2}}\Big(\varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j)\Big) \leq L \cdot \rho_{\mathbb{H}_{\mathbf{s}_1}}(\mathbf{x}_i, \mathbf{x}_j), \quad \text{для } \mathbf{x}_i, \mathbf{x}_j \in \mathbb{H}_{\mathbf{s}_1}.$$

Предлагается изучить связь между рядами не только на траекторных пространствах $\mathbb{H}_{\mathbf{s}_1}$ и $\mathbb{H}_{\mathbf{s}_2}$, но и на их подпространствах. Это позволяет

- Искать связь между рядами, если размерность их траекторных пространств избыточна.
- Более детальное изучить связь между рядами.

Построение сходящегося перекрестного отображения

- Выбираем $\mathbf{x}_{t^*} \in \mathbb{H}_{\mathbf{s}_1}$.
- Пусть $\mathbf{x}_{t_1}, \dots, \mathbf{x}_{t_k}$ – k ближайших соседей вектора $\mathbf{x}_{t^*}$ в пространстве $\mathbb{H}_{\mathbf{s}_1}$.
- Рассмотрим соответствующие им строки матрицы $\mathbf{H}_{\mathbf{s}_2}$ в моменты времмени $t_1, \dots, t_k$.

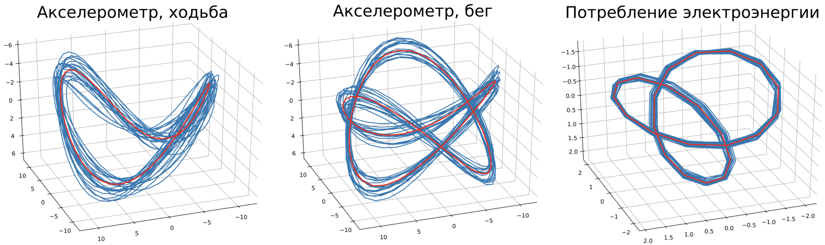


Эксперимент

Эксперимент проводился на трех временных рядах: показатели акселерометра во время ходьбы и во время бега, данные потребления электроэнергии в течение года.

Цель

- Анализ размерности собственного подпространства
- Сравнение определений собственного подпространства
- Проверка связи между рядами на их собственных подпространствах.



Переберем размерность собственного подпространства $q \in [1, 15]$. Для каждой размерности найдем ошибку аппроксимации фазовой траектории.

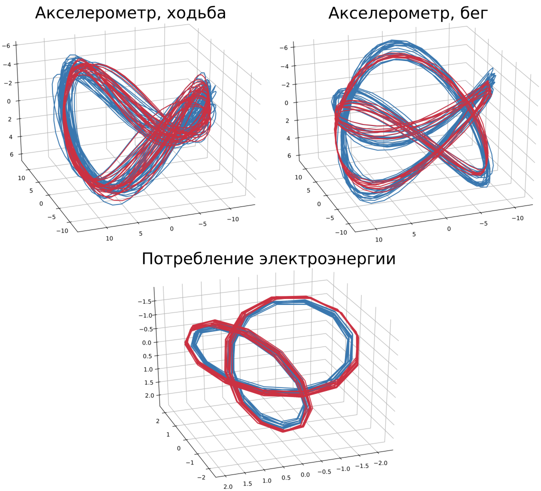
$q = 3$

$q = 6$

$q = 8$

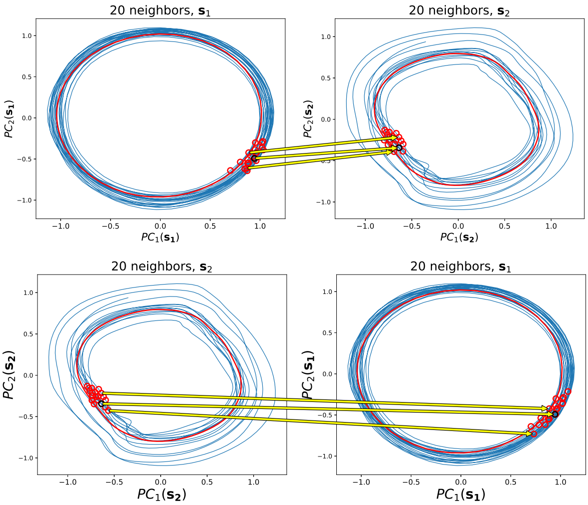
Аппроксимация фазовой траектории

Построим аппроксимацию фазовой траектории в соответствии с найденной размерностью собственного подпространства.



Обнаружения связи между рядами

Проверим наличие связи между показателями акселерометра и гироскопа на собственных подпространствах.



Результаты

- Проведен эксперимент по нахождению собственного подпространства методом сферической регрессии.
- С помощью метода CCM исследована связь между показателями акселерометра и гироскопа на собственных подпространствах.

Публикации

- К. Р. Усманова, В. В. Стрижов. Анализ зависимостей между показателями при прогнозировании объема грузоперевозок // Системы и средства информатики, 2018.
- К. Р. Усманова, В. В. Стрижов. Модели обнаружения зависимостей во временных рядах в задачах построения прогностических моделей. // Системы и средства информатики, 2019.
- К. Р. Усманова, К. В. Рудаков, В. В. Стрижов. Аппроксимация фазовой траектории квазипериодических сигналов методом сферической регрессии // Вестник Московского университета, 2020.