

Институт вычислительной математики
Российская академия наук
119333, Москва, Губкина 8
www.inm.ras.ru

$\left[\begin{array}{c} \text{публикации} \\ \text{И В М} \\ \text{Р А Н} \end{array} \right]$
pub.inm.ras.ru

Дмитрий Савостьянов

ТЕНЗОРНЫЕ АЛГОРИТМЫ СЛЕПОГО РАЗДЕЛЕНИЯ ЭЛЕКТРОМАГНИТНЫХ СИГНАЛОВ

Препринт №09-09

09-12-19



Москва 2009

ТЕНЗОРНЫЕ АЛГОРИТМЫ СЛЕПОГО РАЗДЕЛЕНИЯ ЭЛЕКТРОМАГНИТНЫХ СИГНАЛОВ

Дмитрий Савостьянов

*Институт вычислительной математики РАН, 119333 Москва, Губкина 8,
dmitry.savostyanov@gmail.com*

19 декабря 2009 г.

Аннотация. Задача слепого разделения электромагнитных сигналов сводится к одновременной диагонализации набора матриц-статистик, для решения которой предлагаются комплексные версии алгоритмов разложения тензоров, основанных на мультипликативном обновлении и суперобобщенном разложении Шура. Проводится сравнение полученных результатов с алгоритмами типа FastICA, обсуждаются преимущества тензорных методов при пакетном режиме обработки сигналов.

Ключевые слова: Слепое разделение сигналов, цифровая обработка сигналов, MIMO-каналы, разложение тензоров, каноническое разложение, одновременная диагонализация матриц, суперобобщенное разложение Шура, FastICA

1. Введение

Разделение сигналов, принятых антенной системой, является основной задачей при передаче информации через MIMO (multi-input multi-output) каналы [1, 2, 3]. Эта технология используется в новейших стандартах передачи данных, таких как IEEE 802.11n (Wi-Fi), IEEE 802.16e (WiMAX) и сотовых сетях четвертого поколения.

Антенная система представляет из себя набор принимающих устройств, каждое из которых характеризуется своей *диаграммой направленности* (ДН), связывающей падающий сигнал и наблюдаемый прибором ток:

$$y(t) = a(\phi, \psi)x(t),$$

где $x(t)$ — значение электрического поля в приходящем сигнале, $y(t)$ — ток, наблюдаемый в антенне, $a(\phi, \psi)$ — функция, зависящая от конфигурации антенны,

[§]Работа поддержана грантом фонда «Научный потенциал» (Human Capital Foundation) и ФЦП «Кадры» (гос. контракт П940)

азимута ϕ , угла места ψ и поляризации падающего сигнала. Приняв один сигнал на несколько антенн с разными ДН $\alpha_i(\phi, \psi)$, мы можем определить направление, с которого пришла волна (произвести *пеленгацию* сигнала), определив ϕ, ψ из системы

$$y_i(t)/y_j(t) = \alpha_i(\phi, \psi)/\alpha_j(\phi, \psi).$$

В случае, когда на АС одновременно приходит *смесь* из нескольких сигналов, и разделение их в частотной области невозможно (полосы частот сигналов существенно перекрываются), пеленгация на основе таких простых формул невозможна. Таким образом, встает задача о *разделении сигналов*, то есть определении вектора источников $x(t) \in \mathbb{C}^n$ и смешивающей матрицы $A \in \mathbb{C}^{m \times n}$ по наблюдаемым значениям вектора сигналов $y(t) \in \mathbb{C}^m$, $m \geq n$. Все методы разделения основываются на *линейной модели*

$$y(t) = Ax(t) + \xi(t), \quad (1)$$

причем здесь и далее мы будем полагать, что вектора $x(t)$ и $y(t)$ имеют нулевое среднее, и пренебрегать влиянием шумов $\xi(t)$, если не указано обратное. В уравнении (1) известным является только вектор сигналов $y(t)$. Матрица A подлежит определению, даже если функции ДН всех антенн $\alpha_{ij} = \alpha_i(\phi_j, \psi_j)$ известны — пока не решена задача пеленгации, направления ϕ_j, ψ_j прихода волн, и численные величины матричных элементов α_{ij} нам неизвестны. Источники $x_i(t)$ предполагаются *независимыми*, однако при решении задачи о разделении сигналов это условие может быть по-разному использовано и отображено на алгоритм.

