

Санкт-Петербургский государственный университет
Прикладная математика и информатика

Учебная практика 4 (научно-исследовательская работа) (семестр 7)

«ТЕНЗОРНЫЙ АНАЛИЗ СИНГУЛЯРНОГО СПЕКТРА»

Выполнил:

Хромов Никита Андреевич

20.Б04-мм

Научный руководитель:

д. ф.-м. н., доцент

Голяндина Н. Э.

Санкт-Петербург

2024

Оглавление

| | | |
|------|---|-----------|
| 1. | Введение | 3 |
| 2. | HOSVD и его свойства | 5 |
| 2.1. | Свойства HOSVD | 6 |
| 3. | Описание метода HO-SSA | 10 |
| 4. | Свойства HO-SSA | 13 |
| 4.1. | Разделимость рядов в терминах HO-SSA | 13 |
| 4.2. | Примеры разделимости рядов в тензорном случае | 15 |
| 4.3. | Ранг ряда в терминах HO-SSA | 16 |
| 5. | Примеры использования HO-SSA | 16 |
| 5.1. | Сравнение HO-SSA и SSA | 18 |
| 6. | Описание метода HOSVD MSSA для выделения сигнала | 22 |
| 7. | Свойства HOSVD MSSA | 22 |
| 8. | Численные сравнения полученных методов с известными | 26 |
| 8.1. | Принцип выбора числа компонент, относимых к сигналу | 26 |
| 8.2. | Численные сравнения | 28 |
| 9. | Альтернативные тензорные разложения | 35 |
| 10. | Заключение | 39 |
| | Список литературы | 40 |

1. Введение

Singular spectrum analysis (SSA) [1] является популярным методом анализа временных рядов. Этот метод используется, в частности, для выделения сигнала и выделения тренда и периодических компонент из временного ряда. Метод SSA основан на сингулярном разложении особой матрицы построенной по временному ряду, называемой траекторной.

В работах [2, 3, 4] предлагается обобщение метода SSA, Tensor SSA, который основан на некотором тензорном разложении траекторного тензора, построенного по временному ряду. В работах [2, 3] рассматривается задача выделения сигнала из временного ряда, а в работе [4] — задача оценки частот периодических компонент сигнала. Причём, в этих работах утверждается о преимуществах Tensor SSA над обычным SSA.

Существует множество видов тензорных разложений, например High-Order Singular Value Decomposition (HOSVD) [5], Canonical Polyadic Decomposition (CPD) [6, 7], Tucker decomposition [8]. В частности, в работах [2, 3] используется разложение CPD, а в работе [4] — HOSVD.

Была поставлена задача изучить различные варианты тензорных разложений, реализовать метод Tensor SSA, выбрав некоторые из них, и сравнить с методом Basic SSA по точности выделения сигнала и компонент сигнала, а также рассмотреть расширение метода SSA на многомерные ряды — метод MSSA [9], сформулировать и реализовать тензорную модификацию этого метода и сравнить её с другими методами семейства SSA по точности выделения сигнала. В качестве первого метода разложения был выбран метод HOSVD, который является обобщением метода SVD для матриц.

Способ построения траекторного тензора и его разложение были выбраны из предложенных в статье [4], однако в отличие от этой статьи, в этой работе изучается применение выбранных средств в задаче выделения сигнала из временного ряда и в задаче отделения компонент сигнала.

В разделе 2 приведено описание разложения HOSVD и некоторые свойства этого разложения, необходимые для доказательства ключевых утверждений о тензорных модификациях SSA. В разделах 3, 4 представлено описание тензорных модификаций метода SSA — HOSVD SSA и HOOI SSA (оба этих метода в совокупности будем называть High-Order SSA или HO-SSA), и приведены утверждения, позволяющие применять некоторые определения и свойства из теории базового SSA к модифицированным ал-

горитмам. В разделе 5 приведены примеры применения методов HO-SSA в задачах выделения сигнала из ряда и отделения компонент и численные сравнения этих методов с Basic SSA. По итогам этого раздела была выявлена проблема сложности сопоставления компонент разложения с компонентами сигнала, и было установлено, что в большинстве рассмотренных случаев методы HO-SSA имеют меньшую точность, чем метод Basic SSA. В ходе работы в этом семестре были добавлены разделы 6, 7 и 8, в которых описывается тензорная модификация метода MSSA — HOSVD MSSA и проводятся численные сравнения этого метода с известными. В результате были получены утверждения, позволяющие применять метод HOSVD MSSA для выделения сигнала, и установлено преимущество этого метода над методом MSSA.

2. HOSVD и его свойства

Ключевым этапом в алгоритме HO-SSA является применение тензорного разложения HOSVD [5] к некоторому тензору. Приведём определение этого разложения и некоторые его свойства. Все утверждения из этого раздела и их доказательства приведены в статье [5].

Определение 2.1 (Произведение тензора и матрицы по измерению). Пусть \mathcal{A} — тензор размерности $I_1 \times I_2 \times \dots \times I_M$, \mathbf{U} — матрица размерности $J_n \times I_n$, тогда произведением тензора \mathcal{A} и матрицы \mathbf{U} по измерению n ($\mathcal{A} \times_n \mathbf{U}$) называется тензор размерности $I_1 \times I_2 \times \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_M$, который считается по формуле

$$(\mathcal{A} \times_n \mathbf{U})_{i_1 i_2 \dots i_{n-1} j_n i_{n+1} \dots i_M} = \sum_{i_n=1}^{I_n} a_{i_1 i_2 \dots i_{n-1} i_n i_{n+1} \dots i_M} u_{j_n i_n}.$$

Теорема 1 (Сингулярное разложение порядка M). Любой комплекснозначный тензор \mathcal{A} размерности $I_1 \times I_2 \times \dots \times I_M$ может быть представлен в виде произведения

$$\mathcal{A} = \mathcal{Z} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \dots \times_M \mathbf{U}^{(M)}, \quad (1)$$

в котором

1. $\mathbf{U}^{(n)} = [U_1^{(n)} : U_2^{(n)} : \dots : U_{I_n}^{(n)}]$ — унитарная матрица,
2. \mathcal{Z} — комплекснозначный тензор размерности $I_1 \times I_2 \times \dots \times I_M$, в котором каждый подтензор $\mathcal{Z}_{i_n=\alpha}$, полученный фиксированием индекса $i_n = \alpha$ удовлетворяет следующим свойствам

- а. полная ортогональность: подтензоры $\mathcal{Z}_{i_n=\alpha}$ и $\mathcal{Z}_{i_n=\beta}$ ортогональны для всех возможных значений $n, \alpha, \beta : \alpha \neq \beta$:

$$\langle \mathcal{Z}_{i_n=\alpha}, \mathcal{Z}_{i_n=\beta} \rangle = 0 \quad \alpha \neq \beta,$$

- б. упорядоченность: подтензоры расположены в порядке убывания их норм Фробениуса:

$$\|\mathcal{Z}_{i_n=1}\| \geq \|\mathcal{Z}_{i_n=2}\| \geq \dots \geq \|\mathcal{Z}_{i_n=I_n}\| \quad (2)$$

для всех $n \in \overline{1 : M}$.

Определение 2.2 (Сингулярное разложение тензора). Разложение вида (1) будем называть сингулярным разложением тензора \mathcal{A} порядка M или HOSVD тензора \mathcal{A} .

Определение 2.3 (Сингулярное число тензора). Обозначим $\sigma_i^{(n)} = \|\mathcal{Z}_{i_n=i}\|$ и будем называть $\sigma_i^{(n)}$ i -м сингулярным числом тензора \mathcal{A} по измерению n .

Определение 2.4 (Сингулярный вектор тензора). Векторы $U_i^{(n)}$ будем называть i -м сингулярным вектором тензора \mathcal{A} по измерению n .

Замечание. Представление (1) можно записать в виде

$$\mathcal{A} = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_M=1}^{I_M} \mathcal{Z}_{i_1, i_2, \dots, i_M} U_{i_1}^{(1)} \circ U_{i_2}^{(2)} \circ \dots \circ U_{i_M}^{(M)}. \quad (3)$$

Такое представление удобнее для описания алгоритма HO-SSA.

2.1. Свойства HOSVD

Многие свойства метода SSA являются следствиями свойств SVD. В свою очередь, многие свойства HOSVD являются аналогами свойств SVD. Таким образом, аналогичность свойств SSA и HO-SSA может быть выведена из аналогичности некоторых свойств SVD и HOSVD.

Определение 2.5 (Развёртка тензора по измерению). Пусть \mathcal{A} — тензор размерности $I_1 \times I_2 \times \dots \times I_M$, тогда его развёртка по n -му измерению — это матрица $\mathbf{A}_{(n)}$ (или $[\mathbf{A}]_n$) размерности $I_n \times I_{n+1}I_{n+2} \dots I_M I_1 I_2 \dots I_{n-1}$, в которой элемент $a_{i_1 i_2 \dots i_M}$ тензора содержится в строке i_n и столбце с номером равным

$$(i_{n+1} - 1)I_{n+2}I_{n+3} \dots I_M I_1 I_2 \dots I_{n-1} + (i_{n+2} - 1)I_{n+3}I_{n+4} \dots I_M I_1 I_2 \dots I_{n-1} + \dots \\ + (i_M - 1)I_1 I_2 \dots I_{n-1} + (i_1 - 1)I_2 I_3 \dots I_{n-1} + (i_2 - 1)I_3 I_4 \dots I_{n-1} + \dots + i_{n-1}.$$

Пусть \mathcal{A} — тензор размерности $I_1 \times I_2 \times \dots \times I_M$, применим SVD к развёрткам этого тензора вдоль каждого измерения, в результате чего получим M матриц $\mathbf{U}^{(n)}$, составленных из левых сингулярных векторов соответствующих развёрток. Теперь определим тензор \mathcal{Z} следующим образом

$$\mathcal{Z} = \mathcal{A} \times_1 \mathbf{U}^{(1)\mathrm{H}} \times_2 \mathbf{U}^{(2)\mathrm{H}} \dots \times_M \mathbf{U}^{(M)\mathrm{H}}.$$

Тогда исходный тензор \mathcal{A} можно представить в виде

$$\mathcal{A} = \mathcal{Z} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \dots \times_M \mathbf{U}^{(M)}. \quad (4)$$

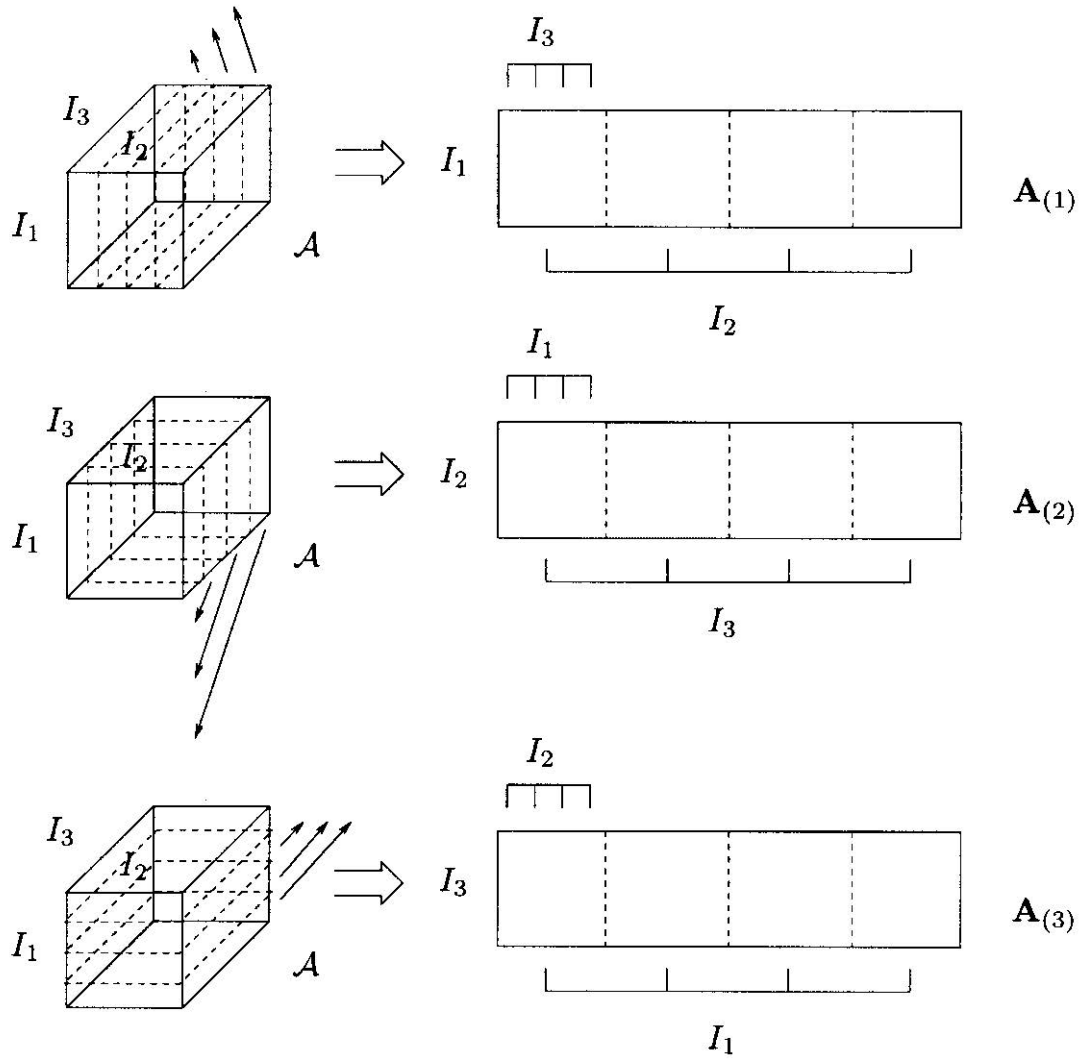


Рис. 1. Развёртка тензора \mathcal{A} размерности $I_1 \times I_2 \times I_3$ в матрицы $\mathbf{A}_{(1)}$, $\mathbf{A}_{(2)}$, $\mathbf{A}_{(3)}$ размерностей $I_1 \times (I_2 I_3)$, $I_2 \times (I_3 I_1)$, $I_3 \times (I_1 I_2)$ соответственно

Утверждение 2.1. *Представление тензора \mathcal{A} в виде (4) является HOSVD этого тензора.*

Из-за этой связи HOSVD со стандартным матричным SVD для многих свойств SVD существуют аналогичные свойства HOSVD.

