

Тензорный анализ сингулярного спектра

Хромов Никита Андреевич, Голяндина Нина Эдуардовна

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Процессы управления и устойчивость
2 апреля 2024, Санкт-Петербург

Постановка задачи

$X = (x_1, x_2, \dots, x_N)$, $x_i \in \mathbb{R}$ — вещественный временной ряд.

$$X = T + P + R.$$

T — тренд, P — сезонность, R — шум.

Возможные задачи:

- 1 Выделение сигнала из ряда: нахождение $S = T + P$,
- 2 Отделение компонент сигнала: нахождение T и P ,
- 3 Нахождение параметров сигнала в параметрической модели.

Методы, основанные на подпространстве сигнала:

- SSA (задачи 1 и 2)
(Golyandina et al. (2001), Analysis of time series structure: SSA and related techniques)
- ESPRIT (задача 3)
(Roy, Kailath (1989), ESPRIT-estimation of signal parameters via rotational invariance techniques)

Имеющиеся результаты

В работе Pary et al. (2005) была предложена тензорная модификация метода ESPRIT и экспериментально показано её преимущество для конкретной модели.

Цель: расширение предложенного Pary алгоритма для решения задачи выделения сигнала, исследование свойств тензорных модификаций методов семейства SSA с точки зрения точности выделения сигнала.

Модель одномерного сигнала

$$X = (x_0, x_1, \dots, x_{N-1}) = S + R,$$

S — сигнал, R — шум.

$$s_n = \sum_{j=1}^R a_j e^{-\alpha_j n} \cos(2\pi\omega_j n + \varphi_j)$$

Параметры:

$a_j \in \mathbb{R} \setminus \{0\}$ — амплитуды, $\alpha_j \in \mathbb{R}$ — степени затухания,

$\omega_j \in [0, 1/2]$ — частоты, $\varphi_j \in [0, 2\pi)$ — фазы.

Модель многомерного сигнала

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_P \end{pmatrix}, \quad X_p = S_p + R_p \text{ — одномерные ряды.}$$

Общий случай:

$$s_n^{(p)} = \sum_{j=1}^{R(p)} a_j^{(p)} e^{-\alpha_j^{(p)} n} \cos \left(2\pi \omega_j^{(p)} n + \varphi_j^{(p)} \right)$$

Рассматриваемый случай:

$$s_n^{(p)} = \sum_{j=1}^R a_j^{(p)} e^{-\alpha_j n} \cos \left(2\pi \omega_j n + \varphi_j^{(p)} \right)$$

Описание алгоритма MSSA

X — P -мерный временной ряд длины N с сигналом S , L — длина окна, $K = N - L + 1$.

Оператор вложения:

$$\mathbb{H}_L(X_p) = \begin{pmatrix} x_0^{(p)} & x_1^{(p)} & x_2^{(p)} & \dots & x_{K-1}^{(p)} \\ x_1^{(p)} & x_2^{(p)} & \ddots & \dots & \vdots \\ x_2^{(p)} & \ddots & \ddots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & x_{N-2}^{(p)} \\ x_{L-1}^{(p)} & \dots & \dots & x_{N-2}^{(p)} & x_{N-1}^{(p)} \end{pmatrix}$$

Описание алгоритма MSSA

Параметры алгоритма: $L, R : R \leq L < N, K \geq L$.

R — число компонент, отнесённых к сигналу.

Схема алгоритма MSSA для выделения сигнала

❶ Вложение

$$\mathbf{X} \mapsto \mathbf{H} = [\mathbb{H}(\mathbf{X}_1) : \mathbb{H}(\mathbf{X}_2) : \cdots : \mathbb{H}(\mathbf{X}_P)] \in \mathbb{R}^{L \times KP},$$

❷ Разложение $\mathbf{H} = \sum_{i=1}^d \sqrt{\lambda_i} U_i V_i^T, \quad d \leq L$

❸ Группировка $\tilde{\mathbf{S}} = \sum_{i=1}^R \sqrt{\lambda_i} U_i V_i^T, \quad R \leq d$

$$\tilde{\mathbf{S}} = [\tilde{\mathbf{S}}_1 : \tilde{\mathbf{S}}_2 : \cdots : \tilde{\mathbf{S}}_P], \quad \tilde{\mathbf{S}}_p \in \mathbb{R}^{L \times K}$$

❹ Восстановление Матрицы $\tilde{\mathbf{S}}_p$ усредняются вдоль побочных диагоналей: $\tilde{s}_n^{(p)} = \text{mean} \left\{ \left(\tilde{\mathbf{S}}_p \right)_{i,j} \mid i + j - 2 = n \right\}.$

Ранг сигнала

S имеет ранг r , если $\forall L : r \leq \min(L, K) \quad \text{rank } \mathbb{H}_L(S) = r$.

Рекомендуемый выбор параметра R в алгоритме: $R = r$.

