

Выбор прогностических моделей в римановых фазовых пространствах

Алтай Эйнуллаев Эльшан оглы

Московский физико-технический институт

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

2025

Выбор моделей в римановых фазовых пространствах

Матрицы ковариации набора временных рядов лежат в римановом пространстве SPD матриц. В этом пространстве решается задача классификации набора 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

Векторное представление матриц ковариации рассматривается как описание взаимосвязи фазовых траекторий набора временных рядов и используется для улучшения качества прогноза.

Матрица ковариации в фазовых пространствах

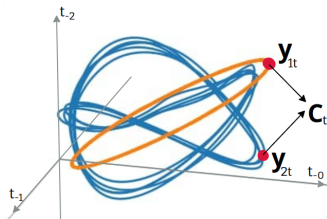
$\mathbf{x}_t = [x_{1t}, \dots, x_{nt}]^T$ – набор n временных рядов,

$\mathbf{y}_{it} = [x_{i(t-L+1)}, \dots, x_{it}]^T$ – точки фазового пространства \mathbf{y}_i векторов задержек ряда x_{it} размерности L .

Матрица ковариации в каждый момент времени:

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

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



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

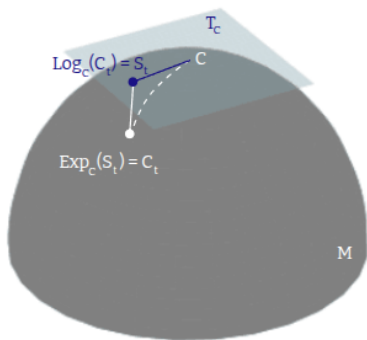
Перевод матриц в касательное пространство T_C

Перевод SPD матрицы C_t в касательное пространство T_C :

$$S_t = \text{Log}_C(C_t) = C^{\frac{1}{2}} \log(C^{-\frac{1}{2}} C_t C^{-\frac{1}{2}}) C^{\frac{1}{2}}.$$

Обратное отображение:

$$C_t = \text{Exp}_C(S_t) = C^{\frac{1}{2}} \exp(C^{-\frac{1}{2}} S_t C^{-\frac{1}{2}}) C^{\frac{1}{2}}.$$



Перевод C_t в пространство T_C и обратно

Зашумленные матрицы ковариации

Theorem (Эйнуллаев, 2025)

Пусть $\mathbf{x}_t = (x_{1t}, x_{2t})$ и $x_{it} = \tilde{x}_{it} + \varepsilon_{it}$, где $i \in \{1, 2\}, \varepsilon_{it} \sim \mathcal{N}(0, \sigma_i^2)$. Пусть $\tilde{\mathbf{C}}_t$ – матрица ковариации для набора временных рядов $\tilde{\mathbf{x}}_t = (\tilde{x}_{1t}, \tilde{x}_{2t})$. Тогда матрица ковариации \mathbf{C}_t для набора \mathbf{x}_t равна:

$$\mathbf{C}_t = \tilde{\mathbf{C}}_t \pm \frac{1}{L-1} \Delta \mathbf{C}_t,$$

где

$$\Delta \mathbf{C}_t = \begin{pmatrix} \frac{\sqrt{4\|\mathbf{y}_{1t}\|^2\sigma_1^2 + 2L\sigma_1^4}}{\sqrt{\sigma_2^2\|\mathbf{y}_{1t}\|^2 + \sigma_1^2\|\mathbf{y}_{2t}\|^2 + L\sigma_1^2\sigma_2^2}} & \frac{\sqrt{\sigma_2^2\|\mathbf{y}_{1t}\|^2 + \sigma_1^2\|\mathbf{y}_{2t}\|^2 + L\sigma_1^2\sigma_2^2}}{\sqrt{4\|\mathbf{y}_{2t}\|^2\sigma_2^2 + 2L\sigma_2^4}} \end{pmatrix}$$

– матрица среднеквадратичных отклонений соответствующих элементов матрицы \mathbf{C}_t .

Риманово расстояние между зашумленными \mathbf{C}_t

Определение метрики в пространстве SPD матриц:

$$d(\mathbf{C}_{t_1}, \mathbf{C}_{t_2}) = \|\text{Log}(\mathbf{C}_{t_1}^{-1} \mathbf{C}_{t_2})\|_F = \left[\sum_{i=1}^L \ln^2 \lambda_i \right]^{\frac{1}{2}},$$

Lemma (Эйнуллаев, 2025)