В простейшем случае независимость сигналов формулируется как их некоррелированность, то есть диагональность матрицы $F_x = E[x(t)x^*(t)]$, где $E[\cdot]$ — осреднение по времени. Тогда можно записать (пренебрегая шумами)

$$F_y = E[y(t)y^*(t)] = E[Ax(t)x^*(t)A^*] = AE[x(t)x^*(t)]A^* = ADA^*, \quad (2)$$

и мы получаем задачу о диагонализации эрмитовой положительно определенной матрицы F_y . У нее есть ответ в виде собственного разложения $F_y = U\Lambda U^*$, но использовать U вместо A нельзя, так как матрица собственных векторов U унитарна, а смешивающая матрица A — не обязательно. Решение задачи (2) не единственно — в качестве A может быть взята матрица вида

$$A = U\Lambda^{1/2}W\Lambda^{\dagger/2}, \quad (3)$$

где W произвольная унитарная. Определить A по данным только о статистике F_y невозможно. В методе MUSIC (multiple signal classification), который был предложен в 1979 году R. O. Schmidt [4] как развитие метода Писаренко [5], в качестве дополнительной информации используется аналитическое представление ДН $A = A(\phi, \psi)$. Фактически, вычисляется подпространство старших собственных векторов матрицы F_y , соответствующих собственным значениям, превышающим выбранный шумовой порог σ^2 , и направления ϕ, ψ выбираются так, чтобы

подпространство, определяемое столбцами $A(\phi, \psi)$, совпало с найденным подпространством.

Однако, точный аналитический расчет ДН сложных антенн невозможен, так как не удастся пренебречь взаимным влиянием антенн друг на друга, учесть влияние свойств подстилающей поверхности, подводных проводов, сторонних тел в зоне действия АС и других случайных факторов. В связи с этим большую популярность приобрел метод ESPRIT (estimation of signal parameters via rotational invariant techniques), предложенный в 1985 году А. Paulraj, R. Roy и Т. Kailath [6]. Метод ESPRIT не требует точного знания ДН всех элементов АС, но предполагает некоторые соотношения между ее частями. В простейшем случае АС представляет собой две идентичные подсистемы Y и Z , сдвинутые друг относительно друга в пространстве на некоторое расстояние. В итоге плоские волны, пришедшие на подсистемы Y, Z приобретают набег фаз:

$$x_j^{(y)}(t) = x_j^{(z)}(t)e^{i\delta(\phi_j, \psi_j)},$$

где $x_j^{(y)}, x_j^{(z)}$ — значения ЭМ волны от источника j , пришедшие на подсистемы Y и Z , а $\delta(\phi_j, \psi_j)$ — разность хода между двумя элементами подсистемы, зависящая от угла прихода волны от источника j . Сигналы, зарегистрированные на подсистемах, имеют вид

$$y(t) = Ax^{(y)}(t) + \xi(t), \quad z(t) = Ax^{(z)} + \eta(t) = ADx^{(y)}(t) + \eta(t).$$

Здесь A — смешивающая матрица для подсистемы (она нам не потребуется), D — диагональная матрица, содержащая элементы $\delta(\phi_j, \psi_j)$, определив которую, мы определим и интересующие направления ϕ_j, ψ_j . Выписав корреляционные матрицы

$$\begin{aligned} F_y &= E[y(t)y^*(t)] = AF_xA^* + \sigma^2I, \\ F_{zy} &= E[z(t)y^*(t)] = ADF_xA^*, \end{aligned} \tag{4}$$

и решив обобщенную собственную задачу

$$F_y = U\Lambda V^*, \quad F_{zy} = U\Lambda V^*,$$

можно после выделения сигнального подпространства найти D как отношение собственных значений M и Λ .

В последние годы особенный интерес вызывает задача *слепого разделения источников* (blind source separation, BSS), при решении которой информация о размешивающей матрице вообще не используется. Алгоритмы, использующие слепое разделение, не теряют применимости в случае, когда нарушена геометрия АС, сигнал испытывает преломления на пути к АС, часть элементов АС вышла из строя. Однако, алгоритмы BSS существенно сложнее, и не могут быть основаны на анализе одной лишь матрице корреляции.