Свойство 2.1 (Единственность).

1. Все сингулярные числа по каждому измерению определяются однозначно.
2. Если сингулярные числа по измерению n различны, то сингулярные векторы по измерению n определены в точности до умножения на коэффициент единичной нормы. Если $U_\alpha^{(n)}$ умножается на $e^{i\theta}$, то $\mathcal{Z}_{i_n=\alpha}$ должен быть умножен на обратный коэффициент $e^{-i\theta}$.

Сингулярные векторы по измерению n , соответствующие одному и тому же сингулярному числу по измерению n , могут быть заменены любой унитарной линейной комбинацией. Соответствующие подтензоры $\{\mathcal{Z}_{in=\alpha}\}$ должны быть пересчитаны обратным образом. Формально $\mathbf{U}^{(n)}$ можно заменить на $\mathbf{U}^{(n)}\mathbf{Q}$, где \mathbf{Q} — блочно-диагональная матрица, состоящая из унитарных блоков, в которой блочное разбиение соответствует разбиению $\mathbf{U}^{(n)}$ на наборы сингулярных векторов по измерению n , соответствующих одинаковым сингулярным значениям по измерению n . При этом тензор \mathcal{Z} должен быть заменён на $\mathcal{Z} \times_n \mathbf{Q}^H$.

В случае вещественнозначных тензоров единственность имеется в точности до знака, что соответствует умножению на унитарную матрицу.

Свойство 2.2 (Обобщение). HOSVD тензора второго порядка сводится к его матричному SVD.

Это свойство означает, что результат применения HOSVD к тензору с двумя измерениями, т.е. матрице, совпадает с результатом применения SVD к этой же матрице, в точности до унитарных преобразований сингулярных векторов и матрицы сингулярных значений.

Перед формулировкой следующих свойств необходимо ввести несколько определений.

Определение 2.6 (n -ранг). n -рангом тензора \mathcal{A} называется размерность векторного пространства, порождённого векторами измерения n этого тензора. Обозначается $R_n = \text{rank}_n(\mathcal{A})$.

Замечание. В отличие от матричного случая, n -ранги тензора порядка выше 2 могут в общем случае отличаться.

Определение 2.7 (Тензорный ранг).

1. Говорят, что тензор \mathcal{A} размерности $I_1 \times I_2 \times \dots \times I_M$ имеет тензорный ранг равный 1, если он представим в виде

$$\mathcal{A} = a_1 \circ a_2 \circ \dots \circ a_M,$$

где $a_k \in \mathbb{C}^{I_k}$, а \circ обозначает внешнее произведение.

2. Говорят, что тензор \mathcal{A} имеет ранг R , если он представим в виде линейной комбинации R тензоров ранга 1, и такое R минимальное. Обозначение: $R = \text{rank}(\mathcal{A})$.

Замечание. В общем случае ранг тензора \mathcal{A} не равен его n -рангам, даже если они все равны между собой. Более того, всегда справедливо неравенство $\text{rank}_n(\mathcal{A}) \leq \text{rank}(\mathcal{A})$.

Свойство 2.3 (Связь n -ранга тензора и ранга его развёртки по измерению n). Векторы измерения n тензора \mathcal{A} являются столбцами его развёртки по измерению n и выполняется равенство

$$\text{rank}_n(\mathcal{A}) = \text{rank}(\mathbf{A}_{(n)}).$$

Свойство 2.4 (Связь n -ранга тензора и его HOSVD). Пусть имеется HOSVD тензора \mathcal{A} размерности $I_1 \times I_2 \times \dots \times I_M$

$$\mathcal{A} = \mathcal{Z} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \dots \times_M \mathbf{U}^{(M)},$$

тогда, по определению, тензор \mathcal{Z} удовлетворяет свойству упорядоченности сингулярных чисел

$$\|\mathcal{Z}_{i_n=1}\| \geq \|\mathcal{Z}_{i_n=2}\| \geq \dots \geq \|\mathcal{Z}_{i_n=I_n}\|$$

для всех $n \in \overline{1:M}$. Обозначим r_n — наибольший индекс такой, что $\|\mathcal{Z}_{i_n=r_n}\| > 0$. Тогда

$$\text{rank}_n(\mathcal{A}) = r_n. \quad (5)$$

Свойство 2.5 (Норма). Пусть имеется HOSVD тензора \mathcal{A} , представленное в виде (1), и пусть $R_n = \text{rank}_n(\mathcal{A})$, $n \in \overline{1:M}$. Тогда справедливо равенство

$$\begin{aligned} \|\mathcal{A}\|^2 &= \sum_{i=1}^{R_1} \left(\sigma_i^{(1)}\right)^2 = \sum_{i=1}^{R_2} \left(\sigma_i^{(2)}\right)^2 = \dots = \sum_{i=1}^{R_M} \left(\sigma_i^{(M)}\right)^2 = \\ &= \|\mathcal{Z}\|^2. \end{aligned}$$

Определение 2.8 (Ориентированная энергия). Ориентированной по измерению n энергией тензора $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_M}$ в направлении вектора $X \in \mathbb{C}^{I_n}$ единичной нормы называют выражение

$$\text{OE}_n(X, \mathcal{A}) = \|X^H \mathbf{A}_{(n)}\|^2.$$

Свойство 2.6 (Об ориентированной энергии). Направления экстремальной ориентированной энергии по измерению n соответствуют сингулярным векторам по измерению n , причем значение экстремальной энергии равно соответствующему квадрату сингулярного значения по измерению n .

Это означает, что векторы тензора \mathcal{A} по измерению n в основном содержат вклады в направлении $U_1^{(n)}$; на это направление приходится $\sigma_1^{(n)^2}$ энергии по отношению к общему количеству энергии в тензоре. Затем ориентированная энергия по измерению n достигает экстремума в направлении $U_2^{(n)}$, перпендикулярном $U_1^{(n)}$, с величиной $\sigma_2^{(n)^2}$, и так далее.

Свойство 2.7 (Приближение). Пусть имеется HOSVD тензора \mathcal{A} , представленное в виде (1), и пусть $R_n = \text{rank}_n(\mathcal{A})$. Определим тензор $\hat{\mathcal{A}}$ отбрасыванием наименьших сингулярных значений $\sigma_{I'_n+1}^{(n)}, \sigma_{I'_n+2}^{(n)}, \dots, \sigma_{R_n}^{(n)}$ для заданных значений I'_n , $n \in \overline{1:M}$, то есть заменяя нулями соответствующие части тензора \mathcal{Z} . Тогда верно

$$\|\mathcal{A} - \hat{\mathcal{A}}\|^2 \leq \sum_{i_1=I'_1+1}^{R_1} \left(\sigma_{i_1}^{(1)}\right)^2 + \sum_{i_2=I'_2+1}^{R_2} \left(\sigma_{i_2}^{(2)}\right)^2 + \dots + \sum_{i_M=I'_M+1}^{R_M} \left(\sigma_{i_M}^{(M)}\right)^2. \quad (6)$$

Это свойство является эквивалентом высшего порядка связи между SVD матрицы и ее наилучшим приближением, в смысле наименьших квадратов, матрицей более низкого ранга. Однако для тензоров ситуация совершенно иная. Отбрасывая наименьшие сингулярные значения измерения n , мы получаем тензор $\hat{\mathcal{A}}$ с рангом столбцов равным I'_1 , рангом строк равным I'_2 и т.д. Но этот тензор в общем случае не является наилучшим приближением при заданных ограничениях на ранги измерений. Тем не менее, предположение об упорядочении (2) подразумевает, что «энергия» \mathcal{A} в основном сосредоточена в части, соответствующей малым значениям i_1, i_2, \dots, i_M . Следовательно, если $\sigma_{I'_n}^{(n)} \gg \sigma_{I'_n+1}^{(n)}$ (например, если $I'_n = \text{rank}_n(\mathcal{A})$, то меньшие сингулярные значения по измерению n не существенны), то $\hat{\mathcal{A}}$ всё ещё можно считать хорошим приближением \mathcal{A} . Ошибка ограничена выражением (6).

3. Описание метода HO-SSA

Пусть дан временной ряд \mathbf{X} длины N

$$\mathbf{X} = (x_1, x_2, \dots, x_N).$$

Определение 3.1 (Траекторный тензор ряда). Траекторным тензором ряда \mathbf{X} с параметрами $I, L : 1 \leq I, L \leq N, I + L \leq N + 1$ будем называть тензор \mathcal{X} размерности $I \times L \times J$, $J = N - I - L + 2$, элементы которого удовлетворяют равенству

$$\mathcal{X}_{i,l,j} = x_{i+l+j-2} \quad i \in \overline{1:I}, l \in \overline{1:L}, j \in \overline{1:J}.$$

Определение 3.2 (Оператор вложения). Оператором вложения \mathcal{H}_L с длиной окна L будем называть отображение, переводящее временной ряд $\mathbf{X} = (x_1, x_2, \dots, x_N)$, $N \geq L$, в матрицу \mathbf{X} размерности $L \times K$, $K = N - L + 1$, такую, что $\mathbf{X}_{l,k} = x_{l+k-1}$.

Другими словами

$$\mathcal{H}_L(\mathbf{X}) = \begin{pmatrix} x_1 & x_2 & \dots & x_K \\ x_2 & x_3 & \dots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \dots & x_N \end{pmatrix}.$$

В терминах оператора вложения слои траекторного тензора ряда \mathbf{X} с параметрами I, L имеют следующий вид

$$\begin{aligned} \mathcal{X}_{:,j} &= \mathcal{H}_I((x_j, x_{j+1}, \dots, x_{j+I+L-2})), \\ \mathcal{X}_{:,l} &= \mathcal{H}_I((x_l, x_{l+1}, \dots, x_{l+L+J-2})), \\ \mathcal{X}_{:,i} &= \mathcal{H}_L((x_i, x_{i+1}, \dots, x_{i+L+J-2})). \end{aligned}$$

На вход алгоритму подаётся временной ряд \mathbf{X} и параметры $I, L : 1 \leq I, L \leq N$, $I + L \leq N + 1$. Так как при замене одного из этих параметров на $J = N - I - L + 2$ или при замене их между собой получаются те же самые траекторные тензоры с точностью до перестановки их измерений, то имеет смысл при рассмотрении нескольких наборов параметров рассматривать только те, которые дают уникальные тройки (I, L, J) . В зависимости от целей определяются разные формулировки алгоритма.

HOSVD-SSA для отделения различных компонент ряда друг от друга представлен в алгоритме 1

Алгоритм HO-SSA для выделения в ряде сигнала из шума сводится к получению как можно более точного приближения траекторного тензора тензором меньшего, заданного пользователем, ранга, и может быть проведён двумя различными способами.

Первый способ заключается в приближении траекторного тензора путём усечения его HOSVD (HOSVD SSA). Благодаря свойствам 2.7, 2.6 такое приближение можно считать достаточно точным, хоть оно и не оптимально. Первые два шага этого алгоритма совпадают с алгоритмом для отделения компонент ряда, поэтому опишем его, начиная с третьего шага. Описание приведено в алгоритме 2.

Алгоритм 1 HOSVD-SSA для отделения компонент ряда.

Входные данные: \mathbf{X} , $I, L : 1 \leq I, L \leq N$, $I + L \leq N + 1$, где N — длина \mathbf{X} , m , $\mathfrak{S}_1, \dots, \mathfrak{S}_m$:

$$\mathfrak{S} = \bigcup_{k=1}^m \mathfrak{S}_k \quad \mathfrak{S}_k \cap \mathfrak{S}_l = \emptyset, k \neq l,$$

где $\mathfrak{S} = \{1, 2, \dots, \min(I, L, J)\}$.

Результат: $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(m)}$:

$$\mathbf{X} = \sum_{k=1}^m \mathbf{X}^{(k)}.$$

- 1: Построение траекторного тензора \mathcal{X} по параметрам I, L .
- 2: Проведение HOSVD траекторного тензора \mathcal{X} , получение его представления в виде (3)

$$\mathcal{X} = \sum_{i=1}^I \sum_{l=1}^L \sum_{j=1}^J \mathcal{Z}_{i,l,j} \mathbf{U}_i^{(1)} \circ \mathbf{U}_l^{(2)} \circ \mathbf{U}_j^{(3)}. \quad (7)$$

- 3: Группировка: построение тензоров

$$\mathcal{X}^{(\mathfrak{S}_k)} = \sum_{i \in \mathfrak{S}_k} \sum_{l \in \mathfrak{S}_k} \sum_{j \in \mathfrak{S}_k} \mathcal{Z}_{i,l,j} \mathbf{U}_i^{(1)} \circ \mathbf{U}_l^{(2)} \circ \mathbf{U}_j^{(3)}.$$

- 4: Восстановление рядов $\mathbf{X}^{(k)} = \mathbf{X}^{(\mathfrak{S}_k)}$ по тензорам $\mathcal{X}^{(\mathfrak{S}_k)}$ посредством их усреднения вдоль плоскостей $i + l + j = \text{const}$:

$$x_n^{(k)} = \frac{1}{\#\mathfrak{M}_n} \sum_{(i,l,j) \in \mathfrak{M}_n} \mathcal{X}_{i,l,j}^{(\mathfrak{S}_k)}, \quad n \in \overline{1:N},$$

$$\mathfrak{M}_n = \{(i, l, j) | 1 \leq i \leq I, 1 \leq l \leq L, 1 \leq j \leq J, i + l + j - 2 = n\}.$$

Алгоритм 2 HOSVD SSA для выделения сигнала.

