Модели согласования скрытого пространства в задаче корреляционного анализа

Роберт Сафиуллин

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

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

Москва, 2023

Задача декодирования временного ряда

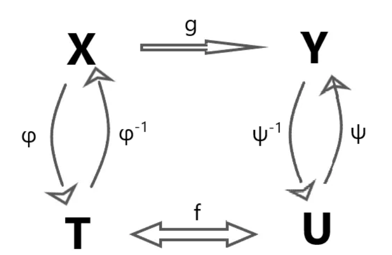
Цель

Рассматривается система, описываемая аттракторами в нескольких фазовых пространствах. Требуется исследовать зависимости объектов в различных фазовых пространствах динамической системы и построить устойчивую модель декодирования временных рядов между ними

Задача

Строятся частные модели, аппроксимирующие измерения состояния системы в каждом пространстве. Строится согласующая мультимодель. Уточняются параметры частных моделей. Для учёта зависимостей в пространствах объектов и ответов предлагается снизить размерность исходного пространства

Схема



$\mathbf{X} = \{x_i\}_{i=1}^N$- заданные временные ряды, $\mathbf{Y} = \{y_i\}_{i=1}^N$ - целевые временные ряды

$$\mathbf{g} = \psi \circ \mathbf{f} \circ \phi$$

Частные модели

$\phi : \mathbf{R}^N \to \mathbf{R}^m$

$\psi : \mathbf{R}^m \to \mathbf{R}^N$

Модель согласования

$\mathbf{f} : \mathbf{R}^m \to \mathbf{R}^m$

Многослойный перцептрон

$$\mathbf{f}(x)^{[N]} = \mathbf{a}^{[N]}(W^{[N]}(\mathbf{a}^{[N-1]}(W^{[N-1]}(\dots \mathbf{a}^{[1]}(W^{[1]}x+b^{[1]})\dots)+b^{[N-1]})+b^{[N]})$$

a = max(0, x)

Для каждого слоя N, пара $(W^{[N]}, b^{[N]})$ является параметрами этого слоя, а $\mathbf{a}^{[N]}$ - функцией активации.

Метод частных наименьших квадратов (PLS)

$$\underset{m\times n}{\mathbf{X}} = \underset{m\times l}{\mathbf{T}} \cdot \underset{l\times n}{\mathbf{P}^{\mathrm{T}}} + \underset{m\times n}{\mathbf{F}} = \sum_{k=1}^{l} \underset{m\times 1}{\mathbf{t}_k} \cdot \underset{1\times n}{\mathbf{p}_k^{\mathrm{T}}} + \underset{m\times n}{\mathbf{F}},$$

$$\underset{m\times r}{\mathbf{Y}} = \underset{m\times l}{\mathbf{U}} \cdot \underset{l\times r}{\mathbf{Q}^{\mathrm{T}}} + \underset{m\times r}{\mathbf{E}} = \sum_{k=1}^{l} \underset{m\times 1}{\mathbf{u}_k} \cdot \underset{1\times r}{\mathbf{q}_k^{\mathrm{T}}} + \underset{m\times r}{\mathbf{E}}.$$

Частные модели

Кодировщик имеет функцию активации ϕ и матрицу весов $\mathbf{W} \in \mathbb{R}^{m \times n}$. Тогда сжатое представление можно выразить следующим образом:

$$\hat{\mathbf{h}} = \phi(\mathbf{W}h + b)$$

Декодировщик с функцией активации ψ и матрицей весов $\mathbf{V} \in \mathbb{R}^{n \times m}$, восстановленный вектор признаков $\hat{\mathbf{x}} \in \mathbb{R}^n$ может быть определен следующим образом:

$$\hat{\mathbf{x}} = \psi(\mathbf{V}x + c)$$

где $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ - вектор смещения.
Функцией потерь является среднеквадратичная ошибка (MSE):

$$\mathrm{MSE}(x, \hat{x}) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{x}_i)^2.$$

Пространство СПО матриц

Симметрично положительно определенные (СПО) матрицы формируют дифференцируемое многообразие **M**. Касательное пространство к этому многообразию в точке C обозначим τCM. Риманово расстояние между двумя СПО матрицами: $\delta_R(\mathbf{C_1}, \mathbf{C_2}) = \|\mathrm{logm}(\mathbf{C_1}^{-1}\mathbf{C_2})\|_F = \left(\sum_{i=1}^{n} \log^2 \lambda_i\right)^{1/2}$, где $\lambda_1, ..., \lambda_n$ - положительные собственные значения матрицы $C_1^{-1} C_2$.