Примеры

- $s_n = Ae^{\alpha n} \cos(2\pi\omega n + \varphi),$
 $A \neq 0, \alpha \in \mathbb{R}, \omega \in [0, 1/2], \varphi \in [0, 2\pi)$

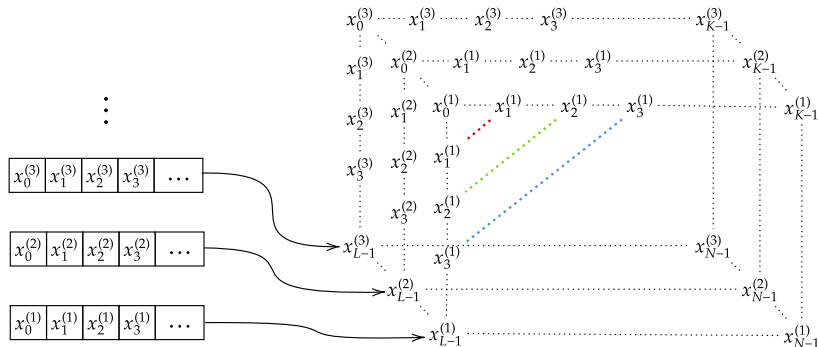
$$r(\omega) = \begin{cases} 1, & \omega \in \{0, 1/2\}, \\ 2, & \omega \in (0, 1/2). \end{cases}$$

-

$$s_n^{(p)} = \sum_{j=1}^R a_j^{(p)} e^{-\alpha_j n} \cos(2\pi\omega_j n + \varphi_j^{(p)})$$

$$r = \sum_{(\omega, \alpha) \in \Omega} r(\omega), \quad \Omega — \text{все уникальные пары } (\omega_j, \alpha_j).$$

Построение траекторного тензора



$$L < N, \quad K = N - L + 1 \geq L$$

Разложение и группировка

- HOSVD траекторного тензора \mathcal{H} имеет вид

$$\mathcal{H} = \sum_{l=1}^L \sum_{k=1}^K \sum_{p=1}^P c_{lkp} \mathbf{U}_l^{(1)} \circ \mathbf{U}_k^{(2)} \circ \mathbf{U}_p^{(3)}.$$

- Этап группировки в алгоритме HO-MSSA имеет вид

$$\tilde{\mathcal{H}} = \sum_{l=1}^{R_1} \sum_{k=1}^{R_2} \sum_{p=1}^{R_3} c_{lkp} \mathbf{U}_l^{(1)} \circ \mathbf{U}_k^{(2)} \circ \mathbf{U}_p^{(3)},$$

$R_i \leq \min(L, K, P)$ — параметры алгоритма.

Ранги тензора

n -ранг тензора: размерность пространства, порождённого векторами вдоль n -го измерения ($\text{rank}_n(\mathcal{A})$).

Теорема

Пусть многомерный временной ряд

$$\left(x_0^{(p)}, x_1^{(p)}, \dots, x_{N-1}^{(p)} \right), \quad p = 1, 2, \dots, P$$

имеет ранг d в терминах MSSA, тогда для траекторного тензора \mathcal{H} , построенного по любой длине окна $L < N$ такой, что $\min(L, K) \geq d$, выполняется $\text{rank}_1(\mathcal{H}) = \text{rank}_2(\mathcal{H}) = d$, а 3-ранг этого тензора равен рангу матрицы, в строках которой записаны заданные одномерные ряды.

Ранги сигнала в HO-MSSA

- 1- и 2-ранги траекторного тензора \mathcal{H} сигнала S совпадают с рангом этого сигнала в терминах MSSA.
- Однако ранг третьего измерения имеет иной смысл.

На этапе группировки рекомендуется брать $R_1 = R_2 = r$ и $R_3 = r_3$, где r — MSSA-ранг сигнала, r_3 — ранг матрицы, составленной из S_p .

Примеры

$$s_n^{(p)} = \sum_{j=1}^R a_j^{(p)} e^{-\alpha_j n} \cos \left(2\pi \omega_j n + \varphi_j^{(p)} \right), \quad n \in \overline{0 : N-1},$$

$$p \in \overline{1 : P}, a_j^{(p)} \neq 0, \alpha_j \in \mathbb{R}, \omega_j \in (0, 1/2), \varphi_j^{(p)} \in [0, 2\pi)$$

- 1 $\omega_i \neq \omega_j, \varphi_j^{(p)} = \varphi_j^{(m)} \implies r = 2R, r_3 = R,$
- 2 $\omega_i \neq \omega_j, \varphi_j^{(p)} = b_j p + c_j \implies r = 2R, r_3 = 2R.$

Численные сравнения

1 Одинаковые фазы

$$x_n^{(p)} = c_1^{(p)} e^{-0.01n} \cos(2\pi 0.2n) + c_2^{(p)} e^{-0.02n} \cos(2\pi 0.22n) + \varepsilon_n^{(p)},$$

2 Линейно меняющиеся фазы

$$x_n^{(p)} = c_1^{(p)} e^{-0.01n} \cos(2\pi 0.2n + p\pi/6) + c_2^{(p)} e^{-0.02n} \cos(2\pi 0.22n + p\pi/9) + \varepsilon_n^{(p)},$$

В обоих случаях $n = 0, 1, \dots, 24$, $p = 1, 2, \dots, 12$, $c_k^{(p)} \sim N(0, 1)$, $\varepsilon_n^{(p)} \sim N(0, 0.02)$ и независимы.

Точность сравнивалась по RMSE по 1000 реализациям шума $\varepsilon_n^{(p)}$, сравнение проводилось на одних и тех же реализациях шума при выборе оптимальных для каждого метода параметров L , R и R_3 .

Численные сравнения

Таблица: RMSE оценки многомерного сигнала

	MSSA	HO-MSSA
равные фазы	0.0107	0.0079
линейные фазы	0.00924	0.00918

Выводы

- HOSVD SSA и HOSVD MSSA являются прямыми обобщениями SSA и MSSA, однако устроены существенно сложнее и имеют большую трудоемкость.
- Оба расширения усложняют алгоритм необходимостью подбора дополнительного параметра.
- HOSVD SSA выделяет сигнал менее точно, чем SSA.
- HOSVD MSSA выделяет сигнал точнее, чем MSSA.