Входные данные: \mathbf{X} , $I, L : 1 \leq I, L \leq N$, $I + L \leq N + 1$, где N — длина \mathbf{X} , $R_1 \in \overline{1:I}$, $R_2 \in \overline{1:L}$, $R_3 \in \overline{1:J}$.

Результат: $\hat{\mathbf{X}}$.

- 3: По параметрам R_1, R_2, R_3 и разложению траекторного тензора \mathcal{X} в виде (7), в тензоре \mathcal{Z} проводится замена матриц-слоёв с номерами $k > R_m$ в соответствующих измерениях на нулевые, и по полученному тензору $\hat{\mathcal{Z}}$ строится приближение траекторного тензора $\hat{\mathcal{X}}$.
 - 4: Усреднение тензора $\hat{\mathcal{X}}$ вдоль плоскостей $i + l + j = \text{const}$, в результате чего получается оценка сигнала $\hat{\mathbf{X}}$.
-

Второй способ использует метод приближения тензора другим тензором с меньшими значениями n -рангов — High-Order Orthogonal Iteration (HOOI) [10]. При заданных тензоре \mathcal{A} и наборе n -рангов (R_1, R_2, \dots, R_M) , результатом метода будет тензор $\hat{\mathcal{A}}$, n -ранги которого совпадают с набором (R_1, R_2, \dots, R_M) , и который решает задачу минимизации

$$\|\mathcal{A} - \hat{\mathcal{A}}\| \rightarrow \min,$$

где минимум берётся по классу тензоров с заданными n -рангами. Исходя из определения, результат метода HOOI является оптимальным, в связи с чем его можно использовать для приближения траекторного тензора ряда.

Приведём вторую реализацию алгоритма HO-SSA для отделения сигнала от шума, используя HOOI (HOOI SSA). Первый шаг алгоритма совпадает с предыдущими алгоритмами, поэтому опишем его начиная со второго шага. Описание приведено в алгоритме 3.

Алгоритм 3 HOOI SSA

Входные данные: \mathbf{X} , $I, L : 1 \leq I, L \leq N$, $I + L \leq N + 1$, где N — длина \mathbf{X} , $R_1 \in \overline{1 : I}$, $R_2 \in \overline{1 : L}$, $R_3 \in \overline{1 : J}$.

Результат: $\hat{\mathbf{X}}$.

- 2: Применение к построенному на первом шаге траекторному тензору \mathcal{X} алгоритма HOOI с набором n -рангов (R_1, R_2, R_3) , в результате чего получается тензор $\hat{\mathcal{X}}$.
 - 3: Восстановление сигнала по тензору $\hat{\mathcal{X}}$ совпадает с четвёртым шагом в алгоритме HOSVD SSA для выделения в ряде сигнала из шума.
-

4. Свойства HO-SSA

В силу аналогичности свойств SVD и HOSVD, многие определения и свойства из теории SSA [1] можно перенести на тензорный случай.

4.1. Разделимость рядов в терминах HO-SSA

Утверждение 4.1. $\tilde{\mathbf{X}} = (\tilde{x}_1, \dots, \tilde{x}_N)$, $\hat{\mathbf{X}} = (\hat{x}_1, \dots, \hat{x}_N)$ — временные ряды длины N . Пусть ряд \mathbf{X} является суммой этих рядов. Траекторные тензоры рядов равны соответственно: $\tilde{\mathcal{X}}$, $\hat{\mathcal{X}}$, \mathcal{X} . Тогда существует сингулярное разложение тензора \mathcal{X} с параметрами I, L , которое можно представить в виде суммы сингулярных разложений

тензоров $\tilde{\mathcal{X}}$ и $\hat{\mathcal{X}}$ с теми же параметрами в том и только том случае, когда взаимно ортогональны все подряды рядов $\tilde{\mathcal{X}}$ и $\hat{\mathcal{X}}$ длины $I, L, J = N - I - L + 2$, то есть

1. $\tilde{x}_k \hat{x}_m + \dots + \tilde{x}_{k+I-1} \hat{x}_{m+I-1} = 0 \quad \forall k, m \in \overline{1 : N - I + 1},$
2. $\tilde{x}_k \hat{x}_m + \dots + \tilde{x}_{k+L-1} \hat{x}_{m+L-1} = 0 \quad \forall k, m \in \overline{1 : N - L + 1},$
3. $\tilde{x}_k \hat{x}_m + \dots + \tilde{x}_{k+J-1} \hat{x}_{m+J-1} = 0 \quad \forall k, m \in \overline{1 : N - J + 1}.$

Доказательство. Сингулярные разложения тензоров $\mathcal{X}, \tilde{\mathcal{X}}, \hat{\mathcal{X}}$ могут быть представлены в виде следующих сумм:

$$\begin{aligned} \mathcal{X} &= \sum_{i=1}^I \sum_{l=1}^L \sum_{j=1}^J \mathcal{Z}_{i,l,j} \mathbf{U}_i^{(1)} \circ \mathbf{U}_l^{(2)} \circ \mathbf{U}_j^{(3)}, \\ \tilde{\mathcal{X}} &= \sum_{i=1}^I \sum_{l=1}^L \sum_{j=1}^J \tilde{\mathcal{Z}}_{i,l,j} \tilde{\mathbf{U}}_i^{(1)} \circ \tilde{\mathbf{U}}_l^{(2)} \circ \tilde{\mathbf{U}}_j^{(3)}, \\ \hat{\mathcal{X}} &= \sum_{i=1}^I \sum_{l=1}^L \sum_{j=1}^J \hat{\mathcal{Z}}_{i,l,j} \hat{\mathbf{U}}_i^{(1)} \circ \hat{\mathbf{U}}_l^{(2)} \circ \hat{\mathbf{U}}_j^{(3)}. \end{aligned}$$

Сумма $\mathcal{X} = \sum_i \sum_l \sum_j \tilde{\mathcal{Z}}_{i,l,j} \tilde{\mathbf{U}}_i^{(1)} \circ \tilde{\mathbf{U}}_l^{(2)} \circ \tilde{\mathbf{U}}_j^{(3)} + \sum_i \sum_l \sum_j \hat{\mathcal{Z}}_{i,l,j} \hat{\mathbf{U}}_i^{(1)} \circ \hat{\mathbf{U}}_l^{(2)} \circ \hat{\mathbf{U}}_j^{(3)}$ является сингулярным разложением \mathcal{X} в том и только том случае, когда пары векторов $\tilde{\mathbf{U}}_k^{(\sigma)}, \hat{\mathbf{U}}_m^{(\sigma)}$ взаимно ортогональны при всех возможных значениях σ, k, m . Это равносильно ортогональности линейных пространств $\mathcal{L}_1^{(\sigma)}, \mathcal{L}_2^{(\sigma)}$, построенных на векторах $\tilde{\mathbf{U}}_k^{(\sigma)}$ и $\hat{\mathbf{U}}_m^{(\sigma)}$ соответственно.

Рассмотрим пространства $\mathcal{L}_1^{(1)}, \mathcal{L}_2^{(1)}$: это пространства первых измерений тензоров $\tilde{\mathcal{X}}$ и $\hat{\mathcal{X}}$, то есть пространства построенные на векторах вида $\tilde{\mathcal{X}}_{i,l,j}$ и $\hat{\mathcal{X}}_{i,l,j}$ соответственно. Вспоминая вид тензоров $\tilde{\mathcal{X}}$ и $\hat{\mathcal{X}}$ получаем, что условие ортогональности этих линейных пространств равносильно первому условию из формулировки утверждения.

Оставшиеся два условия получаются аналогично из условий ортогональности оставшихся двух пар линейных пространств. \square

Замечание. Условия утверждения выше сильнее, чем для аналогичного утверждения для стандартного SSA. Это позволяет предположить, что тензорный вариант может оказаться не лучше стандартного.

Из утверждения 4.1 следует, что понятие слабой разделимости ряда из теории SSA применимо и к тензорному случаю.

Следствие 4.1.1. Если временные ряды \tilde{X} и \hat{X} длины N слабо I - и L -разделимы в смысле теории SSA, то существует такое HOSVD траекторного тензора \mathcal{X} ряда $X = \tilde{X} + \hat{X}$, что его можно разбить на две части, являющиеся HOSVD траекторных тензоров, составленных по рядам \tilde{X} и \hat{X} .

4.2. Примеры разделимости рядов в тензорном случае

Рассмотрим условия разделимости рядов $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N)$, $\hat{X} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N)$ в некоторых частных случаях.

- Отделимость от константного ряда

Пусть $\tilde{x}_n = c \neq 0$ для $n \in \overline{1:N}$. Тогда необходимые и достаточные условия отделимости от него ряда \hat{X} в смысле HO-SSA следующие:

1. Ряд \hat{X} имеет целый период T , и I/T , L/T , J/T — целые;
2. $\hat{x}_1 + \hat{x}_2 + \dots + \hat{x}_T = 0$.

Пример 4.1. Ряд с элементами вида $\tilde{x}_n = \cos(2\pi n/T + \varphi)$ длины N такой, что $N+2$ делится нацело на T , будет слабо отделим от константного ряда при выборе параметров $I, L : I + L < N + 1$, делящихся нацело на T .

- Отделимость от экспоненциального ряда

Пусть $\tilde{x}_n = e^{\alpha n}$ для $n \in \overline{1:N}$. Тогда необходимые и достаточные условия отделимости от него ряда \hat{X} в смысле HO-SSA следующие:

1. Ряд $(\tilde{x}_1\hat{x}_1, \tilde{x}_2\hat{x}_2, \dots, \tilde{x}_N\hat{x}_N)$ имеет целый период T , и I/T , L/T , J/T — целые;
2. $\tilde{x}_1\hat{x}_1 + \tilde{x}_2\hat{x}_2 + \dots + \tilde{x}_N\hat{x}_N = 0$.

Пример 4.2. Ряд с элементами вида $\tilde{x}_n = e^{-\alpha n} \cos(2\pi n/T + \varphi)$ длины N такой, что $N+2$ делится нацело на T , будет слабо отделим от ряда с элементами вида $\hat{x}_n = e^{\alpha n}$ при выборе параметров $I, L : I + L < N + 1$, делящихся нацело на T .

- Отделимость от гармонического ряда

Пусть $\tilde{x}_n = \cos(2\pi\omega n + \varphi)$, где $0 < \omega < 1/2$, и $I, L, J > 2$. Положим $\hat{x}_n = \cos(2\pi\omega' n + \varphi')$, тогда ряд \tilde{X} отделим от ряда \hat{X} в смысле HO-SSA тогда и только тогда, когда $\omega \neq \omega'$ и $I\omega, I\omega', L\omega, L\omega', J\omega, J\omega'$ — целые числа.

Замечание. Результаты выше приведены в случае точной разделимости компонент в смысле теории SSA. В случае, когда точной разделимости нет, а есть только приближённая, в тензорном случае возникает такая ситуация, когда одни и те же одноранговые тензоры в HOSVD траекторного тензора могут быть отнесены сразу к нескольким компонентам. На данный момент это является нерешённой проблемой.

4.3. Ранг ряда в терминах HO-SSA

Утверждение 4.2. Пусть временной ряд X имеет конечный ранг d в терминах SSA. Тогда для любых значений параметров I и L таких, что

$$d \leq \min(I, L, N - I - L + 2),$$

количество ненулевых сингулярных чисел по каждому измерению в HOSVD траекторного тензора \mathcal{X} , построенного по этому ряду с параметрами I и L , будет равно d .

Это утверждение является прямым следствием определения ранга ряда и свойства 2.4 HOSVD.

Следствие 4.2.1. Понятие ранга ряда имеет тот же смысл в терминах HO-SSA, что и в стандартной теории SSA, причём ряды конечного ранга имеют одинаковые ранги в тензорном и стандартном случаях.

5. Примеры использования HO-SSA

Рассмотрим несколько примеров использования HO-SSA для анализа временных рядов.

