

Введение

Важным аспектом в классификации временных рядов является выбор метрики и пространства, в котором происходит классификация. Одним из возможных подходов является построения признакового описания временного ряда $\mathbf{x}(t)$ на основе его кросс-спектра.

Автоковариационной матрицей временного ряда $\mathbf{x}(t)$ называется

$$R_{\mathbf{x}}(\tau) = E [\mathbf{x}(t)\mathbf{x}(t - \tau)^{\top}]$$

Матрицей кросс-спектральных плотностей называется

$$S_{\mathbf{x}}(f) = \sum_{\tau=-\infty}^{+\infty} R_{\mathbf{x}}(\tau)e^{-2\pi i f \tau}$$

В качестве признакового описания временного ряда $\mathbf{x}(t)$ можно использовать $S_{\mathbf{x}}(f)$, $f \in F$, где F - набор частот выбранных априорно. Таким образом, каждый \mathbf{x} временной ряд отображается в пространство положительно определенных симметричных матриц. Возникает вопрос - какую метрику следует использовать, как меру близости временных рядов в новом пространстве?

Риманова метрика на многообразии положительно определенных матриц

В качестве метрики предлагается использовать

$$\delta_2(S_1, S_2) = \|\log S_1^{-1/2} S_2 S_1^{-1/2}\|_2$$

.

Введем эту метрику формально и рассмотрим ее свойства.

Пространство всех симметричных матриц \mathcal{S}_n является гильбертовым пространством с произведением

$$\langle A, B \rangle = \text{tr}(A^{\top} B)$$

и имеет связанную с ним норму $\|A\|_2 = (\text{tr}(A^\top A))^{1/2}$. Множество симметричных положительных матриц представляет собой открытое подмножество \mathcal{P}_n в \mathcal{S}_n . Следовательно, это дифференцируемое многообразие. Касательное пространство к \mathcal{P}_n в любой точке S - это пространство, отождествляемое с \mathcal{S}_n . Произведение на \mathcal{S}_n приводит к римановой метрике на многообразии \mathcal{P}_n . Если $\gamma : [0, 1] \rightarrow \mathcal{P}_n$ - путь между двумя точками, то его длина определяется как

$$L(\gamma) = \int_a^b \|\gamma^{-1/2}(t)\gamma'(t)\gamma^{-1/2}(t)\|_2 dt$$

Для точек S_1 и S_2 определим

$$\delta_2(S_1, S_2) = \inf_{\gamma} \{L(\gamma) : \gamma(0) = S_1, \gamma(1) = S_2\}$$

Лемма 1

Для каждой обратимой матрицы X , которая задает биекцию в \mathcal{P}_n как $\Gamma_X(A) = X^\top A X$, и для каждого дифференцируемого пути γ

$$L(\Gamma_X \circ \gamma) = L(\gamma)$$

Доказательство. Подинтегральное выражение:

$$\begin{aligned} & \| (X^\top \gamma(t) X)^{-1/2} (X^\top \gamma(t) X)' (X^\top \gamma(t) X)^{-1/2} \|_2 = \\ & [\text{tr} (X^\top \gamma(t) X)^{-1} (X^\top \gamma(t) X)' (X^\top \gamma(t) X)^{-1} (X^\top \gamma(t) X)']^{1/2} = \\ & [\text{tr} X^{-1} \gamma^{-1}(t) \gamma'(t) \gamma^{-1}(t) \gamma'(t) X]^{1/2} = \\ & [\text{tr} \gamma^{-1}(t) \gamma'(t) \gamma^{-1}(t) \gamma'(t)]^{1/2} = \|\gamma^{-1/2}(t) \gamma'(t) \gamma^{-1/2}(t)\|_2 \end{aligned}$$

Интегрируя получаем $L(\Gamma_X \circ \gamma) = L(\gamma)$.

□

Производной экспоненты матрицы S называется матрица

$$De^S[H] = \lim_{t \rightarrow 0} \frac{e^{S+tH} - e^S}{t}$$

Лемма 2

Для всех S и H в \mathcal{S}_n выполнено:

$$\|e^{-S/2}De^S(H)e^{-S/2}\|_2 \geq \|H\|_2$$

Доказательство. Выберем ортонормированный базис, в котором $S = \text{diag}(\lambda_1, \dots, \lambda_n)$, тогда верна формула

$$De^S(H) = \left[\left[\frac{e^{\lambda_i} - e^{\lambda_j}}{\lambda_i - \lambda_j} h_{ij} \right] \right]$$

$$\text{Откуда } e^{-S/2}De^S(H)e^{-S/2} = \left[\left[\frac{\sinh(\lambda_i - \lambda_j)/2}{(\lambda_i - \lambda_j)/2} h_{ij} \right] \right]$$

Поскольку $\frac{\sinh x}{x} \geq 1$ для всех вещественных x , то утверждение леммы доказано.