Для набора ковариационных матриц фрагментов временных рядов $\{\mathrm{C}_i\}_{i=1}^{N}, \mathrm{C}_i = \frac{\mathbf{X_i X_i^T}}{size(x_i)})$ рассчитывается риманово среднее **C**
Операция проецирования матрицы $\mathbf{C_i}$ на τCM:
$\mathbf{S_i} = \mathrm{Log}_C(\mathbf{C_i}) = \mathbf{C}^{1/2}\mathrm{logm}(\mathbf{C}^{-1/2}\mathbf{C_i}\mathbf{C}^{-1/2})\mathbf{C}^{1/2}$.
$\mathbf{C} = \arg\min_{\mathbf{C}} \sum_i \delta_R(\mathbf{C}, \mathbf{C}_i)^2$

Новые признаки

$vec(\mathbf{S}) = [\mathbf{S}_{1,1}; \sqrt{\mathbf{S}}_{1,2}; \mathbf{S}_{2,2}; \sqrt{2}\mathbf{S}_{1,3}; \sqrt{2}\mathbf{S}_{1,3}; \mathbf{S}_{3,3}]$

QPFS

Метод выбора признаков, основанный на поиске вектора значимостей признаков $\mathbf{z}$

$$\underbrace{\mathbf{z}^T\mathbf{Q}\mathbf{z}}_{\text{Sim}} - \alpha \cdot \underbrace{\mathbf{b}^T\mathbf{z}}_{\text{Rel}} \to \min_{\substack{\mathbf{z}\in R^N_+ \\ \|\mathbf{z}\|_1=1}}$$

$$\mathbf{b} = \left[\sum_{k=1}^{r} |\text{corr}(\chi_i, \nu_k)|\right]_{i=1}^{N}.$$

где $\mathbf{Q} \in R^{N\times N}$ - матрица коэффициентов попарного сходства между признаками, $\mathbf{b} \in R^N$ - вектор релевантности признака к целевой переменной

Теоретическое обоснование

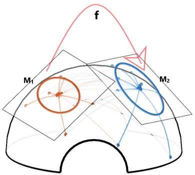
Теорема 1

Имеем система, описываемую аттракторами в двух фазовых пространствах M_1 и M_2. Пусть $\{x_i\}_{i=1}^n$ и $\{y_i\}_{i=1}^n$ обозначают измерения состояния системы в M_1 и M_2 соответственно. Пусть $\phi_\theta : \mathbb{R}^d \to M_1$ и $\psi_\omega : \mathbb{R}^d \to M_2$ являются двумя автокодировщиками, аппроксимирующими измерения в каждом пространстве. Пусть также $f_\gamma : \mathbb{R}^d \to \mathbb{R}^d$ является согласующей мультимоделью, строящей связь между скрытыми представлениями двух пространств, где γ обозначает ее параметры. Если параметры θ, ω и γ обновляются с использованием подходящего алгоритма оптимизации, то они будут сходиться к некоторым оптимальным значениям, при которых функции ϕ_θ, ψ_ω и f_γ будут обратимыми и непрерывными, и они позволят восстановить изначальные временные ряды из одного фазового пространства в другом.