Пример 5.1 (Разделимость синуса и константы). Рассмотрим ряд с элементами $x_n = 3 + \sin(2\pi n/3 + \pi/3)$, где $n \in \overline{0:15}$. После построения траекторного тензора \mathcal{X} с параметрами $I = L = 6$ и его разложения получаем тензор сингулярных чисел \mathcal{Z} и матрицы сингулярных векторов $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, $\mathbf{U}^{(3)}$. Так как все размерности траекторного тензора \mathcal{X} равны, его развёртки по всем измерениям совпадают, а значит совпадают и матрицы

сингулярных векторов $\mathbf{U}^{(1)} = \mathbf{U}^{(2)} = \mathbf{U}^{(3)} = \mathbf{U}$.

$$\mathbf{U} = \begin{pmatrix} -0.41 & 0.00 & 0.58 & 0.70 & -0.10 & 0.01 \\ -0.41 & 0.50 & -0.29 & 0.08 & 0.62 & 0.33 \\ -0.41 & -0.50 & -0.29 & 0.06 & 0.33 & -0.63 \\ -0.41 & -0.00 & 0.58 & -0.70 & 0.10 & -0.01 \\ -0.41 & 0.50 & -0.29 & -0.08 & -0.62 & -0.33 \\ -0.41 & -0.50 & -0.29 & -0.06 & -0.33 & 0.63 \end{pmatrix},$$

$$\mathcal{Z}_{1,1,1} = -44.09,$$

$$-\mathcal{Z}_{2,2,2} = \mathcal{Z}_{3,3,2} = \mathcal{Z}_{2,3,3} = \mathcal{Z}_{3,2,3} = 2.60,$$

$$\mathcal{Z}_{2,3,2} = \mathcal{Z}_{3,2,2} = \mathcal{Z}_{2,2,3} = -\mathcal{Z}_{3,3,3} = -4.50,$$

$$\mathcal{Z}_{i,l,j} = 0 \text{ для всех остальных значений } i, l, j.$$

Видно, что первый сингулярный вектор постоянен, а второй и третий — периодические с периодом 3. Кроме того, по каждому из трёх измерений количество ненулевых сингулярных чисел равно 3 (например $\|\mathcal{Z}_{:,1}\| > \|\mathcal{Z}_{:,2}\| = \|\mathcal{Z}_{:,3}\| > 0$, $\|\mathcal{Z}_{:,j}\| = 0$ для всех остальных j). Исходя из этого, имеет смысл отнести индекс $\{1\}$ к константной компоненте ряда, индексы $\{2, 3\}$ — к гармонической (синус), а остальные проигнорировать. После восстановления тензоров, полученных такой группировкой, получаем два ряда

$$\hat{\mathbf{X}} = (3, 3, \dots, 3),$$

$$\tilde{\mathbf{X}} = (0.86, 0, -0.86, 0.86, \dots, 0, -0.86, 0.86).$$

Таким образом, константный ряд отделился от синуса.

Пример 5.2 (Смещение двух косинусов). Рассмотрим ряд с элементами $x_n = \cos(2\pi n/3) + \cos(2\pi n/4)$, $n \in \overline{0 : 33}$. Выбрав параметры $I = L = 12$, после разложения получаем тензор сингулярных значений \mathcal{Z} и, в силу равенства размерностей траекторного тензора, равные между собой матрицы сингулярных векторов $\mathbf{U}^{(1)} = \mathbf{U}^{(2)} = \mathbf{U}^{(3)} = \mathbf{U}$. Тензор \mathcal{Z} имеет вид тензорного блока \mathcal{Z}' размерности $4 \times 4 \times 4$, окаймлённого нулями, в котором уже нельзя выделить блочно-диагональную структуру. Если рассмотреть матрицы сингулярных векторов, можно увидеть, что никакой сингулярный вектор не имеет периода

равного 3 или 4:

$$\mathbf{U} = \begin{pmatrix} 0 & 0 & -0.58 & 0 & \dots \\ -0.18 & 0.36 & 0.14 & -0.39 & \dots \\ -0.17 & -0.16 & 0.43 & 0.30 & \dots \\ 0.38 & -0.04 & -0.29 & 0.32 & \dots \\ 0.14 & 0.002 & -0.14 & -0.54 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Таким образом, произошло смешение двух косинусов одинаковой амплитуды.

Пример 5.3 (Косинусы одной амплитуды). Рассмотрим ряд с элементами $x_n = 2 \cos(2\pi n/2) + 2 \cos(2\pi n/3 + \pi/2)$, $n \in \overline{1:7}$. Выбрав параметры $I = L = 3$, после разложения получаем тензор сингулярных значений \mathcal{Z} , развёртка которого по первому измерению имеет вид

$$[\mathcal{Z}]_1 = \left(\begin{array}{ccc|ccc|ccc} 8.78 & 0.42 & -0.30 & 0.42 & 1.46 & 1.01 & -0.30 & 1.01 & -0.31 \\ 0.42 & 1.46 & 1.01 & 1.46 & -3.78 & 0.67 & 1.01 & 0.67 & 0.30 \\ -0.30 & 1.01 & -0.31 & 1.01 & 0.67 & 0.30 & -0.31 & 0.30 & -0.15 \end{array} \right).$$

С другой стороны, такой вид может иметь тензор сингулярных значений траекторного тензора ряда, имеющего ранг 3, но не раскладывающегося в сумму рядов меньшего ранга, что затрудняет отнесение конкретных компонент разложения к тем или иным компонентам сигнала.

5.1. Сравнение HO-SSA и SSA

Пусть временной ряд \mathbf{X} имеет вид

$$\begin{aligned} \mathbf{X} &= (x_1, x_2, \dots, x_N), \\ x_n &= s_n + \varepsilon_n, \quad n \in \overline{1:N}, \end{aligned} \tag{8}$$

где $\mathbf{S} = (s_1, s_2, \dots, s_N)$ — сигнал, $\mathcal{E} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)$ — шум. Рассмотрим проблему отделения сигнала от шума на примерах различных рядов.

Пример 5.4 (Выделение экспоненты). Пусть $N = 23$, ε_n — независимые $\mathbf{N}(0, 2.25)$ величины,

$$s_n = 2e^{0.035n}. \tag{9}$$

Таблица 1. SSA: RMSE оценки экспоненциального сигнала (9).

| L | 4 | 6 | 12 |
|-----|------|------|-------------|
| MSE | 0.65 | 0.55 | 0.52 |

Таблица 2. HO-SSA: RMSE оценки экспоненциального сигнала (9).

| Метод приближения \ $I \times L$ | | | | | |
|----------------------------------|------|-------------|--------------|-------------|------|
| | 4×4 | 4×6 | 4×12 | 6×6 | 6×12 |
| Усечение HOSVD | 0.59 | 0.56 | 0.56 | 0.56 | 0.57 |
| HOOI | 0.58 | 0.55 | 0.547 | 0.56 | 0.56 |

В таблицах 1, 2, приведены значения отклонения восстановленного ряда от исходного ряда для различных значений параметров после использования SSA и двух вариаций HO-SSA для выделения сигнала. RMSE здесь и далее высчитывается по 500 реализациям шума, если не указано иное. кроме того, здесь и далее жирным шрифтом выделены минимальные по строке значения RMSE, причём отличие этих минимальных значений от остальных в строке значимо при уровне значимости $\alpha = 0.05$.

Пример 5.5 (Выделение косинуса). Пусть $N = 71$, рассматриваются три варианта шума: белый гауссовский с параметром $\sigma^2 = 25$ и красный гауссовский с параметрами $\delta = \sqrt{5}$, $\varphi = 0.5$ и $\delta = \sqrt{5}$, $\varphi = 0.9$.

$$s_n = 30 \cos(2\pi n/12). \quad (10)$$

В таблицах 3, 4, 5 приведены значения отклонения восстановленного ряда от исход-

Таблица 3. SSA: RMSE оценки синусоидального сигнала (10).

| вид шума \ L | | | | |
|------------------------------|------|------|-------------|------|
| | 12 | 24 | 30 | 36 |
| белый шум, $\sigma^2 = 25$ | 1.82 | 1.42 | 1.40 | 1.42 |
| красный шум, $\varphi = 0.5$ | 1.31 | 1.03 | 1.01 | 1.03 |
| красный шум, $\varphi = 0.9$ | 1.88 | 1.37 | 1.34 | 1.36 |

ного ряда для различных вариантов шума и различных значений параметров после использования SSA и HO-SSA.

Таблица 4. HOSVD SSA: RMSE оценки синусоидального сигнала (10).

| $I \times L$ вид шума | 12×12 | 12×24 | 12×30 | 24×24 | 24×30 | 30×36 |
|------------------------------|-------|-------------|-------|-------|-------|-------------|
| белый шум, $\sigma^2 = 25$ | 1.64 | 1.53 | 1.57 | 1.66 | 1.62 | 1.49 |
| красный шум, $\varphi = 0.5$ | 1.18 | 1.12 | 1.14 | 1.21 | 1.19 | 1.08 |
| красный шум, $\varphi = 0.9$ | 1.58 | 1.44 | 1.47 | 1.57 | 1.54 | 1.46 |

Таблица 5. HOOI SSA: RMSE оценки синусоидального сигнала (10).

| $I \times L$ вид шума | 12×12 | 12×24 | 12×30 | 24×24 | 24×30 | 30×36 |
|------------------------------|-------|-------|-------|-------|-------|-------------|
| белый шум, $\sigma^2 = 25$ | 1.63 | 1.53 | 1.56 | 1.65 | 1.62 | 1.49 |
| красный шум, $\varphi = 0.5$ | 1.17 | 1.12 | 1.14 | 1.21 | 1.19 | 1.08 |
| красный шум, $\varphi = 0.9$ | 1.56 | 1.42 | 1.44 | 1.54 | 1.51 | 1.39 |

Пример 5.6 (Выделение линейного ряда). Пусть $N = 39$, рассматриваются два варианта белого гауссовского шума: с параметром $\sigma^2 = 2.25$ и с параметром $\sigma^2 = 0.04$.

$$s_n = 2 + 0.1n. \quad (11)$$

В каждом из случаев шума будем рассматривать по два варианта восстановления линейного сигнала: по одной компоненте и по двум. В таблицах 6, 7, 8, 9 приведены

Таблица 6. SSA: RMSE оценки линейного сигнала (11), случай $\sigma^2 = 2.25$.

| L Компонент | 10 | 15 | 20 |
|------------------|------|-------------|-------------|
| 1 | 0.44 | 0.42 | 0.42 |
| 2 | 0.70 | 0.66 | 0.68 |

значения отклонения восстановленного линейного ряда от исходного. Таблицам 6, 7 соответствует гауссовский шум с параметром $\sigma^2 = 2.25$, таблицам 8, 9 — гауссовский шум с параметром $\sigma^2 = 0.04$.

Пример 5.7 (Случай, когда HO-SSA срабатывает точнее, чем SSA). Пусть $N = 9$,

Таблица 7. HO-SSA: RMSE оценки линейного сигнала (11), случай $\sigma^2 = 2.25$.

| Компонент | $I \times L$ | 10×10 | 10×15 | 10×20 | 15×15 | 15×20 |
|-----------|----------------------|-------|-------|-------|-------|-------------|
| | Метод восстановления | | | | | |
| 1 | Усечение HOSVD | 0.47 | 0.49 | 0.48 | 0.49 | 0.46 |
| | HOOI | 0.47 | 0.48 | 0.47 | 0.49 | 0.45 |
| 2 | Усечение HOSVD | 0.59 | 0.60 | 0.58 | 0.61 | 0.57 |
| | HOOI | 0.62 | 0.64 | 0.62 | 0.64 | 0.61 |

Таблица 8. SSA: RMSE оценки линейного сигнала (11), случай $\sigma^2 = 0.04$.

| Компонент | L | 10 | 15 | 20 |
|-----------|-----|-------------|--------------|--------------|
| | | | | |
| 1 | | 0.09 | 0.108 | 0.118 |
| 2 | | 0.10 | 0.093 | 0.093 |

Таблица 9. HO-SSA: RMSE оценки линейного сигнала (11), случай $\sigma^2 = 0.04$.

| Компонент | $I \times L$ | 10×10 | 10×15 | 10×20 | 15×15 | 15×20 |
|-----------|----------------------|-------|-------|-------|-------|--------------|
| | Метод восстановления | | | | | |
| 1 | HOSVD SSA | 0.133 | 0.145 | 0.136 | 0.147 | 0.130 |
| | HOOI SSA | 0.132 | 0.144 | 0.136 | 0.146 | 0.130 |
| 2 | HOSVD SSA | 0.123 | 0.130 | 0.125 | 0.133 | 0.112 |
| | HOOI SSA | 0.123 | 0.130 | 0.124 | 0.133 | 0.114 |

шум — красный гауссовский с параметрами $\delta = 0.1$, $\varphi = 0.9$.

$$s_n = \sin(2\pi n/3 + \pi/2). \quad (12)$$

В таблице 10 приведены результаты измерения средних отклонений восстановленного ряда от исходного для разных методов по 1000 реализаций шума.

Таблица 10. Сравнение SSA и HO-SSA: RMSE оценки короткого синусоидального сигнала (12).

| SSA | HOSVD SSA | HOOI SSA |
|-------|-----------|--------------|
| 0.116 | 0.110 | 0.095 |

Причины того, что в этом примере обе вариации HO-SSA дают стабильно лучшие

результаты, чем SSA, пока неизвестны и остаются для дальнейшего изучения.

6. Описание метода HOSVD MSSA для выделения сигнала

Пусть дан P -мерный временной ряд \mathbf{X} длины N

$$\mathbf{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(P)})^T,$$

$$\mathbf{X}^{(i)} = (x_1^{(i)}, \dots, x_N^{(i)}), i \in \overline{1 : N}.$$

Определение 6.1 (Траекторный тензор многомерного ряда). Траекторным тензором ряда \mathbf{X} с длиной окна $L : 1 \leq L \leq N$ будем называть тензор \mathcal{X} размерности $L \times K \times P$, $K = N - L + 1$, элементы которого удовлетворяют равенству

$$\mathcal{X}_{l,k,p} = x_{l+k-1}^{(p)} \quad l \in \overline{1 : L}, k \in \overline{1 : K}, p \in \overline{1 : P}.$$

Из определения следует, что p -й слой вдоль третьего измерения траекторного тензора \mathcal{X} с длиной окна L является траекторной матрицей ряда $\mathbf{X}^{(p)}$, построенной по длине окна L . Пользуясь определением 3.2 оператора вложения, можно записать следующее представление

$$\mathcal{X}_{:,p} = \mathcal{H}_L(\mathbf{X}^{(p)}).$$

Метод HOSVD MSSA для выделения в ряде сигнала из шума, по аналогии с алгоритмом HO-SSA, сводится к получению как можно более точного приближения траекторного тензора тензором меньших, заданных пользователем, n -рангов. Для получения такого приближения используется усечение HOSVD траекторного тензора. Описание метода приведено в алгоритме 4.

