

# Обзор литературы по Tensor SSA

## 1 Статьи с теорией тензорных разложений

### 1.1 [A Multilinear Singular Value Decomposition](#)

Базовая теория по **HOSVD** (определения, свойства).

### 1.2 [On the Best Rank-1 and Rank- \$\(R\_1, R\_2, \dots, R\_N\)\$ Approximation of Higher-Order Tensors](#)

Про наилучшее приближение тензора меньшими рангами, описание алгоритма **HOOI**, некоторые его свойства.

### 1.3 [Tensor SVD: Statistical and Computational Limits](#)

Рассматривается точность приближения **HOOI** произвольного тензора по его зашумлённому варианту при различных случаях отношения минимального сингулярного числа тензора к уровню шума (SNR).

### 1.4 [Factorization strategies for third-order tensors и Third-Order Tensors as Operators on Matrices: A Theoretical and Computational Framework with Applications in Imaging](#)

В основном теория по трёхмерным тензорам, вводится определение **TSVD** и его свойства.

## 2 Tensor SSA с использованием HOSVD или HOOI

### 2.1 [Exponential data fitting using multilinear algebra: the single-channel and multi-channel case](#)

Рассматривается задача оценки параметров комплексного сигнала (одномерный и многомерный случаи), состоящего из суммы комплексных экспонент с близкими частотами. Приводится описание и обоснование тензорной модификации алгоритма ESPRIT с применением **HOOI**. Проводится численное сравнение этой модификации с базовым ESPRIT. Выявлено преимущество тензорного метода по точности оценки параметров сигнала в смысле RRMSE, причём с увеличением уровня шума, преимущество увеличивается.

Траекторным тензором одномерного ряда  $\mathbf{X}$  с параметрами  $I, L : 1 < I, L < N, I + L < N + 1$  считается тензор  $\mathcal{X}$  размера  $I \times L \times J$ ,  $J = N - I - L + 2$ , элементы которого удовлетворяют равенству

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

Визуализация на рисунке 1.

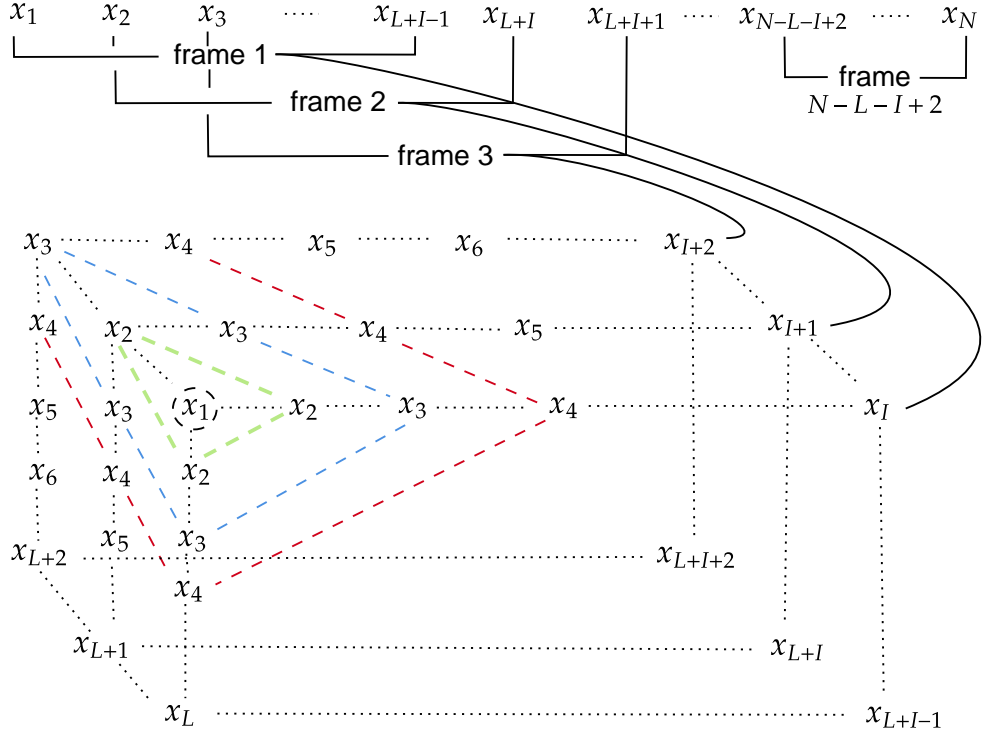


Рис. 1: Траекторный тензор одномерного ряда в HOSVD-SSA.