□

Лемма 3

Пусть $S(t), 0 \leq t \leq 1$ - некоторый путь в \mathcal{S}_n и пусть $\gamma(t) = e^{S(t)}$. Тогда

$$L(\gamma) \geq \int_a^b \|S'(t)\|_2 dt$$

Доказательство. По формуле производной сложной функции $\gamma'(t) = De^{S(t)}(S'(t))$. Следовательно

$$L(\gamma) = \int_a^b \|e^{-S/2}De^S(S')e^{-S/2}\|_2 dt \geq \int_a^b \|S'\|_2 dt$$

□

Если $\gamma(t)$ - путь, соединяющий A и B в \mathcal{P}_n , то $S(t) = \log \gamma(t)$ - путь, соединяющий $\log A$ и $\log B$ в \mathcal{S}_n . Правая часть неравенства - длина этого пути в евклидовом пространстве \mathcal{S}_n . Иными словами, $\delta_2(A, B) = L(\gamma) \geq \|\log A - \log B\|_2$.

То есть отображение

$$(\mathcal{S}_n, || \cdot ||_2) \xrightarrow{\exp} (\mathcal{P}_n, \delta_2)$$

увеличивает или не изменяет расстояние между точками.

Линия, соединяющая точки S и H в \mathcal{S}_n задается

$$\gamma(t) = (1 - t)S + tH, 0 \leq t \leq 1$$

Лемма 4

Пусть A и B - коммутирующие матрицы в \mathcal{P}_n . Экспоненциальное отображение сопоставляет отрезок прямой $[\log A, \log B]$ в \mathcal{S}_n с геодезической линией $[A, B]$ в \mathcal{P}_n . В этом случае

$$\delta_2(A, B) = || \log A - \log B ||_2.$$

Доказательство. Линия $S(t)$, соединяющая точки $\log A$ и $\log B$ в \mathcal{S}_n задается

$$S(t) = (1 - t) \log A + t \log B, 0 \leq t \leq 1$$

Необходимо убедиться, что путь $\gamma(t) = e^{S(t)} = \exp((1 - t) \log A + t \log B)$, $0 \leq t \leq 1$, является единственным путем кратчайшей длины, соединяющим A и B в пространстве $(\mathcal{P}_n, \delta_2)$. Поскольку A и B коммутируют, $\gamma(t) = A^{1-t} B^t$ и $\gamma'(t) = (\log B - \log A) \gamma(t)$. По определению

$$L(\gamma) = \int_a^b || \log A - \log B ||_2 dt = || \log A - \log B ||_2$$

Лемма 3 говорит, что ни один путь не может быть короче этого. Таким образом, рассматриваемый путь γ имеет наименьшую возможную длину. Предположим, что $\bar{\gamma}$ - это другой путь, соединяющий A и B и имеющий ту же длину, что и путь γ . Тогда $\bar{S}(t) = \log \bar{\gamma}(t)$ - это путь, соединяющий $\log A$ и $\log B$ в \mathcal{S}_n , и, согласно Лемме 3, этот путь имеет

длину $\|\log A - \log B\|_2$. Но в евклидовом пространстве отрезок прямой - единственный кратчайший путь между двумя точками. Таким образом, $\bar{S}(t)$ является репараметризацией отрезка прямой $[\log A, \log B]$.

□

Применение этого рассуждения к любому подинтервалу $[0, a]$ из $[0, 1]$ приводит к тому, что параметризация $H(t) = (1-t) \log A + t \log B$ отрезка прямой $[\log A, \log B]$ - это тот отрезок, который изометрически отображается на $[A, B]$ вдоль всего интервала. Другими словами, естественная параметризация геодезической линии $[A, B]$ при коммутировании A и B задается $\gamma(t) = A^{1-t} B^t$, $0 \leq t \leq 1$, в том смысле, что $\delta_2(A, \gamma(t)) = t \delta_2(A, B)$ для каждого t . Общий случай получается из случая коммутирующих матриц с помощью применения изометрий Γ_X .