7. Свойства HOSVD MSSA

Пусть $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(P)}$ — временные ряды длины N , $\mathbf{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(P)})^T$ — многомерный временной ряд длины N , \mathbf{H}_p — траекторные матрицы рядов $\mathbf{X}^{(p)}$ с длиной окна $L < N$, $p \in \overline{1 : P}$. Обозначим $K = N - L + 1$.

Построим траекторную матрицу многомерного ряда \mathbf{X} : $\mathbf{X} = [\mathbf{H}_1 : \mathbf{H}_2 : \dots : \mathbf{H}_P] \in \mathbb{R}^{L \times KP}$, её SVD имеет вид

$$\mathbf{X} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T.$$

Алгоритм 4 HOSVD MSSA

Входные данные: $\mathbf{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(P)})^T$, $L : 1 \leq L \leq N$, где N — длина \mathbf{X} , $R_1 \in \overline{1:L}$, $R_2 \in \overline{1:K}$, $R_3 \in \overline{1:P}$, где $K = N - L + 1$.

Результат: $\tilde{\mathbf{X}}$.

- 1: Построение по P -мерному временному ряду \mathbf{X} траекторного тензора \mathcal{X} с длиной окна L .
- 2: проведение HOSVD траекторного тензора \mathcal{X} , получение его представления в виде (3)

$$\mathcal{X} = \sum_{l=1}^L \sum_{k=1}^K \sum_{p=1}^P \mathcal{Z}_{l,k,p} \mathbf{U}_l^{(1)} \circ \mathbf{U}_k^{(2)} \circ \mathbf{U}_p^{(3)}. \quad (13)$$

- 3: Построение по параметрам R_1, R_2, R_3 усечённого тензора

$$\tilde{\mathcal{X}} = \sum_{l=1}^{R_1} \sum_{k=1}^{R_2} \sum_{p=1}^{R_3} \mathcal{Z}_{l,k,p} \mathbf{U}_l^{(1)} \circ \mathbf{U}_k^{(2)} \circ \mathbf{U}_p^{(3)}.$$

- 4: Восстановление многомерного ряда $\tilde{\mathbf{X}} = (\tilde{\mathbf{X}}^{(1)}, \dots, \tilde{\mathbf{X}}^{(P)})$ по тензору $\tilde{\mathcal{X}}$, которое происходит следующим образом: ряды $\tilde{\mathbf{X}}^{(p)}$ получаются усреднением матриц-слоёв $\tilde{\mathcal{X}}_{:,p}$ вдоль псевдодиагоналей $l + k = \text{const}$.
-

Теперь построим траекторный тензор $\mathcal{X} \in \mathbb{R}^{L \times K \times P}$ этого многомерного ряда с длиной окна L , как описано в определении 6.1: $\mathcal{X}_{:,p} = \mathbf{H}_p$, $p \in \overline{1:P}$, его HOSVD имеет вид

$$\mathcal{X} = \mathcal{Z} \times_1 \hat{\mathbf{U}}_1 \times_2 \hat{\mathbf{U}}_2 \times_3 \hat{\mathbf{U}}_3. \quad (14)$$

Утверждение 7.1. *Существуют такие SVD матрицы \mathbf{X} и HOSVD тензора \mathcal{X} разложения, что $\mathbf{U} = \hat{\mathbf{U}}_1$.*

Доказательство. Из свойства 2.1 известно, что в качестве $\hat{\mathbf{U}}_1$ можно выбрать матрицу $\hat{\mathbf{U}}$ левых сингулярных векторов из SVD развёртки тензора \mathcal{X} по первому измерению $[\mathcal{X}]_1$. В свою очередь, $[\mathcal{X}]_1 = [\mathbf{H}_1 : \mathbf{H}_2 : \dots : \mathbf{H}_P] = \mathbf{X}$. \square

Следствие 7.1.1. *Из этого утверждения и из того, что разложения SVD и HOSVD единственны с точностью до, возможно, некоторых ортогональных преобразований матриц сингулярных векторов, следует, что пространства, порождаемые левыми сингулярными векторами матрицы \mathbf{X} и сингулярными векторами первого измерения тензора \mathcal{X} совпадают.*

Замечание. При вычислении HOSVD тензора используется алгоритм, который последовательно вычисляет SVD развёрток этого тензора по каждому измерению, и получившиеся матрицы левых сингулярных векторов использует в качестве матриц сингулярных векторов соответствующего измерения в представлении (14). Таким образом, на практике совпадают не только пространства, порождаемые описанными выше сингулярными векторами, но и матрицы, составленные из этих векторов.

В теории MSSA для выделения сигнала из ряда с шумом важным понятием является понятие ранга ряда.

Определение 7.1 (Ранг многомерного ряда в теории MSSA). Число d называется рангом P -мерного ряда \mathbf{X} , если для любой длины окна $L : d \leq \min(L, N - L + 1)$ ранг траекторной матрицы \mathbf{X} , построенной по этой длине окна, будет равен d .

Таким образом, исходя из утверждения 7.1, ранг первого измерения траекторного тензора \mathcal{X} , построенного по многомерному ряду \mathbf{X} с длиной окна $L : d \leq \min(L, N - L + 1)$, будет равен рангу d этого ряда из теории MSSA. Кроме того, для таких L ранг второго измерения этого траекторного тензора тоже будет равен d , так как $\text{rank}_2 \mathcal{X} = \text{rank}[\mathcal{X}]_2$, а матрица $[\mathcal{X}]_2$ по построению совпадает с траекторной матрицей ряда \mathbf{X} , построенной

по длине окна $K = N - L - 1$. Однако, ранг третьего измерения траекторного тензора \mathcal{X} не связан с рангом ряда и имеет смысл структурного различия рядов между собой, и может быть посчитан как ранг матрицы, составленной из одномерных рядов, образующих данный многомерный ряд.

Замечание. По построению траекторного тензора набор векторов вдоль третьего измерения этого тензора не зависит от выбора длины окна L , поэтому и 3-ранг траекторного тензора не зависит от выбора длины окна.

Пример 7.1. Пусть двумерный временной ряд $\mathbf{X} = (\mathbf{X}^{(1)}, \mathbf{X}^{(2)})^T$ состоит из временных рядов следующего вида:

$$\begin{aligned}\mathbf{X}^{(m)} &= (x_1^{(m)}, x_2^{(m)}, \dots, x_N^{(m)}), \\ x_n^{(m)} &= a_m \sin(2\pi n \omega_m + \psi_m),\end{aligned}$$

где $m \in \{1, 2\}$, $n \in \overline{1:N}$, $a_m \neq 0$, $0 < \omega_m < 1/2$, $0 \leq \psi < 2\pi$. Обозначим за r ранг ряда \mathbf{X} , а за r_3 ранг третьего измерения этого ряда. В таблице 11 представлены значения рангов этого ряда в терминах HOSVD MSSA в зависимости от различных значений параметров ψ_m и ω_m .

Таблица 11. Ранги рядов в зависимости от параметров.

| Параметры | $\psi_1 = \psi_2$ | $\psi_1 \neq \psi_2$ | $\omega_1 \neq \omega_2$ |
|-----------|-----------------------|--------------------------|--------------------------|
| | $\omega_1 = \omega_2$ | $\omega_1 \neq \omega_2$ | |
| r | 2 | 2 | 4 |
| r_3 | 1 | 2 | 2 |

Утверждение 7.2. (О симметричности относительно замены длины окна) Пусть дан P -мерный временной ряд \mathbf{X} длины N и выбрана некоторая длина окна L , \mathcal{X} — траекторный тензор этого ряда, построенный по длине окна L , а \mathcal{Y} — по длине окна $K = N - L + 1$, и пусть R_1, R_2, R_3 — параметры, выбранные в третьем шаге алгоритма HOSVD MSSA, соответствующие \mathcal{X} , а R'_1, R'_2, R'_3 — соответствующие \mathcal{Y} . Тогда если $R_1 = R'_1, R_2 = R'_2, R_3 = R'_3$, то оценки сигнала $\tilde{\mathbf{X}}$ и $\tilde{\mathbf{Y}}$, построенные по \mathcal{X} и \mathcal{Y} соответственно, совпадут.

Доказательство. По построению $\mathcal{X}_{l,k,p} = x_{l+k-1}^{(p)} = x_{k+l-1}^{(p)} = \mathcal{Y}_{k,l,p}$, где $l \in \overline{1:L}$, $k \in \overline{1:K}$, $p \in \overline{1:P}$. Это означает, что $[\mathcal{X}]_1 = [\mathcal{Y}]_2$ и $[\mathcal{X}]_2 = [\mathcal{Y}]_1$, а $[\mathcal{X}]_3 = [\mathcal{Y}]_3$. Пусть усечение

HOSVD \mathcal{X} , описанное в третьем шаге алгоритма, имеет вид

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

тогда усечение HOSVD \mathcal{Y} будет иметь вид

$$\tilde{\mathcal{Y}} = \sum_{l=1}^{R'_1} \sum_{k=1}^{R'_2} \sum_{p=1}^{R'_3} \mathcal{Z}_{l,k,p} \mathbf{U}_l^{(2)} \circ \mathbf{U}_k^{(1)} \circ \mathbf{U}_p^{(3)} = \sum_{k=1}^{R_2} \sum_{l=1}^{R_1} \sum_{p=1}^{R_3} \mathcal{Z}_{k,l,p} \mathbf{U}_k^{(2)} \circ \mathbf{U}_l^{(1)} \circ \mathbf{U}_p^{(3)},$$

Отсюда следует, что $[\tilde{\mathcal{X}}]_1 = [\tilde{\mathcal{Y}}]_2$, $[\tilde{\mathcal{X}}]_2 = [\tilde{\mathcal{Y}}]_1$ и $[\tilde{\mathcal{X}}]_3 = [\tilde{\mathcal{Y}}]_3$, а значит выполнено равенство $\tilde{\mathcal{X}}_{l,k,p} = \tilde{\mathcal{Y}}_{k,l,p}$ или по другому $\tilde{\mathcal{X}}_{:,p} = \tilde{\mathcal{Y}}_{:,p}^T$, следовательно усреднения матриц $\tilde{\mathcal{X}}_{:,p}$ и $\tilde{\mathcal{Y}}_{:,p}$ вдоль псевдодиагоналей совпадут для всех p . \square

8. Численные сравнения полученных методов с известными

В этой части представлены сравнения методов SSA, MSSA, 2D-SSA [9], HOOI SSA и HOSVD MSSA по точности восстановления сигнала из многомерных рядов вида

$$\begin{aligned} \mathbf{X} &= (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(P)}), \\ \mathbf{X}^{(p)} &= (s_1^{(p)} + \varepsilon_1^{(p)}, \dots, s_N^{(p)} + \varepsilon_N^{(p)}), \\ s_n^{(p)} &= a_p \cos(2\pi n \omega_p + \psi_p), \end{aligned} \tag{15}$$

где $\varepsilon_n^{(p)}$ — независимые $N(0, 25)$ величины (белый гауссовский шум), а $s_n^{(p)}$ — элементы искомого сигнала. В качестве оценки точности восстановления сигнала было взято средне-квадратичное отклонение (RMSE) восстановленного сигнала от истинного его значения по 500 реализациям шума. Под применением одномерных методов (SSA и HOOI SSA) для выделения многомерного сигнала \mathbf{X} понимается применение этих методов к каждому из рядов $\mathbf{X}^{(p)}$ отдельно.

8.1. Принцип выбора числа компонент, относимых к сигналу

В каждом из методов семейства SSA важным этапом является выбор числа компонент разложения траекторной матрицы или тензора ряда, по которым будет восстанавливаться сигнал. В теории SSA утверждается, что в общем случае для получения наибольшей точности восстановления сигнала нужно брать число компонент, равное рангу этого сигнала в терминах конкретного метода. Приведём выражения, позволяющие находить ранги сигналов вида (15), для каждого рассматриваемого метода. Ранг

сигнала в терминах метода Т будем обозначать r_T , например: r_{SSA} — ранг сигнала в терминах SSA, r_{MSSA} — в терминах MSSA и т.д.

Тензорные алгоритмы выделения сигнала сводятся к приближению траекторного тензора некоторым тензором (R_1, R_2, R_3) -ранга. Утверждения 4.2 и 7.1 позволяют свести задачу выбора R_1, R_2, R_3 к нахождению r_{SSA} в одномерном случае и r_{MSSA} и r_3 в многомерном, где r_3 — ранг матрицы, составленной из одномерных сигналов, образующих данный многомерный сигнал. Тогда рекомендуемым выбором (R_1, R_2, R_3) в алгоритме HO-SSA будет $R_1 = R_2 = R_3 = r_{\text{SSA}}$, а в алгоритме HOSVD-MSSA — $R_1 = R_2 = r_{\text{MSSA}}$, $R_3 = r_3$.

SSA В терминах SSA любой одномерный сигнал вида $s_n = a \cos(2\pi n\omega + \psi)$, где $a \neq 0$, $0 \leq \psi \leq 2\pi$ и $0 \leq \omega \leq 1/2$ имеет ранг $r_{\text{SSA}} = 2$, если $0 < \omega < 1/2$ и $r_{\text{SSA}} = 1$ при $\omega \in \{0, 1/2\}$ и $\psi \notin \{\pi/2, 3\pi/2\}$.