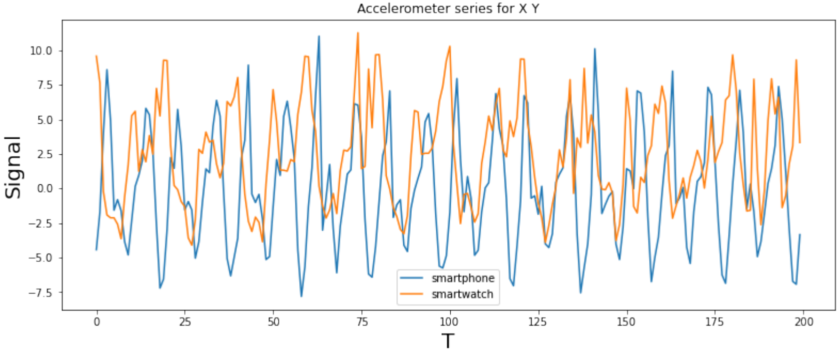
Теоретическое обоснование

Теорема 2

Пусть $\mathbf{M}_1$ и $\mathbf{M}_2$ - два римановых многообразия, представляющие фазовые пространства динамической системы, и пусть $\epsilon > 0$ - заданная точность аппроксимации. Мы предполагаем, что существует изометрическая функция перехода $\mathbf{T} : \mathbf{M}_1 \to \mathbf{M}_2$, такая что функция сохраняет риманову метрику и может быть аппроксимирована нейронной сетью $\mathbf{f}$ с точностью ϵ, достигая взаимно однозначного преобразования между фазовыми пространствами.



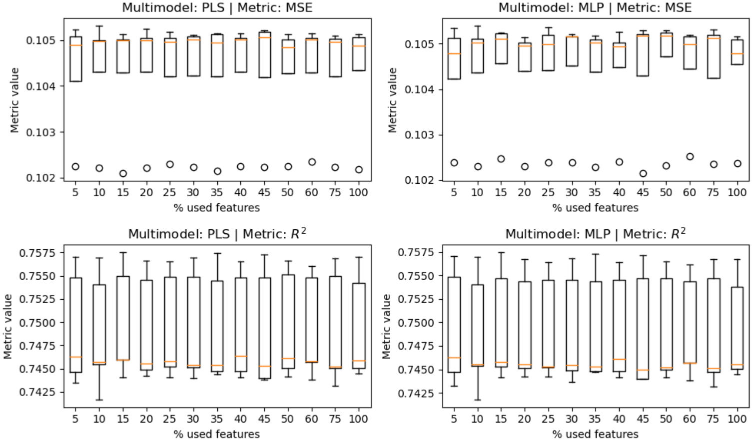
Данные: Smartphone and Smartwatch Activity and Biometrics Dataset



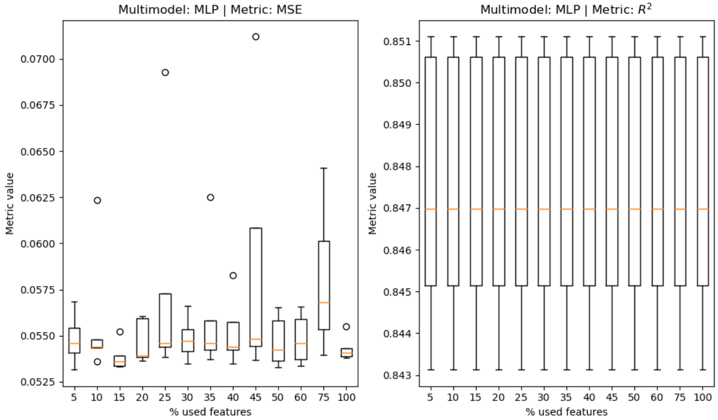
[1]

[1] Gary M. Weiss, Kenichi Yoneda, and Thaier Hayajneh. Smartphone and Smartwatch-Based Biometrics Using Activities of Daily Living. IEEE Access, 7:133190-133202, Sept. 2019.

Скрытое представление автокодировщика



Проецирование на многообразие



Итоги

- Предложен и теоретически обоснован метод восстановления временных рядов
- Выполнен численный эксперимент для проверки качества метода
- Показано, что предложенный метод слабо зависит от числа использованных признаков