Теорема 1

Пусть A и B - любые два элемента \mathcal{P}_n . Тогда существует единственная геодезическая линия $[A, B]$, соединяющая A и B . Эта геодезическая линия имеет параметризацию

$$\gamma(t) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right) A^{1/2}, 0 \leq t \leq 1,$$

что естественно в том смысле, что для каждого t

$$\delta_2(A, \gamma(t)) = t \delta_2(A, B)$$

Кроме того

$$\delta_2(A, B) = \|\log A^{-1/2} B A^{-1/2}\|_2.$$

Доказательство. Матрицы I и $A^{-1/2} B A^{-1/2}$ коммутируют. Поэтому геодезическая линия $[I, A^{-1/2} B A^{-1/2}]$ естественно параметризуется

$$\gamma_0(t) = \left(A^{-1/2} B A^{-1/2} \right)^t$$

Применяя изометрию, задаваемую матрицей $A^{1/2}$, получаем путь

$$\gamma(t) = \Gamma_{A^{1/2}}(\gamma_0(t)) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^t A^{1/2},$$

соединяющий точки $\Gamma_{A^{1/2}}(I) = A$ и $\Gamma_{A^{1/2}}(A^{-1/2}BA^{-1/2}) = B$. Этот путь является геодезической линией $[A, B]$. Из леммы 4 следует

$$\delta_2(A, B) = \delta_2(I, A^{-1/2}BA^{-1/2}) = \|\log I - \log A^{-1/2}BA^{-1/2}\|_2 = \|\log A^{-1/2}BA^{-1/2}\|_2$$

Эта формула дает явное представление для метрики δ_2 , которая является римановой метрикой на многообразии \mathcal{P}_n .

□

Лемма 5

Для любых матриц $A, B \in \mathcal{P}_n$ выполнено

$$\delta_2(A, B) = \left(\sum_{i=1}^n \log^2 \lambda_i(A^{-1}B) \right)^{1/2},$$

где λ_i - собственные значения матрицы $A^{-1}B$.

Доказательство. Так как $A^{-1/2}BA^{-1/2}$ - симметричная матрица, то существует такое разложение, что $A^{-1/2}BA^{-1/2} = QDQ^\top$, тогда $\log A^{-1/2}BA^{-1/2} = Q \log DQ^\top$, где D - диагональная матрица из собственных чисел матрицы $A^{-1/2}BA^{-1/2}$.

$$\begin{aligned} \delta_2^2(A, B) &= \|\log A^{-1/2}BA^{-1/2}\|_2^2 = \|Q \log DQ^\top\|_2^2 = \text{tr}(Q \log DQ^\top)(Q \log DQ^\top) = \\ &= \text{tr} Q \log^2 DQ^\top = \text{tr} Q^\top Q \log^2 D = \text{tr} \log^2 D = \sum_{i=1}^n \log^2 \lambda(A^{-1/2}BA^{-1/2}) = \\ &= \sum_{i=1}^n \log^2 \lambda(A^{-1}B) \end{aligned}$$

В последнем переходе был использован тот факт, что $A^{-1/2}BA^{-1/2}$ и $A^{-1}B$ - подобные матрицы, а значит имеют одинаковые собственные числа. Матрицы C_1 и C_2 называются подобными, если существует обратимая матрица X такая, что $C_2 = X^{-1}C_1X$.

□

Элементы пространства центрированных нормальных распределений параметризуются своими ковариационными матрицами. То есть существует отображение между пространством центрированных нормальных распределений и симметричных положительно определенных матриц \mathcal{P}_n . Информационная метрика Фишера, индуцируемая на пространстве параметров, оказывается совпадает с введенной выше римановой метрикой в \mathcal{P}_n .

Список литературы

- [1] Bhatia Rajendra. Positive Definite Matrices. — Princeton University Press 41 William St. Princeton, NJ United States, 2015. — September.
- [2] Pedro Luiz Coelho Rodrigues Marco Congedo Christian Jutten. Multivariate Time-Series Analysis Via Manifold Learning. — 2018. — Jun.
- [3] Costa Sueli I. R., Santos Sandra A., Strapasson João E. Fisher information distance: a geometrical reading. — 2014. — 1210.2354.