Траекторным тензором многомерного ряда  $\mathbf{X}$  с длиной окна  $L : 1 < L < N$  считается тензор  $\mathcal{X}$  размерности  $L \times K \times P$ ,  $K = N - L + 1$ , элементы которого удовлетворяют равенству

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

Визуализация на рисунке 2.

Алгоритм заключается в нахождении наилучшего приближения траекторного тензора ряда с  $n$ -рангами  $(R, R, R)$  с помощью алгоритма **HOOF**.  $R$  задаётся равным числу экспонент с различными показателями входящих в сигнал. Затем оценка строится по матрице сингулярных векторов одного из направлений (1-го направления в одномерном случае, и 3-го в многомерном) тем же способом, что и в базовом ESPRIT.

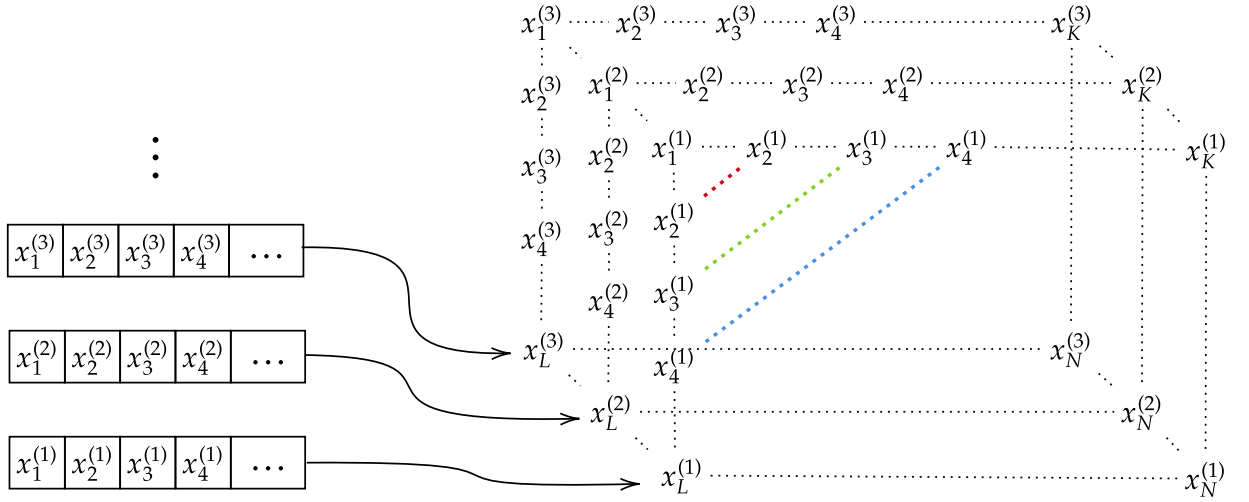


Рис. 2: Траекторный тензор многомерного ряда в HOSVD-MSSA.

### Наши результаты

1. В одномерном случае оценки параметров множество длин окна, при которых тензорный метод даёт более точный результат, довольно мало. Большинство параметров длин окна в тензорном методе дают гораздо менее точный результат. Когда преимущество есть, оно невелико.
2. В одномерном случае выделения сигнала тензорный метод не даёт преимущества над SSA ни при каких длинах окна. Наибольшей точности тензорный метод достигает при усечении траекторного тензора только по направлению наибольшего размера, особенно когда у одного из оставшихся направлений размер наименьший (ранг + 1).
3. В многомерном случае оценки параметров тензорный метод точнее базового на одних и тех же длинах окна.
4. В многомерном случае выделения сигнала ситуация аналогична задаче оценки параметров многомерного случая.

