Выбор прогностических моделей в римановых фазовых пространствах

A Preprint

Эйнуллаев Алтай Эльшан оглы
Кафедра интеллектуальных систем
einullaev.ae@phystech.edu

Вадим Викторович Стрижов
д. ф-м. н.
strijov@forecsys.ru

Abstract

Матрицы ковариации многомерных временных рядов лежат в римановом пространстве SPD матриц. Свойства этого пространства используются для решения задач классификации многомерных временных рядов. Проблема с использованием этого же подхода для прогнозирования многомерных временных рядов заключается в том, что в отличие от задач классификации, нет возможности решать задачу в пространстве SPD матриц. Предлагаются прогностические модели, использующие риманову геометрию SPD матриц, и исследуется качество их прогноза в зависимости от свойств прогнозируемых временных рядов.

Keywords SPD матрицы · Касательное пространство · tSSA · Матрица ковариаций

1 Введение

Матрицы ковариаций многомерных временных рядов принадлежат гладкому многообразию SPD матриц. Свойства этого пространства [1] могут быть применены для решения различных задач, связанных с многомерными временными рядами. В частности, были предложены методы классификации EEG сигналов, основанные на римановой геометрии пространства SPD матриц [2], [3], [4]. Главная идея состоит в том, чтобы перейти в пространство SPD матриц и использовать риманову метрику для решения задачи. Также был разработан метод, в которых матрицы ковариации отображались в касательное пространство, являющееся евклидовым, и в нем классифицировались с помощью известных методов классификации, например SVM. [5].

В отличие от классификации, для прогнозирования многомерных временных рядов не разработаны прогностические модели, использующие риманову геометрию матриц ковариации. В первую очередь, это связано с тем, что нельзя решить задачу в пространстве матриц ковариации, т.к. модель должна прогнозировать исходные временные ряды, а не метку, как в задаче классификации. Однако, можно рассматривать векторное представление матриц ковариации как описание взаимосвязи фазовах траекторий многомерных временных рядов и использовать эту информацию для улучшения качества прогноза.

2 Постановка задачи

Ставится задача прогнозирования многомерного временного ряда $\mathbf{x}_t = [x_{1t}, \dots, x_{nt}]^{\mathrm{T}}$. Каждой компоненте временному ряду можно поставить в соответствие фазовое пространство векторов задержек размерности L. Матрица ковариации для $\mathbf{x}_t$ формально определяется следующим образом:

$$\Sigma = \mathrm{E}(\mathbf{x}_t - \mathrm{E}(\mathbf{x}_t)(\mathbf{x}_t - \mathrm{E}(\mathbf{x}_t))^{\mathrm{T}}. \quad (1)$$

Будем использовать выборочую матрицу ковариации, в качестве оценки матрицы ковариации в каждый момент времени:

$$\mathbf{P}_t = \frac{1}{L-1}\mathbf{X}_t\mathbf{X}_t^{\mathrm{T}}, \tag{2}$$

где $\mathbf{X}_t = [\mathbf{x}_{t-L+1}, \dots \mathbf{x}_t]$. Таким образом, $\mathbf{P}_t$ описывает взаимосвязь между точками фазовых пространств различных компонент временного ряда.

2.1 Риманова геометрия

Введем определения и инструменты римановой геометрии, необходимые для описания предлагаемых прогностических моделей. Определенные выборочные матрицы ковариации принадлежат к некоторому многообразию.