В самом деле, если не обладать никакой дополнительной информацией о смешивающей матрице A и взять в качестве решения (3) матрицу $A = U$, мы получим предложенный в 1901 году метод *главных компонент* (principal component analysis, PCA) [7], который широко применяется для понижения размерности при статистической обработке больших массивов данных (см. напр., [8, 9, 10]) и по сей день используется для предобработки (так называемого «обеления») входных данных в большинстве методов слепого разделения сигналов, но не позволяет качественно разделить даже модельную пару независимых сигналов «синусоида + меандр» (см. рис. 1). Чтобы уточнить решение, не привлекая дополнительной информации о матрице A , можно рассмотреть другие статистики принятых сигналов, обладающих тем же разложением

$$F_k = AD_kA^*, \quad D_k \text{ — диагональны,} \quad k = 1, \dots, p, \quad (5)$$

учитывая тем самым, что независимость сигналов не ограничивается только их некоррелированностью. Следуя [11], в качестве F_k можно выбрать

- корреляционные матрицы $F_k = E[y(t)y^*(t)]_{t \in [t_k, t_k + \tau]}$, отвечающие разным интервалам времени (для сигналов с сильно меняющейся во времени амплитудой);
- корреляции

$$F_k = E[y(t)y^*(t + \tau_k) + y(t + \tau_k)y^*(t)]/2, \quad (6)$$

параметризованные набором временных сдвигов τ_k (для сигналов с меняющимся спектром);

- статистику высокого порядка, например *кумулянт*

$$(C_y)_{ijkl} = E[y_i \bar{y}_j y_k \bar{y}_l] - E[y_i \bar{y}_j]E[y_k \bar{y}_l] - E[y_i y_k]E[\bar{y}_j \bar{y}_l] - E[y_i \bar{y}_l]E[\bar{y}_j y_k],$$

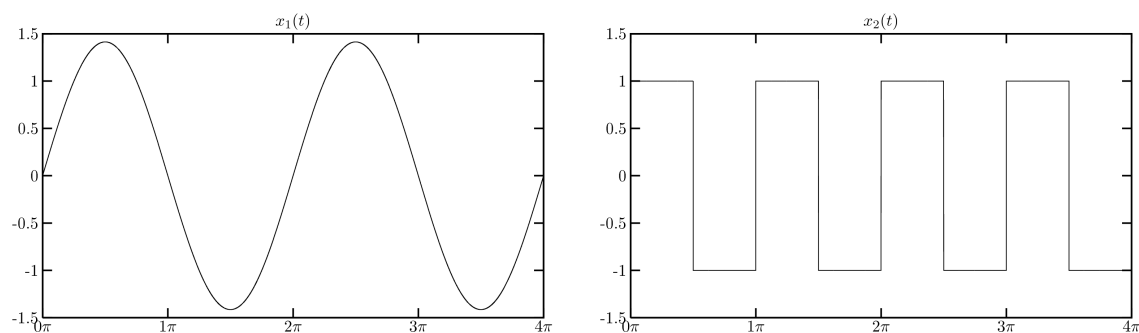
то есть четырехмерный массив размера $m \times m \times m \times m$, который можно считать набором из m^2 матриц размера $m \times m$, для которых выполняется одновременное разложение (5).

На основе BSS решается множество практических задач в самых различных областях. Помимо MIMO-каналов связи можно назвать, например

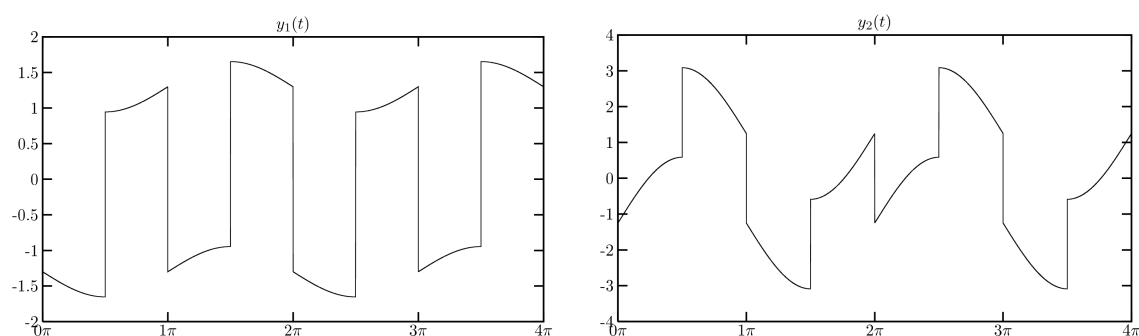
- медицинское исследование функций мозга на основе данных магнитной энцефалографии, магнитно-резонансной томографии или спектроскопии (см., напр. [12, 13, 14, 15, 16]);
- анализ смесей в химии и спектроскопии (см. обзор [17]);
- исследование числовых рядов в финансовой математике [18, 19], социологии, статистике, при анализе спроса;
- контроль транспортных и производственных процессов;