## 2.2 Exponential data fitting using multilinear algebra: the decimative case

Рассматривается задача оценки параметров одномерного комплексного сигнала (только одномерный случай), состоящего из суммы экспоненциально-модулированных гармоник с близкими частотами. Исследуется алгоритм HTLSDstack: модификация ESPRIT, в которой по одномерному ряду строится  $D$  прореженных рядов длины  $M = N/D$  (считается, что длина ряда  $N$  делится на  $D$  нацело). Затем они считаются отдельными каналами одного многомерного ряда, и применяется многомерный вариант ESPRIT. Это уменьшает трудоёмкость алгоритма при небольшом уменьшении точности. Предлагается тензорная модификация этого алгоритма: HO-HTLSDstack, в которой к полученному многомерному ряду применяется тензорная модификация многомерного ESPRIT из 2.1.

Исходный ряд  $\mathbf{X}$  длины  $N$  разбивается на  $D$  прореженных подрядов  $\mathbf{X}^{(d)}$  длины  $M$  так, что  $x_m^{(d)} = x_{(m-1)D+d}$ . Другими словами, в ряд с номером  $d$  входит каждый  $D$ -й элемент исходного ряда, начиная с  $x_d$ . По полученному многомерному ряду строится траекторный тензор так же, как в 2.1. Затем ищется наилучшее приближение этого тензора с  $n$ -рангами  $(R, \min(R, M), R')$ , где  $R$  задаётся равным числу экспонент с различными показателями входящих в сигнал, а  $R' \leq \min(R, D)$ . Авторы утверждают, что если частоты гармоник близки, то выбор  $R' < \min(R, D)$  улучшает точность оценки параметров.

Для оценки параметров используются сингулярные векторы 1-го направления полученного разложения аппроксимирующего тензора.

Численно показано, что тензорный вариант (НО-HTLSDstack) с выбором  $R' = 1$  оказывается точнее HTLSDstack во всех тестах, причём преимущество увеличивается с увеличением уровня шума. Также в одном тесте НО-HTLSDstack сравнивается с базовым ESPRIT и так же оказывается более точной. Во всех тестах рассматривался случай близких частот.

Кроме того, авторы показывают, что метод НО-HTLSDstack меньше чем на порядок более трудоёмкий, чем HTLSDstack, но на порядок менее трудоёмкий, чем базовый ESPRIT.

Графики со сравнением методов по точности и трудоёмкости, приведённые в статье, указаны на рисунке 3.

**Наши результаты** Результаты статьи пока не проверялись.

### 3 Tensor SSA с использованием $(L_r, L_r, 1)$ -разложения

#### 3.1 Blind Separation of Exponential Polynomials and the Decomposition of a Tensor in Rank- $(L_r, L_r, 1)$ Terms

Приводится теоретическая информация про разложение в сумму тензоров с  $n$ -рангами  $(L_r, L_r, 1)$ , в частности определение и условия единственности.

Рассматривается задача выделения сигнала в многомерном ряде

$$\mathbf{Y} = \mathbf{MS} + \mathbf{N},$$

где  $\mathbf{Y}$  — наблюдаемый ряд,  $\mathbf{S}$  — искомый сигнал,  $\mathbf{M}$  — коэффициенты линейных комбинаций, с которыми сигнал составляет наблюдаемый ряд,  $\mathbf{N}$  — шум. Траекторный тензор ряда определяется так же, как в 2.1.

Для модели, в которой сигнал составляют суммы произведений полиномов и комплексных экспонент, доказаны условия единственности  $(L_r, L_r, 1)$ -разложения траекторного тензора ряда.

Метод заключается в построении траекторного тензора по ряду  $\mathbf{Y}$ , аппроксимации этого тензора меньшими  $n$ -рангами (но большими, чем  $n$ -ранги самого сигнала), и применении  $(L_r, L_r, 1)$ -разложения к этой аппроксимации. Далее по этому разложению можно построить оценку  $\mathbf{S}$ .

Проводятся численные сравнения точности выделения сигнала предложенным методом при различных выборах параметров рангов аппроксимации и  $L_r$ . Сравнения с другими методами выделения сигнала не проводится.