MSSA Пусть $\omega \in [0, 1/2]$, обозначим

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

тогда MSSA-ранг многомерного сигнала вида (15) равен

$$r_{\text{MSSA}} = \sum_{\omega \in \Omega} r(\omega),$$

где Ω — множество различных частот ω_p .

2D-SSA Элементы многомерного сигнала вида (15) можно представить в виде суммы

$$s_n^{(p)} = \sum_{m=1}^d A_m(p, n),$$

$$A_m(p, n) = \begin{pmatrix} \cos(2\pi p\omega'_m) \\ \sin(2\pi p\omega'_m) \end{pmatrix}^T \begin{pmatrix} a_m & b_m \\ c_m & d_m \end{pmatrix} \begin{pmatrix} \cos(2\pi n\omega''_m) \\ \sin(2\pi n\omega''_m) \end{pmatrix},$$

где $(\omega'_m, \omega''_m) \neq (\omega'_j, \omega''_j)$ для $m \neq j$ и $\omega'_m, \omega''_m \in (0, 1/2)$. Тогда ранг данного многомерного сигнала в терминах 2D-SSA равен сумме

$$r_{\text{2D-SSA}} = \sum_{m=1}^d \nu_m,$$

где

$$\nu_m = 2 \operatorname{rank} \begin{pmatrix} a_m & b_m & c_m & d_m \\ d_m & -c_m & -b_m & a_m \end{pmatrix}.$$

8.2. Численные сравнения

Рассмотрим некоторые варианты сигнала вида (15). В таблицах 12, 14, 16 приведены значения рангов сигналов в каждом из методов в зависимости от значения параметров ω_p , ψ_p для случаев $P = 2$, $P = 5$ и $P = 9$ соответственно. Для одномерных методов (SSA и HOOI SSA) указаны ранги каждого одномерного сигнала $S^{(p)}$.

В таблицах 13, 15, 17 приведены минимальные значения среднеквадратичного отклонения восстановленного сигнала от истинного для каждого метода и при различных значениях параметров для случаев $P = 2$, $P = 5$ и $P = 9$ соответственно. Жирным шрифтом выделены минимальные по методам значения RMSE оценки сигнала для фиксированного набора параметров. Отличие этих минимальных значений от остальных значимо при уровне значимости $\alpha = 0.05$.

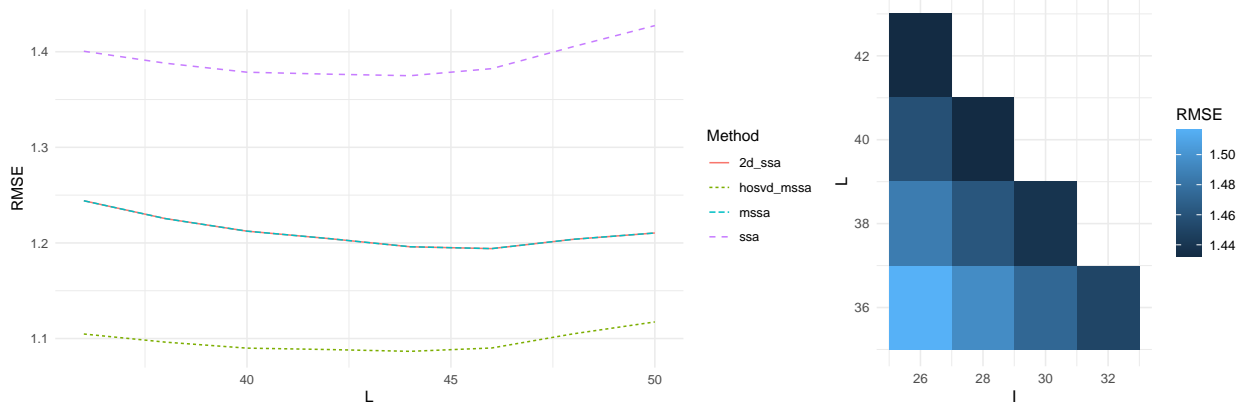
На графиках 2, 3, 4 приведены зависимости точности восстановления сигнала от выбора параметра L для методов SSA, MSSA, HOSVD MSSA и 2D-SSA, и от выбора параметров I, L для метода HOOI SSA для случаев $P = 2$, $P = 5$ и $P = 9$ соответственно.

Замечание. В методе 2D-SSA параметр длины окна двумерный, то есть $L_{xy} = (L_x, L_y)$. Будем полагать $L_y = [(P + 1)/2]$ и $L_x = L$, где $[t]$ обозначает взятие целой части числа t .

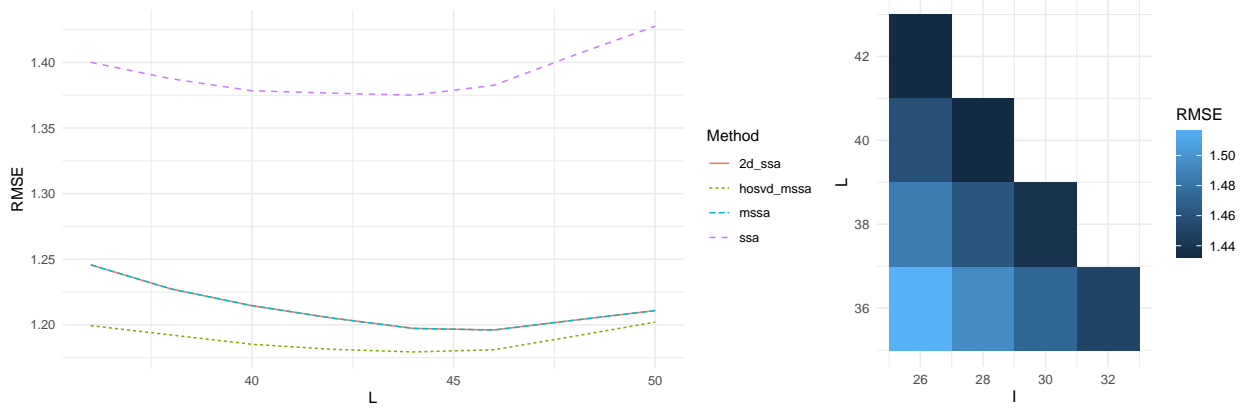
Таблица 12. Сравнение SSA, HOOI SSA, MSSA, HOSVD MSSA и 2D-SSA: ранги сигнала вида (15), $P = 2$, $a_1 = 30$, $a_2 = 20$.

| Параметры | (HOOI) SSA | (HOSVD) MSSA | r_3 | 2D-SSA |
|---|------------|--------------|-------|--------|
| $\omega_1 = \omega_2 = \frac{1}{12}$ $\psi_1 = \psi_2 = 0$ | 2 | 2 | 1 | 2 |
| $\omega_1 = \omega_2 = \frac{1}{12}$ $\psi_1 = 0, \psi_2 = \frac{\pi}{4}$ | 2 | 2 | 2 | 2 |
| $\omega_1 = \frac{1}{12}, \omega_2 = \frac{1}{8}$ $\psi_1 = 0, \psi_2 = \frac{\pi}{4}$ | 2 | 4 | 2 | 4 |

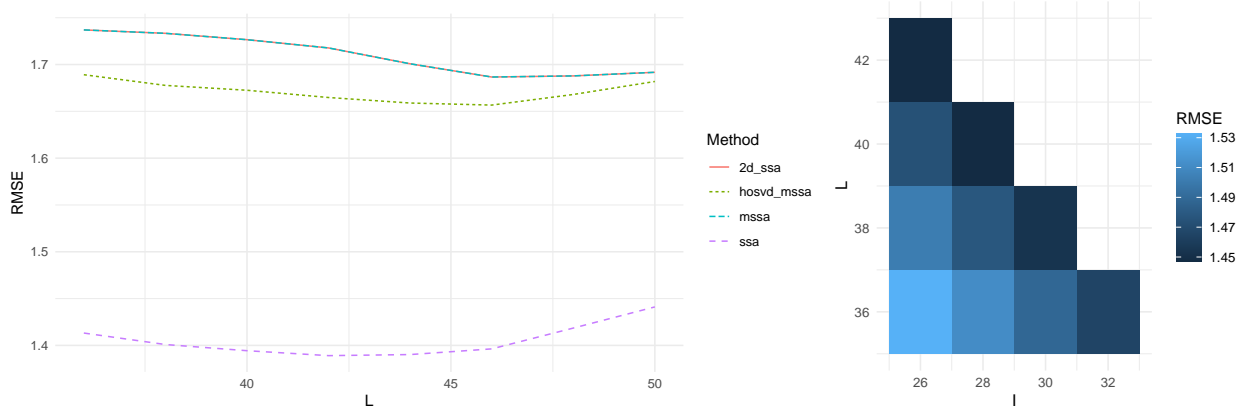
Итоги численного сравнения Таким образом, метод HOSVD MSSA выделил сигнал точнее, чем метод MSSA во всех примерах, и оказался точнее метода 2D-SSA в случае малого числа рядов и в случае различающихся амплитуд у сигналов, причём



$$(a) \omega_1 = \omega_2 = 1/12, \psi_1 = \psi_2 = 0$$



$$(b) \omega_1 = \omega_2 = 1/12, \psi_1 = 0, \psi_2 = \pi/4$$



$$(c) \omega_1 = 1/12, \omega_2 = 1/8, \psi_1 = 0, \psi_2 = \pi/4$$

Рис. 2. Зависимости точности выделения сигнала от длин окна для каждого из методов SSA, MSSA, HOSVD MSSA, 2D-SSA (слева), HOOI SSA (справа), $P = 2$, $a_1 = 30$, $a_2 = 20$.

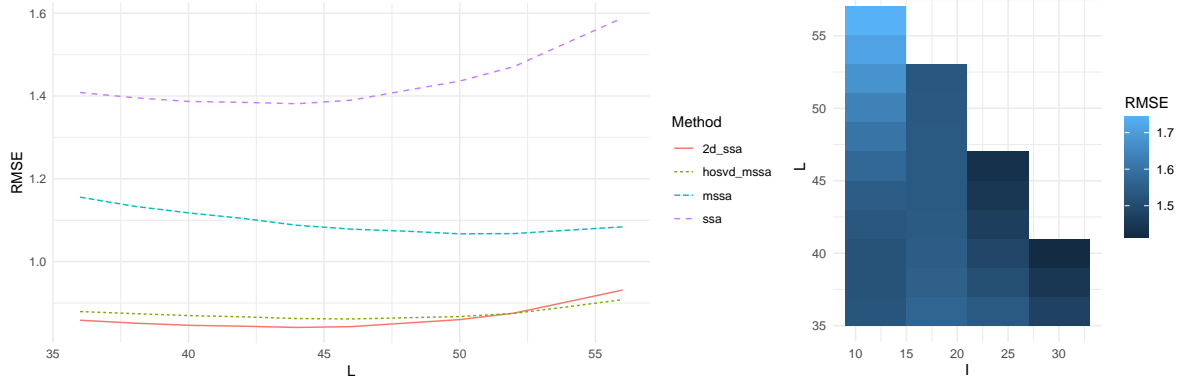
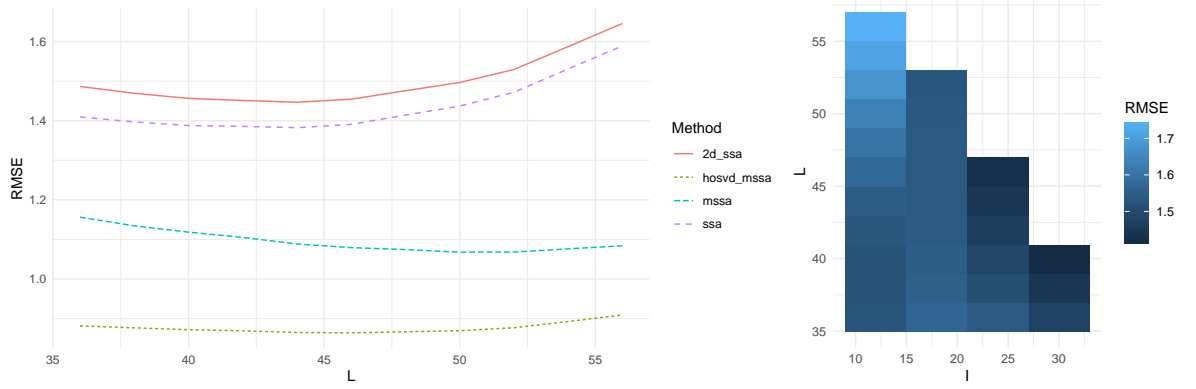
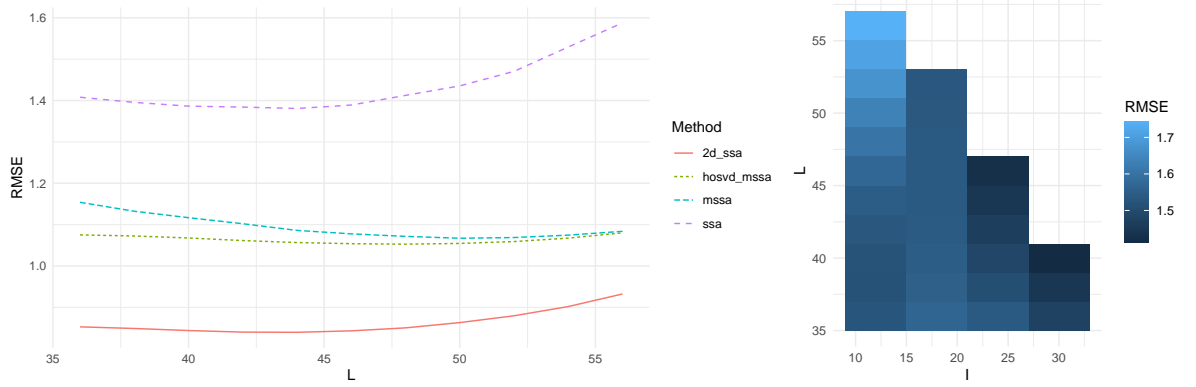
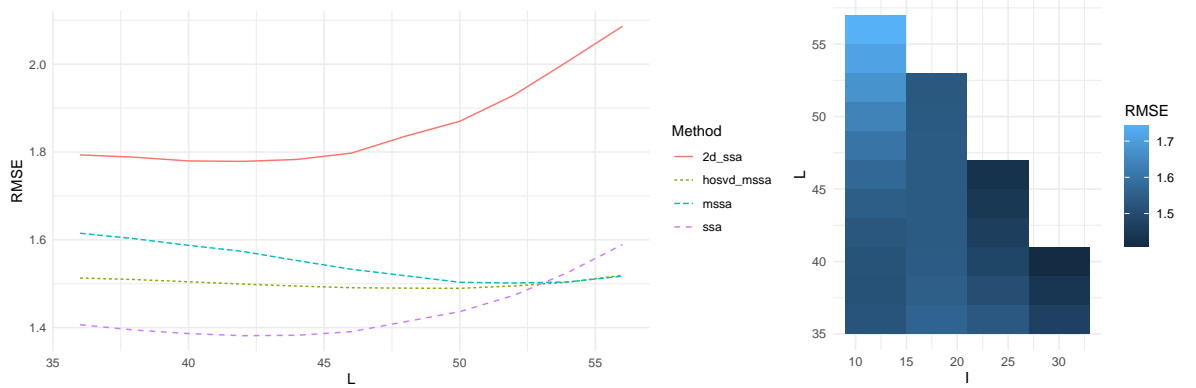
(a) $a_p = 30$, $\omega_p = 1/12$, $\psi_p = 0$ (b) $a_1 = a_4 = 30$, $a_2 = a_5 = 20$, $a_3 = 25$, $\omega_p = 1/12$, $\psi_p = 0$ (c) $a_p = 30$, $\omega_p = 1/12$, $\psi_p = 2(p-1)\pi/3$ (d) $a_p = 30$, $\omega_1 = \omega_2 = \omega_3 = 1/12$, $\omega_4 = \omega_5 = 1/8$, $\psi_p = 2(p-1)\pi/3$