Рис. 1. Классический пример разделения синусоиды и меандра

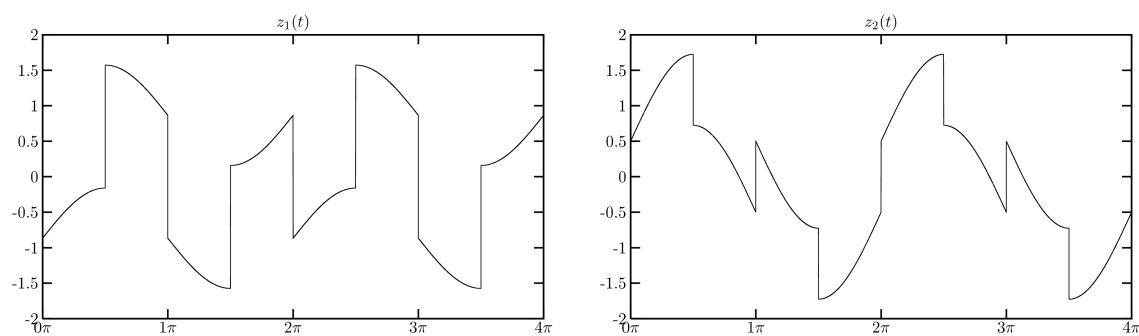
Источники



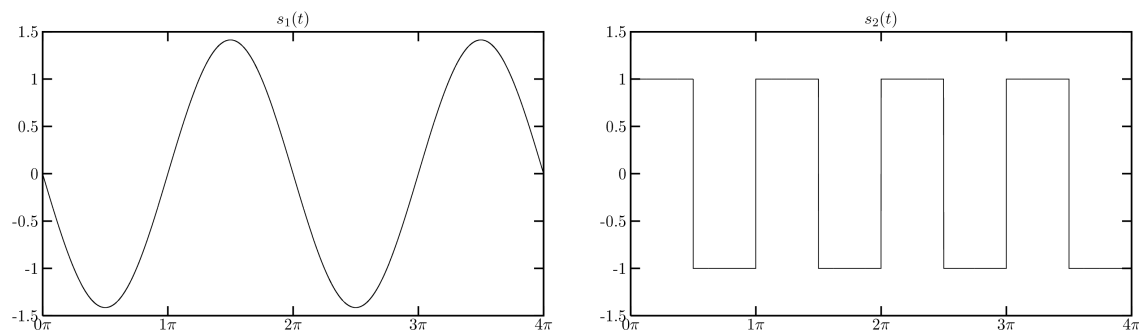
Сигналы (смеси источников)



Главные компоненты



Независимые компоненты



- мультимедийные приложения, распознавание текстов и изображений [20, 21].

В основе слепого разделения лежит одновременная диагонализация нескольких (более, чем двух) матриц — в этом смысле можно считать BSS развитием методов MUSIC и ESPRIT, при котором устранена необходимость использования информации о смешивающей матрице A за счет более тонкого использования свойств независимости источников $x(t)$. Вычислительно задача (5) является существенно более сложной, чем (2) и (4). Похожая задача возникает при обработке трехмерных массивов (*тензоров*) и носит название *трилинейного* или *канонического* разложения тензора, модели CANDECOMP/PARAFAC [22, 23, 24]. Для ее решения предложено множество алгоритмов (см. [23, 24, 25, 26, 27, 28, 29] а также сравнение некоторых методов в [30]) но ни один из них не может считаться наилучшим в общем случае, так как каноническое разложение неустойчиво [31] и успешность вычисления существенно зависит как от используемых алгоритмов, так и от свойств обрабатываемых массивов данных.

В данной работе мы рассмотрим некоторые алгоритмы одновременной диагонализации матриц, применим их для разделения электромагнитных сигналов в эфире и сравним полученные результаты. Мы также обсудим преимущества методов, основанных на решении задачи (5), с популярным методом FastICA [32, 33], в котором разделение сигналов происходит с помощью выбора линейных комбинаций исходных сигналов, максимизирующих степень их *негауссовости*. При этом мы будем полагать, что число источников и приемников *совпадают*, то есть $n = m$ — для выполнения этого условия достаточно, например, произвести обеление методом PCA и перейти от вектора сигналов к вектору главных компонент, отбросив шумовое подпространство.