Пусть $\mathbf{C}_{t_1} = \tilde{\mathbf{C}} \pm \Delta \mathbf{C}_{t_1}$, $\mathbf{C}_{t_2} = \tilde{\mathbf{C}} \pm \Delta \mathbf{C}_{t_2}$. Тогда, при условии приближения $\mathbf{C}_{t_1}^{-1} \mathbf{C}_{t_2} \approx \mathbf{E} \pm (\tilde{\mathbf{C}}^{-1} \Delta \mathbf{C}_{t_1} - \tilde{\mathbf{C}}^{-2} \Delta \mathbf{C}_{t_1} \tilde{\mathbf{C}})$, в первом порядке теории возмущений, собственные числа матрицы $\mathbf{C}_{t_1}^{-1} \mathbf{C}_{t_2}$ равны:

$$\lambda_1 = 1 \pm (\tilde{\mathbf{C}}^{-1} \Delta \mathbf{C}_{t_1} - \tilde{\mathbf{C}}^{-2} \Delta \mathbf{C}_{t_1} \tilde{\mathbf{C}})_{11},$$

$$\lambda_2 = 1 \pm (\tilde{\mathbf{C}}^{-1} \Delta \mathbf{C}_{t_1} - \tilde{\mathbf{C}}^{-2} \Delta \mathbf{C}_{t_1} \tilde{\mathbf{C}})_{22}.$$

Расстояние, на которое отдаются матрицы $\mathbf{C}_{t_1}, \mathbf{C}_{t_2}$:

$$\delta(\mathbf{C}_{t_1}, \mathbf{C}_{t_2}) \approx \sqrt{\sum_i (\tilde{\mathbf{C}}^{-1} \Delta \mathbf{C}_{t_1} - \tilde{\mathbf{C}}^{-2} \Delta \mathbf{C}_{t_1} \tilde{\mathbf{C}})_{ii}^2}.$$

Метод улучшения базового прогноза

Требуется выбрать две модели f , F :

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

$$(\hat{\mathbf{x}}_{t+1}^f, \mathbf{S}_t) \xrightarrow{F} \hat{\mathbf{x}}_{t+1}.$$

Оптимизационная задача

$$\hat{\mathbf{w}}_f = \arg \min_{\mathbf{w} \in \mathbb{R}^k} \mathcal{L}(\mathbf{w}, \hat{\mathbf{x}}_{t+1} - \mathbf{x}_{t+1}),$$

$$\hat{\mathbf{w}}_F = \arg \min_{\mathbf{w} \in \mathbb{R}^m} \|F(\hat{\mathbf{x}}_{t+1}^f, \mathbf{S}_t | \mathbf{w}) - \mathbf{x}_{t+1}\|_2^2.$$

Критерий качества

$$\text{MSE} = \frac{1}{h} \sum_{i=T+1}^{T+h} \frac{1}{2} \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|_2^2.$$

где h – горизонт прогноза.

Разбиение временного ряда для обучения

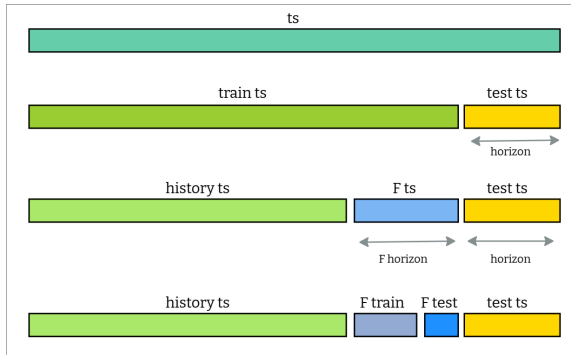


Схема разделения временного ряда для обучения модели:

history ts – используется для обучения f ,

F train – используется для обучения F ,

F test – оценка качества F ,

test ts – сравнение качества F и f .

Метод сходящегося перекрестного отображения (ССМ)

Идея

x_{1t}, x_{2t} причинно связаны если порождены одной и той же динамической системой.

Реализация

M_{x_1}, M_{x_2} – фазовые пространства временных рядов, L – длина временных рядов:

- **Перекрестное отображение** : Возможность предсказать x_i по M_{x_j} : $\hat{x}_i | M_{x_j}$,
- **Сходимость** : Улучшение и сходимость качества прогноза к некоторому значению при увеличении L .

При выполнении условий: x_j – причина x_i .

Метод ковариационного перекрестного отображения

Algorithm Метод ковариационного, перекрестного отображения

Require: $\mathbf{x}_t = (x_{1t}, \dots, x_{nt})^\top$, \mathbf{C}_t ; $t = 1, \dots, N$