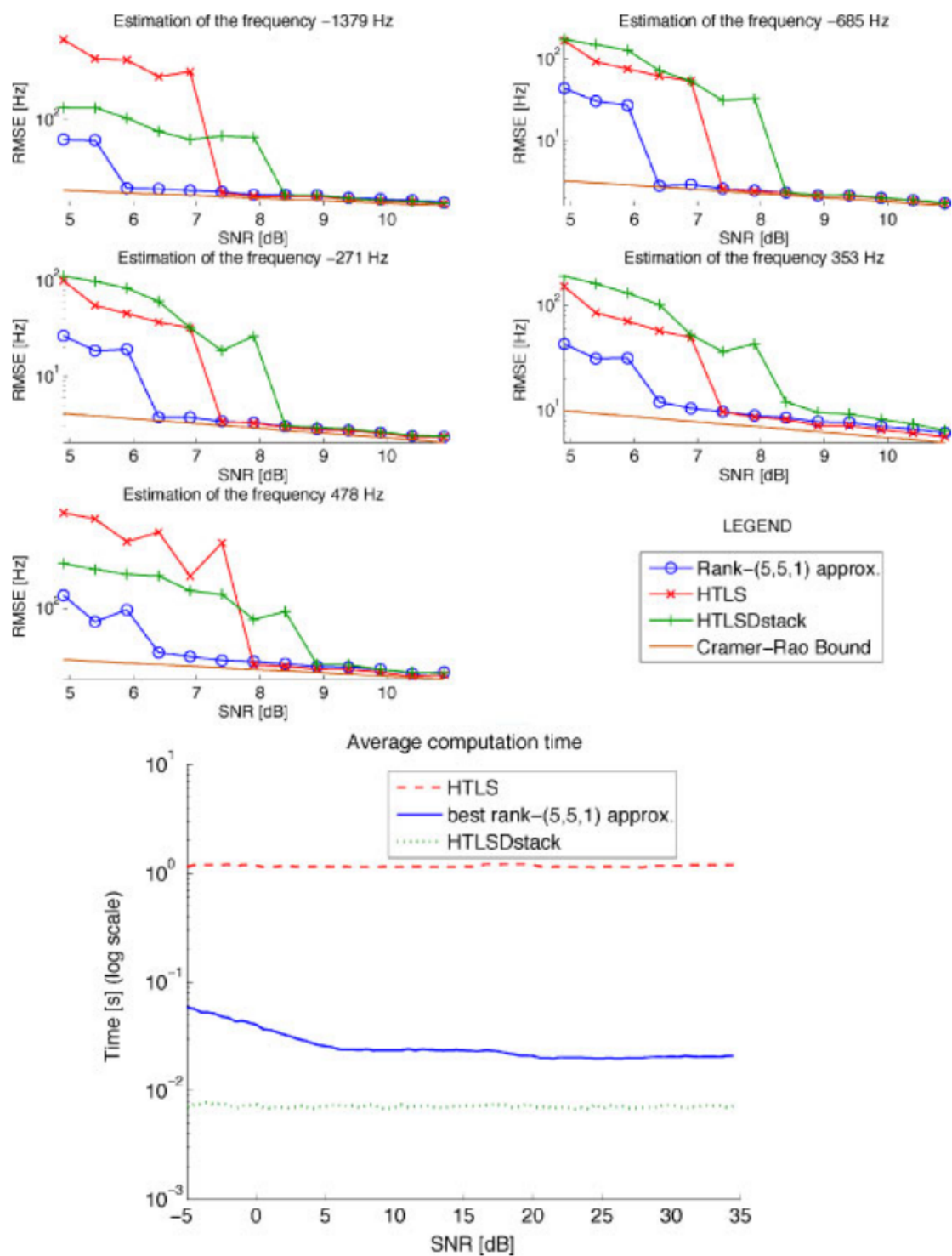


Рис. 3: Сравнение HTLS (ESPRIT), HTLSDstack и HO-HTLSDstack по точности оценки параметров и по трудоёмкости.

**Наши результаты** Результаты статьи пока не проверялись.

## 4 Tensor SSA с использованием CPD

### 4.1 Tensor Based Singular Spectrum Analysis for Automatic Scoring of Sleep EEG

Рассматривается задача выделения одномерного вещественного сигнала из ряда с нестационарным шумом.