2. Метод мультипликативного обновления

Метод мультипликативного обновления для одновременной диагонализации действительных матриц (5) был предложен в [34, 35]. Опишем версию этого метода в комплексной арифметике. Итак, пусть имеется p матриц G_k и требуется привести их к максимально диагональному виду:

$$G_k = AM_kA^*,$$

где M_k максимально близки к диагональным. Эту близость будем характеризовать мерой

$$\mathfrak{M} = \sum_{k=1}^p \mathfrak{M}(M_k), \quad \mathfrak{M}(M) = \|\text{off } M\|_F^2, \quad (7)$$

где $\text{off } M := M - \text{diag } M$ означает внедиагональную часть M , а $\text{diag } M$ — диагональ M .

Пусть $m = n$ и существует $W = A^{-1}$, тогда

$$M_k = WG_kW^*, \quad W = \arg \min \sum_{k=1}^p \|\text{off } M_k\|_F^2. \quad (8)$$

Для решения этой задачи будем применять шаг мультипликативного обновления

$$W_{\{t+1\}} = (I + Q_{\{t\}})W_{\{t\}}, \quad (9)$$

тогда в предположении малости величины шага $Q_k^{\{t\}}$ и нормы погрешности $E_k^{\{t\}} = \text{off } M_k^{\{t\}}$, запишем

$$\begin{aligned} M_k^{\{t+1\}} &= (I + Q_{\{t\}})(D_k^{\{t\}} + E_k^{\{t\}})(I + Q_{\{t\}})^* \\ &\approx D_k^{\{t\}} + Q_{\{t\}}D_k^{\{t\}} + D_k^{\{t\}}Q_{\{t\}}^* + E_k^{\{t\}}. \end{aligned}$$

Таким образом, на каждом шаге алгоритма нам потребуется находить

$$Q = \arg \min \sum_{k=1}^p \mathfrak{M}(QD_k + D_kQ^* + E_k).$$

В поэлементной записи это означает, что

$$Q = \arg \min \sum_{k=1}^p \sum_{i \neq j} |q_{ij}d_j^k + d_i^k \bar{q}_{ji} + e_{ij}^k|^2.$$

Таким образом, необходимо в смысле минимальных квадратов решить систему уравнений

$$q_{ij}d_j^k + d_i^k \bar{q}_{ji} \stackrel{\text{ls}}{=} -e_{ij}^k, \quad i \neq j, \quad i, j = 1, \dots, n, \quad k = 1, \dots, p.$$

Рассмотрим те уравнения этой системы, которые отвечают $i > j$. В оставшихся уравнениях формально поменяем местами индексы i и j , а затем сопряжем правую и левую части, получив

$$\bar{q}_{ji}\bar{d}_i^k + \bar{d}_j^k q_{ij} \stackrel{\text{ls}}{=} -\bar{e}_{ji}^k.$$

Таким образом, система преобразована к виду

$$\begin{cases} q_{ij}d_j^k + \bar{q}_{ji}d_i^k \stackrel{\text{ls}}{=} -e_{ij}^k, \\ q_{ij}\bar{d}_j^k + \bar{q}_{ji}\bar{d}_i^k \stackrel{\text{ls}}{=} -\bar{e}_{ji}^k, \end{cases} \quad i > j, \quad i, j = 1, \dots, n, \quad k = 1, \dots, p.$$

В «полуматричном» виде

$$\begin{bmatrix} d_j^k & d_i^k \\ \bar{d}_j^k & \bar{d}_i^k \end{bmatrix} \begin{bmatrix} q_{ij} \\ \bar{q}_{ji} \end{bmatrix} \stackrel{\text{ls}}{=} - \begin{bmatrix} e_{ij}^k \\ \bar{e}_{ji}^k \end{bmatrix};$$

Обозначая

$$D_k = \begin{bmatrix} d_j^k & d_i^k \\ \bar{d}_j^k & \bar{d}_i^k \end{bmatrix}, \quad e_k = \begin{bmatrix} e_{ij}^k \\ \bar{e}_{ji}^k \end{bmatrix}, \quad q = \begin{bmatrix} q_{ij} \\ \bar{q}_{ji} \end{bmatrix},$$

