Выбор прогностических моделей в римановых фазовых пространствах

Алтай Эйнуллаев Эльшан оглы

Московский физико-технический институт

Научный руководитель: д.ф-м.н. Стрижов Вадим Викторович

2024

Выбор прогностических моделей в римановых фазовых пространствах

Мотивация

Матрицы ковариации многомерных временных рядов лежат в римановом пространстве SPD матриц. Переходя в это пространство, решают[1] задачу классификации многомерных временных рядов.

Проблема

В отличие от задачи классификации, задачу прогнозирования нельзя решить в пространстве матриц ковариации, т.к. требуется прогнозировать исходные временные ряды.

Предлагаемый подход

Рассматривать векторное представление матриц ковариации как описание взаимосвязи фазовах траекторий многомерных временных рядов и использовать эту информацию для улучшения качества прогноза.

---

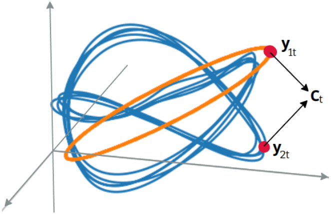
[1] Classification of covariance matrices using a Riemannian-based kernel for BCI applications, Alexandre Barachant et al., *Neurocomputing*, 2013

Матрица ковариации в фазовых пространствах

Ставится задача прогнозирования многомерного временного ряда $\mathbf{x}_t = [x_{1t}, \ldots, x_{nt}]^\mathsf{T}$. Каждому временному ряду ставится в соответствие фазовое пространство Y векторов задержек $\mathbf{y}_t = [x_{t-L+1}, \ldots, x_t]^\mathsf{T}$ размерности L. Матрица ковариации в каждый момент времени:

$$\mathbf{C}_t = \frac{1}{L-1}\mathbf{X}_t\mathbf{X}_t^\mathsf{T},$$

где $\mathbf{X}_t = [\mathbf{y}_{1t} \ldots \mathbf{y}_{nt}]^\mathsf{T}$.



Точкам фазовых траекторий $\mathbf{y}_{1t}, \mathbf{y}_{2t}$ ставится в соответствие матрица ковариации $\mathbf{C}_t$.

Пространство SPD матриц

Пусть

$$S(n) = \{\mathbf{S} \in M(n), \mathbf{S}^\top = \mathbf{S}\}$$

– пространство всех $n \times n$ симметричных матриц в пространстве квадратных вещественных матриц $M(n)$ и

$$P(n) = \{\mathbf{C} \in S(n), \mathbf{u}^\top \mathbf{C} \mathbf{u} > 0, \forall \mathbf{u} \in \mathbb{R}\}.$$

Пространство SPD матриц $P(n)$ – гладкое многообразие $\mathcal{M}$. Касательное пространство к $P(n)$ в $\mathbf{C}$ – векторное пространство $T_{\mathbf{C}} \subset S(n)$.

Метрика и среднее в пр-ве $P(n)$

Скалярное произведение в касательном пространстве:

$$\langle \mathbf{S}_1, \mathbf{S}_2 \rangle_{\mathbf{C}} = \mathrm{tr}(\mathbf{S}_1 \mathbf{C}^{-1} \mathbf{S}_2 \mathbf{C}^{-1}).$$

Кратчайшая линия $\mathbf{\Gamma}(t) : [0,1] \to P(n)$, соединяющая две точки многообразия называется геодезической линией и расстояние между ними равно ее длине $l(\mathbf{\Gamma}(t))$. Таким образом, геодезическое расстояние:

$$\delta(\mathbf{C}_1, \mathbf{C}_2) = \| \log(\mathbf{C}_1^{-1} \mathbf{C}_2) \|_F = \left(\sum_{i=1}^{n} \log^2 \lambda_i \right)^{1/2},$$

где λ_i, $i = 1, \ldots, n$ – собственные числа матрицы $\mathbf{C}_1^{-1} \mathbf{C}_2$.
Среднее SPD матриц, соответствующее δ:

$$G(\mathbf{C}_1, \ldots, \mathbf{C}_I) = \arg \min_{\mathbf{C} \in P(n)} \sum_{i=1}^{I} \delta(\mathbf{C}_i, \mathbf{C}).$$

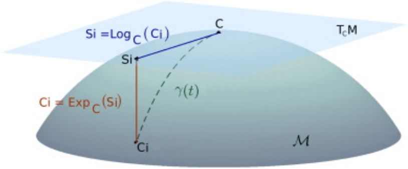
Переход в касательное пространство $T_{\mathbf{C}}$

Для точки $\mathbf{C} \in P(n)$, определим касательное пространство $T_{\mathbf{C}}$. Каждый касательный вектор $\mathbf{S}_i$ является производной в $t = 0$ геодезической линии $\mathbf{\Gamma}_i(t)$ между $\mathbf{C}$ и экспоненциальным отображением $\mathbf{C}_i = \mathrm{Exp}_{\mathbf{C}}(\mathbf{S}_i)$, определенным следующим образом:

$$\mathbf{C}_i = \mathrm{Exp}_{\mathbf{C}}(\mathbf{S}_i) = \mathbf{C}^{\frac{1}{2}} \exp(\mathbf{C}^{-\frac{1}{2}}\mathbf{S}_i\mathbf{C}^{-\frac{1}{2}})\mathbf{C}^{\frac{1}{2}}.$$