Рис. 3. Зависимости точности выделения сигнала от длин окна для каждого из методов SSA, MSSA, HOSVD MSSA, 2D-SSA (слева), HOOI SSA (справа), $P = 5$.

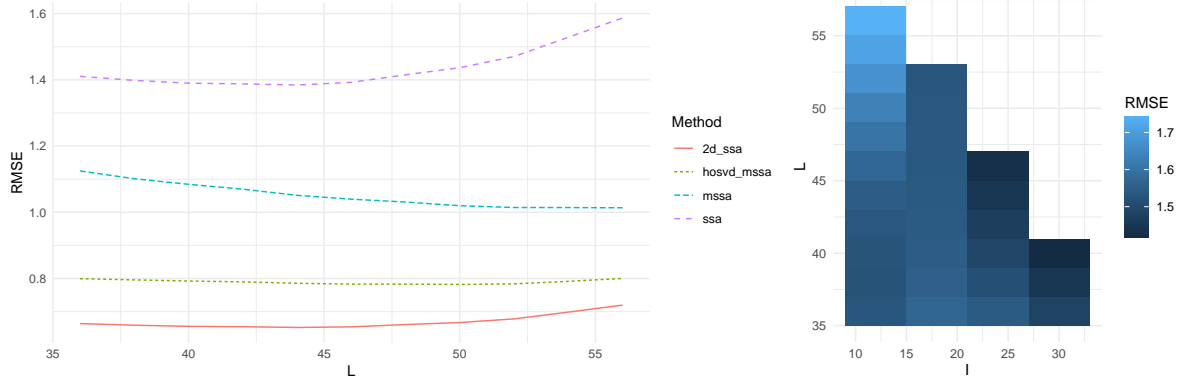
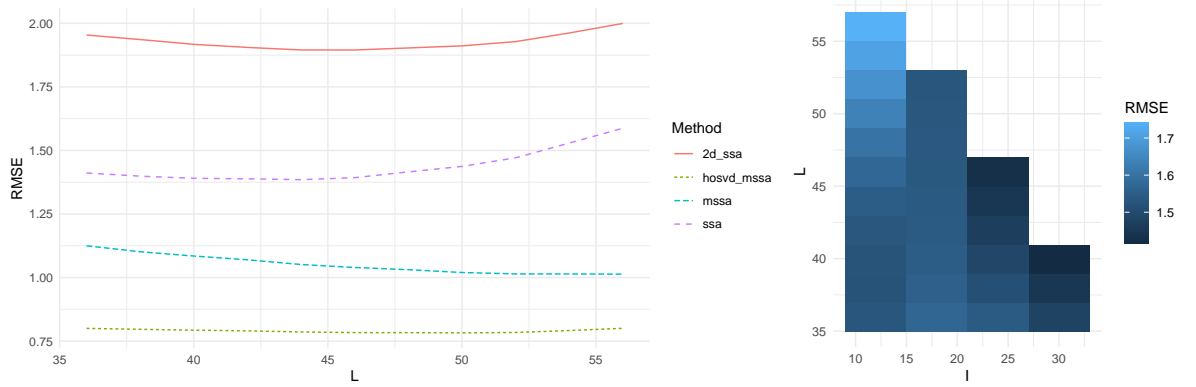
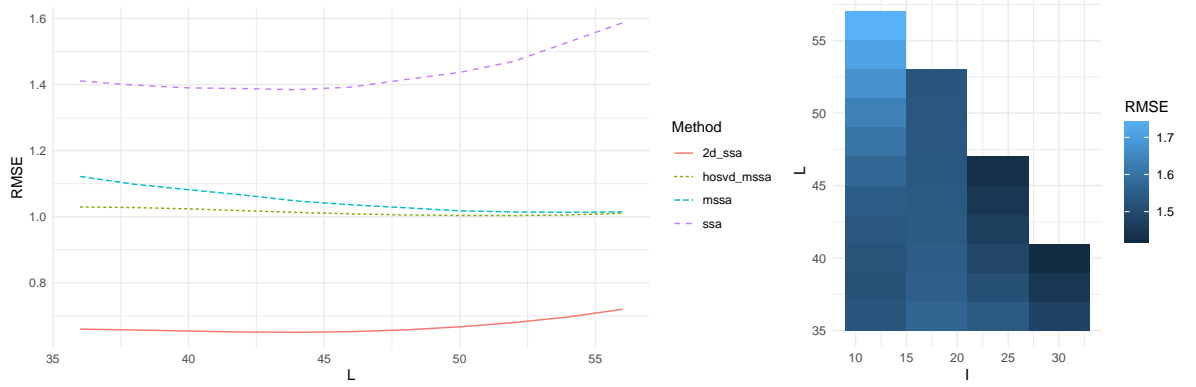
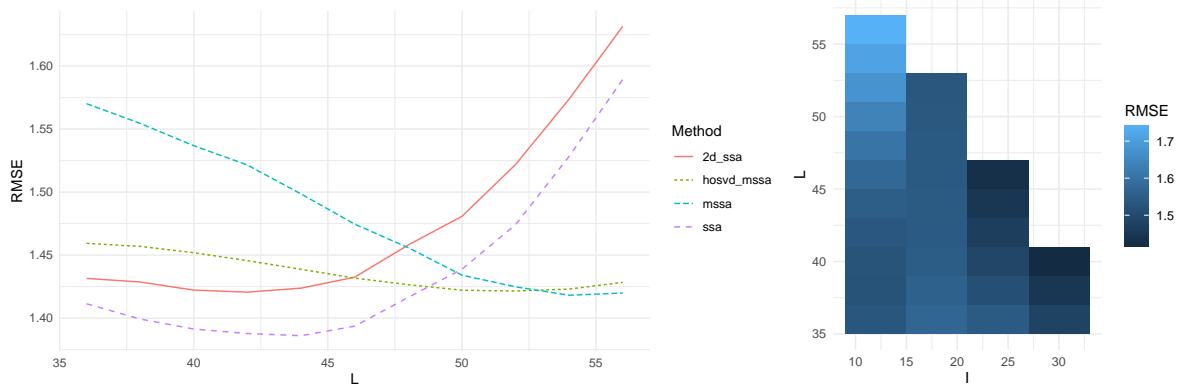
(a) $a_p = 30$, $\omega_p = 1/12$, $\psi_p = 0$ (б) $a_p = a_{p+5} = 30 - 2(p-1)$, $1 \leq p \leq 4$, $a_5 = 22$, $\omega_p = 1/12$, $\psi_p = 0$ (с) $a_p = 30$, $\omega_p = 1/12$, $\psi_p = 2(p-1)\pi/5$ (д) $a_p = 30$, $\omega_p = 1/12$, $1 \leq p \leq 7$, $\omega_8 = \omega_9 = 1/8$, $\psi_p = 2(p-1)\pi/5$

Рис. 4. Зависимости точности выделения сигнала от длин окна для каждого из методов SSA, MSSA, HOSVD MSSA, 2D-SSA (слева), HOOI SSA (справа), $P = 9$.

Таблица 13. Сравнение SSA, HOOI SSA, MSSA, HOSVD MSSA и 2D-SSA: минимальные RMSE оценки сигнала вида (15), $P = 2$, $a_1 = 30$, $a_2 = 20$.

| Параметры | SSA | HOOI SSA | MSSA | HOSVD MSSA | 2D-SSA |
|---|-------------|----------|------|-------------|--------|
| $\omega_1 = \omega_2 = \frac{1}{12}$ $\psi_1 = \psi_2 = 0$ | 1.37 | 1.43 | 1.19 | 1.09 | 1.19 |
| $\omega_1 = \omega_2 = \frac{1}{12}$ $\psi_1 = 0, \psi_2 = \frac{\pi}{4}$ | 1.38 | 1.43 | 1.20 | 1.18 | 1.20 |
| $\omega_1 = \frac{1}{12}, \omega_2 = \frac{1}{8}$ $\psi_1 = 0, \psi_2 = \frac{\pi}{4}$ | 1.39 | 1.45 | 1.69 | 1.66 | 1.69 |

Таблица 14. Сравнение SSA, HOOI SSA, MSSA, HOSVD MSSA и 2D-SSA: ранги сигнала вида (15), $P = 5$.

| Параметры | (HOOI) SSA | (HOSVD) MSSA | r_3 | 2D-SSA |
|--|------------|--------------|-------|--------|
| $a_p = 30$ $\omega_p = \frac{1}{12}, \psi_p = 0$ | 2 | 2 | 1 | 2 |
| $a_1 = a_4 = 30$ $a_2 = a_5 = 20, a_3 = 25$ $\omega_p = \frac{1}{12}, \psi_p = 0$ | 2 | 2 | 1 | 6 |
| $a_p = 30, \omega_p = \frac{1}{12},$ $\psi_p = 2(p-1)\pi/3$ | 2 | 2 | 2 | 2 |
| $a_p = 30, \psi_p = 2(p-1)\pi/3$ $\omega_1 = \omega_2 = \omega_3 = \frac{1}{12}$ $\omega_4 = \omega_5 = \frac{1}{8}$ | 2 | 4 | 4 | 10 |

последнее преимущество увеличилось с увеличением размерности ряда. Однако в случаях равных сигналов и сигналов с меняющейся линейно фазой метод 2D-SSA оказался точнее, чем HOSVD MSSA, причём это преимущество увеличилось при увеличении размерности ряда.

Трудоёмкости алгоритмов Так как шаг разложения является самым трудоёмким шагом в алгоритмах HOSVD SSA и HOSVD MSSA, то для оценки трудоёмкости алгоритма достаточно оценить трудоёмкость HOSVD траекторного тензора. Оценим трудо-

Таблица 15. Сравнение SSA, HOOI SSA, MSSA, HOSVD MSSA и 2D-SSA: минимальные RMSE оценки сигнала вида (15), $P = 5$.

| Параметры | SSA | HOOI SSA | MSSA | HOSVD MSSA | 2D-SSA |
|--|-------------|----------|------|-------------|-------------|
| $a_p = 30$ $\omega_p = \frac{1}{12}, \psi_p = 0$ | 1.34 | 0.00 | 0.95 | 0.75 | 0.73 |
| $a_1 = a_4 = 30$ $a_2 = a_5 = 20, a_3 = 25$ $\omega_p = \frac{1}{12}, \psi_p = 0$ | 1.38 | 1.41 | 1.07 | 0.86 | 1.45 |
| $a_p = 30, \omega_p = \frac{1}{12}$ $\psi_p = 2(p-1)\pi/3$ | 1.38 | 1.41 | 1.07 | 1.05 | 0.84 |
| $a_p = 30, \psi_p = 2(p-1)\pi/3$ $\omega_1 = \omega_2 = \omega_3 = \frac{1}{12}$ $\omega_4 = \omega_5 = \frac{1}{8}$ | 1.38 | 1.41 | 1.50 | 1.49 | 1.78 |

Таблица 16. Сравнение SSA, HOOI SSA, MSSA, HOSVD MSSA и 2D-SSA: ранги сигнала вида (15), $P = 9$.

| Параметры | (HOOI) SSA | (HOSVD) MSSA | r_3 | 2D-SSA |
|---|------------|--------------|-------|--------|
| $a_p = 30$ $\omega_p = \frac{1}{12}, \psi_p = 0$ | 2 | 2 | 1 | 2 |
| $a_p = a_{p+5} = 30 - 2(p-1), 1 \leq p \leq 4$ $a_5 = 22, \omega_p = \frac{1}{12}, \psi_p = 0$ | 2 | 2 | 1 | 10 |
| $a_p = 30, \omega_p = \frac{1}{12},$ $\psi_p = 2(p-1)\pi/5$ | 2 | 2 | 2 | 2 |
| $a_p = 30, \psi_p = 2(p-1)\pi/5$ $\omega_p = \frac{1}{12}, 1 \leq p \leq 7$ $\omega_8 = \omega_9 = \frac{1}{8}$ | 2 | 4 | 4 | 10 |

ёмкость алгоритма HOSVD MSSA, трудоёмкость HOSVD SSA получается аналогично подстановкой других обозначений размерностей.