Ensure: $[\rho_1, \dots, \rho_n]$

for $t \in [L + 2, N - 1]$ **do**

$\mathbf{C}_{t_1}, \dots, \mathbf{C}_{t_{L+1}} \leftarrow \text{nn}(\mathbf{C}_t)$

for $i \in [1, n]$ **do**

$$[\alpha_1, \dots, \alpha_{L+1}] \leftarrow \frac{\exp(-\rho(\mathbf{C}_t, \mathbf{C}_{t_j})/\rho(\mathbf{C}_t, \mathbf{C}_{t_1}))}{\sum_{l=1}^{L+1} \exp(-\rho(\mathbf{C}_t, \mathbf{C}_{t_l})/\rho(\mathbf{C}_t, \mathbf{C}_{t_1}))}$$

$$\hat{x}_{i(t+1)} \leftarrow \sum_{k=1}^{L+1} d_k x_{it_k}$$

end for

end for

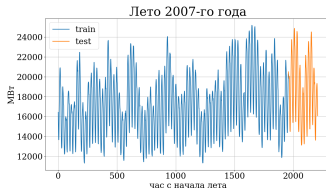
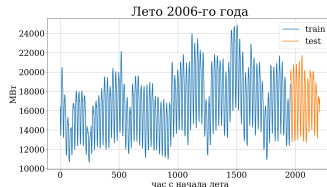
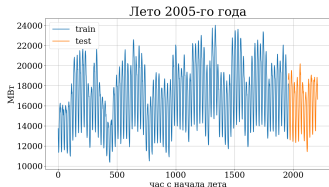
for $i \in [1, n]$ **do**

$$\rho_i \leftarrow \rho([x_{i(L+2)}, \dots, x_{iN}], [\hat{x}_{i(L+2)}, \dots, \hat{x}_{iN}])$$

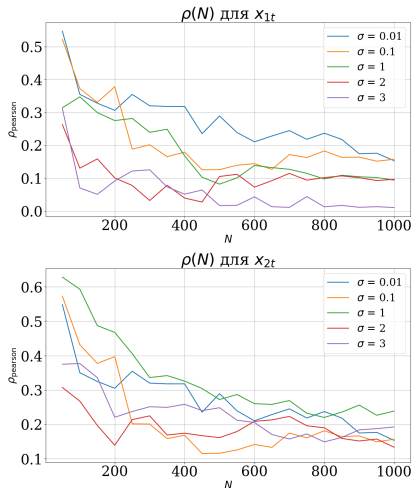
end for

Данные для вычислительного эксперимента

- Набор из 2 зашумленных синусов различной амплитуды и различной дисперсией нормального шума
- Набор потребления электроэнергии за одинаковый промежуток различных годов, набор акселерометров, набор стационарных ряды

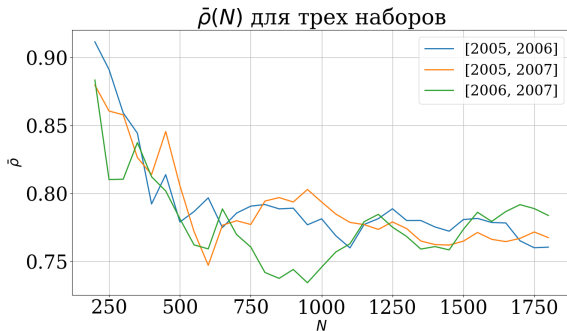


Метод КПО на синтетических данных



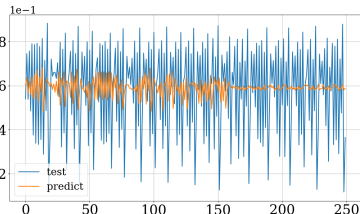
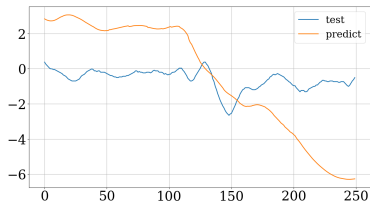
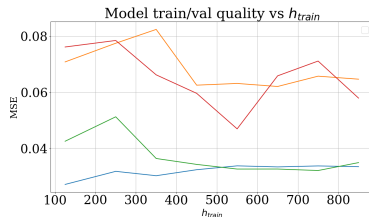
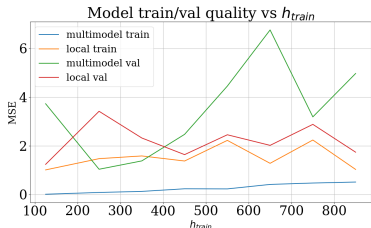
Результаты метода ковариационного, перекрестного отображения для $x_{1t} = \sin(t) + \varepsilon_{1t}$ и $x_{2t} = 3 \sin(t) + \varepsilon_{2t}$, $\varepsilon_{it} \sim \mathcal{N}(0, \sigma^2)$ при различных длинах истории N и различных значениях дисперсии шума

Метод КПО наборах потребления электроэнергии



Результаты метода ковариационного перекрестного отображения для трех наборов рядов

Метод улучшения базового прогноза



Результаты обучения модели в случае: а) невыполнения условий стационарности, ССМ, теста на ковариации б) при выполнении условий стационарности, при наличии ССМ-причинности, однако коэффициент корреляции низок (≈ 0.1).

1. Постановка задачи улучшения прогноза с помощью матрицы ковариации, дана интерпретация,
2. Анализ влияния шума на матрицы ковариации и расстояния между ними в римановом пространстве,
3. Метод ковариационного, перекрестного отображения,
3. Проведенные эксперименты с выбором прогностической модели и условия ее применимости.