Вводится понятие траекторного тензора следующим образом: исходный одномерный ряд делится на непересекающиеся подряды длины  $l$ , затем траекторный тензор строится аналогично траекторному тензору многомерного ряда из 2.1, как если бы эти подряды были каналами многомерного сигнала. Предлагается алгоритм Tensor-Based SSA, основанный на применении **CPD** к полученному траекторному тензору ряда.

Кроме того, предлагается алгоритм TSSA-EMD, основанный на использовании метода EMD для адаптивной группировки компонент **CPD**. Метод заключается в получении **CPD** траекторного тензора, использовании EMD для выделения сигнала, построении траекторного тензора  $\mathcal{F}$  этой оценки сигнала, и решении оптимизационных задач

$$J(g_{ijk}) = \left\| \mathcal{F}_{ijk} - \sum_{r=1}^R g_{ijk} a_{ir} b_{jr} c_{kr} \right\|^2 \rightarrow \min_{g_{ijk}}$$

для всех  $i, j, k$ , где  $a_{ir}$ ,  $b_{jr}$  и  $c_{kr}$  — компоненты CPD. После нахождения тензора адаптивных весов  $\mathcal{G} = \{g_{ijk}\}$  строится оценка траекторного тензора сигнала

$$\mathcal{S}_{ijk} = \sum_{r=1}^R \mathcal{G}_{ijk} a_{ir} b_{jr} c_{kr},$$

по которому восстанавливается оценка сигнала.

Метод TSSA-EMD численно сравнивался с методами SCICA, EEMD-ICA и SSA на моделированном ряде с нестационарным шумом (явная формула рассматриваемого ряда в статье не приведена). Было показано значимое преимущество TSSA-EMD над остальными рассматриваемыми методами в смысле RMSE, причём для каждого алгоритма рассматривался лишь один набор параметров.

Также было рассмотрено применение TSSA-EMD к реальному ряду.

**Наши результаты** Применение предложенного метода без адаптивной группировки с помощью EMD давало стабильно менее точные оценки сигнала, чем базовый SSA. Вариант TSSA-EMD не рассматривался.

### 4.2 Improved Tensor-Based Singular Spectrum Analysis Based on Single Channel Blind Source Separation Algorithm and Its Application to Fault Diagnosis

Рассматривается задача выделения одномерного вещественного сигнала из ряда.

Предлагается алгоритм, совпадающий с TSSA из 4.1 (без адаптивной группировки) во всём, кроме метода для вычисления CDP. В 4.1 используется ALS метод минимизации ошибки аппроксимации

$$\|\mathcal{X} - \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket\|^2,$$

где  $\mathcal{X}$  — аппроксимируемый тензор,  $\llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket = \sum_{r=1}^R A_r \circ B_r \circ C_r$ . Авторы предлагают использовать метод CP-WOPT, который заключается в построении весового тензора  $\mathcal{W}$  и минимизации функционала

$$\|\mathcal{W} * \mathcal{X} - \mathcal{W} * \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket\|^2,$$

где знак  $*$  — поэлементное умножение. Эта модификация призвана улучшить сходимость разложения.

Кроме того предлагалось использовать алгоритм EMD-SVD-BIC для определения числа компонент в CPD. Этот алгоритм заключается в нахождении с помощью EMD сигнальных компонент (IMF), построении их ковариационной матрицы, проведении SVD этой матрицы, вычислении BIC по сингулярным числам ковариационной матрицы при различных предположениях о числе сигнальных компонент.

Проводятся численные сравнения предложенного метода CP-WOPT TSSA с немодифицированным TSSA из 4.1, Fast-ICA и EMD-ICA на ряде не конечного ранга (пульсирующая затухающая гармоника). Утверждается преимущество предложенного метода, но без подтверждения какой либо метрикой (лишь визуальные подтверждения).

**Наши результаты** CP-WOPT метод не был применён, так как из статьи неясно по какому принципу строится весовой тензор  $\mathcal{W}$ . Было проведено сравнение TSSA из 4.1 с SSA и MSSA на ряде предложенного в статье вида. Оказалось, что на таком ряде TSSA выделяет сигнал точнее, чем SSA, но менее точно, чем MSSA. Под применением MSSA к одномерному ряду имеется в виду разделение ряда на непересекающиеся подряды одной длины, и рассмотрение их как каналов многомерного ряда.