Из утверждения 2.1 следует, что вычисление HOSVD траекторного тензора размерности $L \times K \times P$ сводится к вычислению SVD на трёх матрицах размерностей $L \times$

Таблица 17. Сравнение SSA, HOOI SSA, MSSA, HOSVD MSSA и 2D-SSA: минимальные RMSE оценки сигнала вида (15), $P = 9$.

| Параметры | SSA | HOOI SSA | MSSA | HOSVD MSSA | 2D-SSA |
|---|-------------|----------|------|-------------|-------------|
| $a_p = 30$ $\omega_p = \frac{1}{12}, \psi_p = 0$ | 1.38 | 1.42 | 1.01 | 0.78 | 0.65 |
| $a_p = a_{p+5} = 30 - 2(p - 1)$ $1 \leq p \leq 4, a_5 = 22$ $\omega_p = \frac{1}{12}, \psi_p = 0$ | 1.38 | 1.42 | 1.01 | 0.78 | 1.90 |
| $a_p = 30, \omega_p = \frac{1}{12}$ $\psi_p = 2(p - 1)\pi/5$ | 1.38 | 1.42 | 1.01 | 1.00 | 0.65 |
| $a_p = 30, \psi_p = 2(p - 1)\pi/5$ $\omega_p = \frac{1}{12}, 1 \leq p \leq 7$ $\omega_8 = \omega_9 = \frac{1}{8}$ | 1.39 | 1.42 | 1.42 | 1.42 | 1.42 |

$KP, K \times LP, P \times LK$. Вычисление SVD матрицы размерности $m \times n$ имеет асимптотическую трудоёмкость, равную $O(\min(mn^2, m^2n))$. Таким образом, асимптотическая сложность HOSVD траекторного тензора, а следовательно и всего алгоритма, имеет вид

$$O(LKP(\min(L, KP) + \min(K, LP) + \min(P, LK))). \quad (16)$$

Самым трудоёмким шагом в алгоритме HOOI SSA является применение к траекторному тензору алгоритма HOOI. Описание этого алгоритма для приближения тензора размерности $I \times J \times K$ тензором с n -рангами (r_1, r_2, r_3) можно найти в статье [11]. Этот алгоритм является итерационным, и в обозначениях выше каждая итерация имеет асимптотическую сложность

$$O(r_1^2 \min(I, r_2 r_3) + r_2^2 \min(J, r_1 r_3) + r_3^2 \min(K, r_1 r_2)). \quad (17)$$

Если предположить, что пользователь задаёт верхнюю границу количества итераций C , то асимптотическая сложность алгоритма HOOI будет иметь вид (17) с константой, пропорциональной C .

9. Альтернативные тензорные разложения

Помимо HOSVD, существует ещё один тип тензорных разложений: ранговое разложение тензора. Идея заключается в представлении тензора \mathcal{A} в виде линейной комбинации R тензоров ранга 1, где $R = \text{rank}(\mathcal{A})$. Однако нахождение этого ранга в общем случае является NP-трудной задачей [12].

CANDECOMP-PARAFAC [6, 7]— итерационный метод приближения тензора суммой заданного пользователем числа тензоров ранга 1. То есть по параметру K этот метод считает наилучшее приближение входного тензора суммой K тензоров ранга 1. Заметим, что из-за отсутствия каких-либо требований к ортогональности в определении рангового разложения тензора, многие свойства, верные в теории SSA, могут потерять справедливость при использовании этого разложения.

Рассмотрим ряды $\tilde{x}_n = 3$, $\hat{x}_n = \sin(2\pi n/3)$, $n \in \overline{0:15}$. Построим по этим рядам траекторные тензоры с параметрами $I = L = 6$. Тогда ранг траекторного тензора $\tilde{\mathcal{X}}$, соответствующего константному ряду, равен 1, так как его можно представить в виде

$$\tilde{\mathcal{X}} = 3X \circ X \circ X,$$

где $X = (1, 1, 1, 1, 1, 1)$. Ранг траекторного тензора $\hat{\mathcal{X}}$, соответствующего синусу, равен 3, так как его можно представить в виде

$$\hat{\mathcal{X}} = \sum_{k=1}^3 \lambda_k X_k \circ Y_k \circ Z_k,$$

где

$$\lambda_1 = 160.56, \lambda_2 = 65.69, \lambda_3 = 123.76,$$

$$\mathbf{X} = [X_1, X_2, X_3] = \begin{pmatrix} -0.16 & 0.25 & 0.06 \\ 0.25 & -0.06 & -0.25 \\ -0.09 & -0.19 & 0.19 \\ -0.16 & 0.25 & 0.06 \\ 0.25 & -0.06 & -0.25 \\ -0.09 & -0.19 & 0.19 \end{pmatrix},$$

$$\mathbf{Y} = [Y_1, Y_2, Y_3] = \begin{pmatrix} -0.25 & -0.15 & 0.21 \\ 0.18 & -0.10 & -0.25 \\ 0.07 & 0.25 & 0.04 \\ -0.25 & -0.15 & 0.21 \\ 0.18 & -0.10 & -0.25 \\ 0.07 & 0.25 & 0.04 \end{pmatrix},$$

$$\mathbf{Z} = [Z_1, Z_2, Z_3] = \begin{pmatrix} -0.10 & -0.25 & -0.01 \\ 0.25 & 0.12 & -0.24 \\ -0.15 & 0.13 & 0.25 \\ -0.10 & -0.25 & -0.01 \\ 0.25 & 0.12 & -0.24 \\ -0.15 & 0.13 & 0.25 \end{pmatrix},$$

притом точных приближений двумя тензорами ранга 1 нет.

Траекторный тензор ряда $x_n = \tilde{x}_n + \hat{x}_n$, построенный с параметрами $I = L = 6$ представим в виде суммы

$$\mathcal{X} = \sum_{k=1}^4 \lambda_k X_k \circ Y_k \circ Z_k,$$

где

$$\lambda_1 = 320.17, \lambda_2 = 120.97, \lambda_3 = 209.38, \lambda_4 = 648$$

$$\mathbf{X} = [X_1, X_2, X_3, X_4] = \begin{pmatrix} -0.25 & -0.25 & 0.25 & -0.17 \\ 0.11 & 0.25 & -0.04 & -0.17 \\ 0.14 & 0.00 & -0.21 & -0.17 \\ -0.25 & -0.25 & 0.25 & -0.17 \\ 0.11 & 0.25 & -0.04 & -0.17 \\ 0.14 & 0.00 & -0.21 & -0.17 \end{pmatrix},$$

$$\mathbf{Y} = [Y_1, Y_2, Y_3, Y_4] = \begin{pmatrix} 0.25 & -0.25 & -0.25 & -0.17 \\ -0.08 & 0.21 & -0.00 & -0.17 \\ -0.17 & 0.04 & 0.25 & -0.17 \\ 0.25 & -0.25 & -0.25 & -0.17 \\ -0.08 & 0.21 & -0.00 & -0.17 \\ -0.17 & 0.04 & 0.25 & -0.17 \end{pmatrix},$$

$$\mathbf{Z} = [Z_1, Z_2, Z_3, Z_4] = \begin{pmatrix} -0.00 & -0.14 & -0.08 & 0.17 \\ -0.25 & -0.11 & 0.25 & 0.17 \\ 0.25 & 0.25 & -0.17 & 0.17 \\ -0.00 & -0.14 & -0.08 & 0.17 \\ -0.25 & -0.11 & 0.25 & 0.17 \\ 0.25 & 0.25 & -0.17 & 0.17 \end{pmatrix},$$

притом точных приближений тремя тензорами ранга 1 нет.

По виду векторов видно, что четвёртая компонента разложения соответствует константному ряду, а остальные три имеют период равный 3. Таким образом, несмотря на отсутствие ограничений на ортогональность в определении ранговых разложений тензора, наблюдается отделимость константного ряда от периодического ряда при наличии условий слабой разделимости в терминах SSA и отсутствия шума. Однако понятия ранга в терминах SSA и в терминах CPD различаются, так как в терминах SSA синус с периодом 3 имеет ранг 2, а в терминах рангового разложения, как показано выше, такой синус имеет ранг 3.

Другим недостатком CPD является то, что это итерационный метод, причём процесс итерации начинается с генерации случайной матрицы, в связи с чем на одних и тех же данных он может выдавать разные результаты, в том числе может как сойтись, так и нет.

Возможно можно добиться лучших результатов, используя CPD или его модификации, если строить тензор по ряду другим образом и подбирать другие параметры разложения. Этот вопрос предлагается изучить в будущих работах.

10. Заключение

В работе было показано, что по своим свойствам HO-SSA и HOSVD MSSA с использованием HOSVD имеют много общего с Basic SSA и MSSA соответственно. Однако, есть и особенности, кардинально меняющие свойства методов, которые можно использовать для анализа временных рядов.

В результате исследования методов HO-SSA и HOSVD MSSA с HOSVD были сделаны следующие выводы: методы можно использовать для выделения сигнала, однако остался неизученным вопрос о том, как с их помощью разделять компоненты сигнала. Также, в большинстве случаев, как HOSVD SSA, так и HOOI SSA, оказались хуже, чем Basic SSA в точности выделения сигнала. Удалось построить только один пример, где Basic SSA менее точно выделяет сигнал. Это противоречит результатам работы [4], где показано преимущество Tensor SSA с использованием HOOI. Однако, в этой статье сравнение идет для сигнала в виде суммы двух комплексных экспонент по точности оценке частоты сигналов.

С другой стороны, метод HOSVD MSSA выделил сигнал точнее метода MSSA во всех рассмотренных случаях, но в некоторых из них данный метод оказался менее точным, чем метод 2D-SSA.

Таким образом, остались открытыми вопросы: подтвердить преимущество HOSVD-SSA для оценивания частот сигнала, о котором утверждается в статье [4], найти методы, которые разделяют компоненты сигнала, например в работе [13] предлагается метод CPD для разделения комплексных экспонент, а также расширить теорию HOSVD MSSA.

Список литературы

1. Golyandina Nina, Nekrutkin Vladimir, Zhigljavsky Anatoly. Analysis of time series structure: SSA and related techniques. — Chapman & Hall/CRC, 2001.
2. Kouchaki Samaneh, Sanei Saeid. Tensor based singular spectrum analysis for nonstationary source separation // 2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP). — IEEE. — 2013.
3. Improved Tensor-Based Singular Spectrum Analysis Based on Single Channel Blind Source Separation Algorithm and Its Application to Fault Diagnosis / Yang Dan, Yi Cancan, Xu Zengbin, Zhang Yi, Ge Mao, and Liu Changming // Applied Sciences. — 2017. — Vol. 7, no. 4.
4. Papy J. M., De Lathauwer L., Van Huffel S. Exponential data fitting using multilinear algebra: the single-channel and multi-channel case // Numerical Linear Algebra with Applications. — 2005. — Vol. 12, no. 8. — P. 809–826.
5. De Lathauwer Lieven, De Moor Bart, Vandewalle Joos. A Multilinear Singular Value Decomposition // SIAM Journal on Matrix Analysis and Applications. — 2000. — Vol. 21, no. 4. — P. 1253–1278.
6. Harshman Richard A. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-model factor analysis. — 1970. — Vol. 16. — P. 1–84.
7. Carroll J. Douglas, Chang Jih Jie. Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition // Psychometrika. — 1970. — Vol. 35. — P. 283–319.
8. Tucker Ledyard R. Some mathematical notes on three-mode factor analysis // Psychometrika. — 1966. — Vol. 31. — P. 279–311.
9. Golyandina Nina, Zhigljavsky Anatoly. Singular Spectrum Analysis for Time Series. — Springer Berlin, Heidelberg, 2020.
10. De Lathauwer Lieven, De Moor Bart, Vandewalle Joos. On the Best Rank-1 and Rank- (R_1, R_2, \dots, R_N) Approximation of Higher-Order Tensors // SIAM Journal on Matrix Analysis and Applications. — 2000. — Vol. 21, no. 4. — P. 1324–1342.
11. Sheehan Bernard, Saad Yousef. Higher Order Orthogonal Iteration of Tensors (HOOI) and its Relation to PCA and GLRAM. — 2007. — 04.
12. Hillar Christopher J., Lim Lek-Heng. Most Tensor Problems Are NP-Hard // J. ACM. —

2013. — Vol. 60, no. 6.

13. De Lathauwer Lieven. Blind Separation of Exponential Polynomials and the Decomposition of a Tensor in Rank- $(L_r, L_r, 1)$ Terms // SIAM Journal on Matrix Analysis and Applications. — 2011. — Vol. 32, no. 4. — P. 1451–1474.