2.1.1 Обозначения

Пусть $S(n) = \{\mathbf{S} \in M(n), \mathbf{S}^{\mathrm{T}} = \mathbf{S}\}$ – пространство всех $n \times n$ симметричных матриц в пространстве квадратных вещественных матриц $M(n)$ и $P(n) = \{\mathbf{P} \in S(n), \mathbf{u}^{\mathrm{T}}\mathbf{P}\mathbf{u} > 0, \forall \mathbf{u} \in \mathbb{R}\}$. Для SPD матриц из $P(n)$, матричная экспонента вычисляется с помощью разложения по собственным значениям матрицы $\mathbf{P}$:

$$\mathbf{P} = \mathbf{U}\,\mathrm{diag}(\sigma_1, \dots, \sigma_n)\,\mathbf{U}^{\mathrm{T}}, \tag{3}$$

где $\sigma_1 > \sigma_2 > \dots > \sigma_n > 0$ – собственные числа и $\mathbf{U}$ – матрица собственных векторов $\mathbf{P}$. Тогда

$$\exp(\mathbf{P}) = \mathbf{U}\,\mathrm{diag}(\exp(\sigma_1), \dots, \exp(\sigma_n))\,\mathbf{U}^{\mathrm{T}}. \tag{4}$$

Обратная операция:

$$\log(\mathbf{P}) = \mathbf{U}\,\mathrm{diag}(\log(\sigma_1), \dots, \log(\sigma_n))\,\mathbf{U}^{\mathrm{T}}. \tag{5}$$

Для матриц $\mathbf{P} \in P(n)$, $\mathbf{S} \in S(n)$ справедливы следующие утверждения: $\log(\mathbf{P}) \in S(n)$, $\exp(\mathbf{S}) \in P(n)$. Будем обозначать $\mathbf{P}^{\frac{1}{2}}$ такую симметричную матрицу $\mathbf{A}$, что $\mathbf{A}\mathbf{A} = \mathbf{P}$.

2.1.2 Метрика в римановом пространстве

Пространство SPD матриц $P(n)$ является гладким многообразием $\mathcal{M}$. Касательное пространство к $P(n)$ в $\mathbf{P}$ – векторное пространство $T_{\mathbf{P}}$, лежащее в $S(n)$. Размерности многообразия и касательного пространства равны $m = n(n+1)/2$.

В касательном пространстве определено скалярное произведение:

$$\langle \mathbf{S}_1, \mathbf{S}_2 \rangle_{\mathbf{P}} = \mathrm{tr}(\mathbf{S}_1\mathbf{P}^{-1}\mathbf{S}_2\mathbf{P}^{-1}). \tag{6}$$

Пусть $\mathbf{\Gamma}(t) : [0, 1] \to P(n)$ дифференцируемая кривая из $\Gamma(0) = \mathbf{P}_1$ в $\Gamma(1) = \mathbf{P}_2$. Длина кривой $\mathbf{\Gamma}(t)$ равна

$$L(\mathbf{\Gamma}(t)) = \int_0^1 \|\dot{\mathbf{\Gamma}}(t)\|_{\mathbf{\Gamma}(t)} dt, \tag{7}$$

с нормой, порожденной скалярным произведением, определенным выше. Кривая минимальной длины, соединяющая две точки многообразия называется геодезической линией и риманово расстояние между ними равно ее длине. Таким образом, геодезическое расстояние [1]:

$$\delta(\mathbf{P}_1, \mathbf{P}_2) = \|\log(\mathbf{P}_1^{-1}\mathbf{P}_2)\|_F = \left(\sum_{i=1}^{n} \log^2 \lambda_i\right)^{1/2}, \tag{8}$$

где λ_i, $i = 1, \ldots, n$ – собственные числа матрицы $\mathbf{P}_1^{-1}\mathbf{P}_2$.

2.1.3 Касательное пространство

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

$$\mathbf{P}_i = \mathrm{Exp}_{\mathbf{P}}(\mathbf{S}_i) = \mathbf{P}^{\frac{1}{2}} \exp(\mathbf{P}^{-\frac{1}{2}}\mathbf{S}_i\mathbf{P}^{-\frac{1}{2}})\mathbf{P}^{\frac{1}{2}}. \tag{9}$$

Обратное отображение:

$$\mathbf{S}_i = \mathrm{Log}_{\mathbf{P}}(\mathbf{P}_i) = \mathbf{P}^{\frac{1}{2}} \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{P}_i\mathbf{P}^{-\frac{1}{2}})\mathbf{P}^{\frac{1}{2}}. \tag{10}$$

Таким образом, с помощью логарифмического и экспоненциального отображения можем переходить в касательное пространство и обратно.