получаем

$$D_k q \stackrel{\text{ls}}{=} -e_k.$$

Действуя согласно методу наименьших квадратов, получаем

$$q = - \left(\sum_{k=1}^p D_k^* D_k \right)^{-1} \left(\sum_{k=1}^p D_k^* e_k \right).$$

Вычисляем

$$D_k^* D_k = \begin{bmatrix} \bar{d}_j^k & d_j^k \\ \bar{d}_i^k & d_i^k \end{bmatrix} \begin{bmatrix} d_j^k & d_i^k \\ \bar{d}_j^k & \bar{d}_i^k \end{bmatrix} = 2 \begin{bmatrix} |d_j^k|^2 & \Re(\bar{d}_j^k d_i^k) \\ \Re(\bar{d}_i^k d_j^k) & |d_i^k|^2 \end{bmatrix};$$

$$D^* D := \sum_{k=1}^p D_k^* D_k = 2 \begin{bmatrix} \|d_j\|^2 & \Re(d_j, d_i) \\ \Re(d_i, d_j) & \|d_i\|^2 \end{bmatrix} = \begin{bmatrix} z_{jj} & z_{ji} \\ z_{ij} & z_{ii} \end{bmatrix}. \quad (10)$$

Заметим, что матрица полностью состоит из действительных элементов. Далее

$$(D^* D)^{-1} = \begin{bmatrix} z_{jj} & z_{ji} \\ z_{ij} & z_{ii} \end{bmatrix}^{-1} = \frac{1}{z_{jj}z_{ii} - z_{ij}z_{ji}} \begin{bmatrix} z_{ii} & -z_{ij} \\ -z_{ji} & z_{jj} \end{bmatrix}.$$

$$D^* e = \sum_{k=1}^p D_k^* e_k = \begin{bmatrix} (d_j, e_{ij}) + (e_{ji}, d_j) \\ (d_i, e_{ij}) + (e_{ji}, d_i) \end{bmatrix} = \begin{bmatrix} y_{ij} \\ \bar{y}_{ji} \end{bmatrix}. \quad (11)$$

Окончательно

$$q_{ij} = \frac{z_{ij}\bar{y}_{ji} - z_{ii}y_{ij}}{z_{jj}z_{ii} - z_{ij}z_{ji}}, \quad \bar{q}_{ji} = \frac{z_{ji}y_{ij} - z_{jj}\bar{y}_{ji}}{z_{jj}z_{ii} - z_{ij}z_{ji}}. \quad (12)$$

Алгоритм 1: Решение (8) методом мультипликативного обновления

Выберем начальное приближение $W_{\{0\}}$

- 1: **for** $t = 0, 1, \dots$ **do**
- 2: $M_k^{\{t\}} := W_{\{t\}} G_k W_{\{t\}}^*$, $D_k^{\{t\}} := \text{diag}(M_k^{\{t\}})$, $E_k^{\{t\}} := \text{off}(M_k^{\{t\}})$
- 3: Найдем $Q_{\{t\}} := \arg \min \sum_{k=1}^p \Re(QD_k^{\{t\}} + D_k^{\{t\}}Q^* + E_k^{\{t\}})$ по формулам (10)–(12)
- 4: **if** $\|Q_{\{t\}}\| > q_{\max}$ **then**
- 5: $Q_{\{t\}} := q_{\max} Q_{\{t\}} / \|Q_{\{t\}}\|$.
- 6: **end if**
- 7: **if** $\|Q_{\{t\}}\| < q_{\min}$ **then**
- 8: **return** $W := W_{\{t\}}$
- 9: **end if**
- 10: $W_{\{t+1\}} = (I + Q_{\{t\}})W_{\{t\}}$
- 11: **end for**

Приведем вышеизложенное кратко в алгоритме 1, параметрами которого являются величины q_{\max} и q_{\min} , ограничивающие норму матрицы $Q_{\{t\}}$ сверху и снизу. При превышении верхнего порога величина шага принудительно уменьшается до значения q_{\max} , если же шаг становится слишком малым, алгоритм завершается.