Обратное отображение:

$$\mathbf{S}_i = \mathrm{Log}_{\mathbf{C}}(\mathbf{C}_i) = \mathbf{C}^{\frac{1}{2}} \log(\mathbf{C}^{-\frac{1}{2}}\mathbf{C}_i\mathbf{C}^{-\frac{1}{2}})\mathbf{C}^{\frac{1}{2}}.$$



Касательное пространство в точке $\mathbf{C}$, $\mathbf{S}_i$ – касательный вектор в $\mathbf{C}$, $\gamma_i(t)$ – геодезическая линия, соединяющая $\mathbf{C}$, $\mathbf{C}_i$ в $P(n)$.

Построение прогностической модели

Требуется выбрать две модели f, h:

$$\mathbf{x}_t \xrightarrow{f} \hat{\mathbf{x}}_{t+1},$$

$$\mathbf{x}_t \xrightarrow{g} \mathbf{S}_t \xrightarrow{h} \hat{\mathbf{S}}_{t+1},$$

где f – модель прогнозирования временного ряда в фазовом пространстве, g – отображение, ставящее в соответствие временному ряду представление матрицы ковариации в касательном пространстве $\mathbf{S}_t$, h – модель прогнозирования $\hat{\mathbf{S}}_{t+1}$ в касательном пространстве. Для улучшение прогноза f строится мультимодель:

$$\hat{\mathbf{x}}_{t+1} = F(\hat{\mathbf{x}}_{t+1}, \hat{\mathbf{S}}_{t+1}) = F(f(\mathbf{x}_t), h(g(\mathbf{x}_t))).$$

План вычислительного эксперимента

В силу того, что $\mathbf{S}_t$ описывает взаимосвязь точек фазовых траекторий временных рядов, в качестве f рассматриваются прогностические модели в фазовых пространствах:

- Локальные модели прогнозирования,
- SSA, tSSA, mSSA.

В качестве h можно рассматривать произвольные модели прогнозирования, например:

- LSTM,
- SSA, tSSA, mSSA.

В качестве F рассматриваются:

- Линейная модель,
- Двухслойная нейронная сеть.

Разбиение временного ряда для обучения

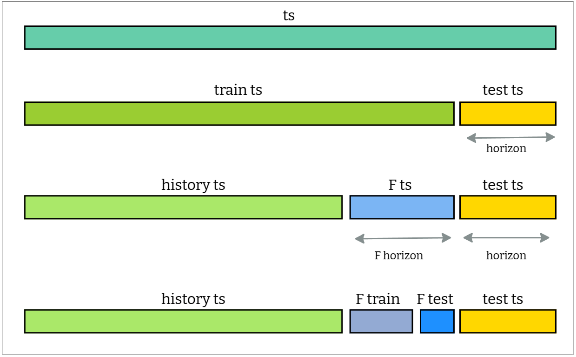
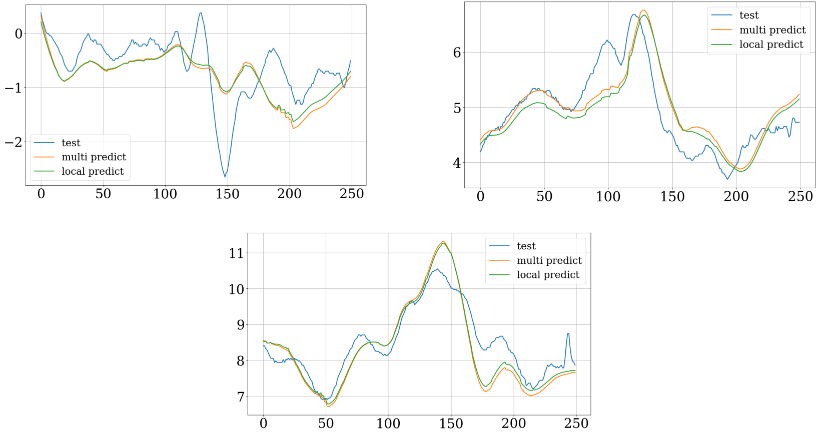


Схема разделения временного ряда для обучения модели:

- history ts – используется для обучения f, g,
- F train – используется для обучения F,
- F test – оценка качества F

Результаты

- f – локальная модель, усреднение $k = 20$ соседей
- h – I, тождественное преобразование
- F – линейная регрессия



Трехмерный ряд акселерометра. $\text{MAPE}_{\text{local}} = 0.683$, $\text{MAPE}_{\text{multi}} = 0.67$

Дальнейшие исследования

- Полностью реализовать план эксперимента
- Получить условия применимости метода
- Попробовать другие способы подсчета матрицы $\mathbf{C}_t$.