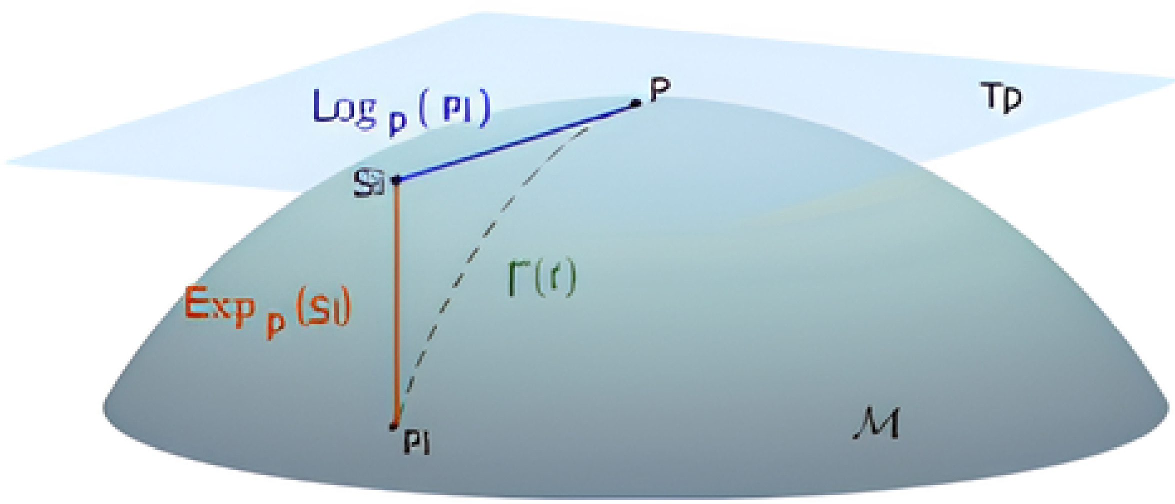

Рис. 1: Касательное пространство в точке $\mathbf{P}$, $\mathbf{S}_i$ – касательный вектор в $\mathbf{P}$, $\Gamma_i(t)$ – кратчайшая линия, соединяющая $\mathbf{P}$, $\mathbf{P}_i$ в $P(n)$.

2.2 Прогностическая модель

Требуется выбрать две модели f, h:

$$\mathbf{x}_t \xrightarrow{f} \hat{\mathbf{x}}_{t+1}, \tag{11}$$

$$\mathbf{x}_t \xrightarrow{g} \mathbf{S}_t \xrightarrow{h} \hat{\mathbf{S}}_{t+1}, \tag{12}$$

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:

$$\hat{\mathbf{x}}_{t+1} = F(\hat{\mathbf{x}}_{t+1}, \hat{\mathbf{S}}_{t+1}) = F(f(\mathbf{x}_t), h(g(\mathbf{x}_t))). \tag{13}$$

3 Вычислительный эксперимент

4 Сравнение результатов

5 Заключение

Список литературы

[1] Maher Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM journal on matrix analysis and applications*, 26(3):735–747, 2005.

[2] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Riemannian geometry applied to bci classification. In *International conference on latent variable analysis and signal separation*, pages 629–636. Springer, 2010.

[3] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Multiclass brain–computer interface classification by riemannian geometry. *IEEE Transactions on Biomedical Engineering*, 59(4):920–928, 2011.

[4] Marco Congedo, Alexandre Barachant, and Rajendra Bhatia. Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces*, 4(3):155–174, 2017.

[5] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Classification of covariance matrices using a riemannian-based kernel for bci applications. *Neurocomputing*, 112:172–178, 2013.