3. Суперобобщенное разложение Шура

3.1. Одновременное приведение матриц и разложение тензоров

Одновременная диагонализация матриц (5) тесно связана с задачей *канонического разложения* трехмерных массивов (*тензоров*), называемой также моделью

CANDECOMP/PARAFAC. Каноническое разложение тензора $\mathcal{G} = [g_{ijk}]$ имеет вид

$$g_{ijk} = \sum_{\alpha} u_{i\alpha} v_{j\alpha} w_{k\alpha}, \quad (13)$$

а выписав его относительно каждой «срезки» $G_k = [g_{ij,k}]$, получаем

$$G_k = U W_k V^T, \quad U = [u_{i\alpha}], \quad V = [v_{j\alpha}], \quad W_k = \text{diag}(w_{k,:}),$$

где $\text{diag}(w_{k,:})$ — диагональная матрица, определенная строкой $w_{k,:}$. При дополнительных ограничениях $U = A$, $V = \bar{A}$ мы получили задачу (5). Для построения (13) предложено множество алгоритмов, однако в силу формальной неустойчивости канонического разложения ни один из методов не является абсолютно надежным. Численная неустойчивость обычно проявляется в том, что факторы U, V, W в (13) становятся почти вырожденными. Поэтому задачу (13) иногда решают в два этапа:

1. Найти суперобобщенное разложение Шура, то есть унитарные матрицы Y, Z , что

$$Y G_k Z = X_k, \quad \sum_{k=1}^p \|X_k\|_{F(\text{low})}^2 \rightarrow \min, \quad Y^* Y = I, \quad Z^* Z = I. \quad (14)$$

Здесь $\|\cdot\|_{F(\text{low})}$ означает фробениусову норму нижнетреугольной части матрицы, то есть задача состоит в одновременном приведении матриц G_k к верхнетреугольному виду.

2. Найти верхнетреугольные матрицы S, T , такие что

$$X_k = S D_k T + E_k, \quad \sum_{k=1}^p \|E_k\|_F^2 \rightarrow \min. \quad (15)$$

Задача (15) тривиально выражается через решение серии переопределенных линейных систем [36]. Для решения задачи (14) известно несколько методов. Первый, очень общий подход, предложен в [37] и требует интегрирования большой и часто очень жесткой системы обыкновенных дифференциальных уравнений на большом промежутке времени, что вызывает вычислительные трудности. Вторым подход, предложенный в [36], состоит в использовании вращений Якоби, при котором уже насчитанные матрицы Y и Z умножаются на матрицы вращений Y_i, Z_i , чтобы на каждом шаге минимизировать $\sum_{k=1}^p \|Y_i(Y A_k Z) Z_i\|_{F(\text{low})}^2$, для чего необходимо найти корень полинома восьмой степени, что является достаточно неустойчивой задачей. Самым эффективным методом вычисления обобщенного разложения Шура (при $p = 2$) является QZ-алгоритм. Его аналог для случая $p > 2$ был построен в [38], однако обобщить на случай многих матриц метод сдвигов, делающий QZ-алгоритм таким эффективным, по-видимому, не удастся.

В этой работе мы предложим комплексную версию алгоритма решения задачи (14), предложенную недавно в [29]. Задача (14) будет сведена к набору суперобобщенных задач на собственные значения, для которых мы предложим версию метода Гаусса-Ньютона.

3.2. Одновременное приведение матриц к треугольному виду

Итак, рассмотрим задачу (14) для p матриц G_k размера $n \times n$. Найдем сначала унитарные матрицы Y и Z такие, что матрицы YG_iZ перейдут в блочно-треугольные матрицы

$$YG_kZ = \begin{pmatrix} x_k & x \\ o & H_k \end{pmatrix}, \quad (16)$$

где H_k — матрицы размера $(n-1) \times (n-1)$, а $\|o\| \rightarrow \min$. Решив (16), мы можем перейти к аналогичной задаче для H_k , т.е. (16) описывает шаг *дефляции*, а $n-1$ таких шагов приводят матрицы G_k к треугольному виду. Задача (16) эквивалентна следующей системе

$$G_k z \stackrel{\text{ls}}{=} x_k y^*, \quad (17)$$

где z — первый столбец Z , а y — первая строка Y . При $p = 2$ система (17) становится обобщенной задачей на собственные значения для пары матриц, поэтому задачу (17) естественно называть *суперобобщенной задачей на собственные значения*. Заметим, что поскольку на каждом шаге дефляции унитарные матрицы Y, Z определены только своим первым столбцом или строкой, причем с точностью до комплексного множителя единичной нормы, мы можем не ограничивая общности искать Y, Z как матрицы вращений [39]

$$Y = I - 2uu^*, \quad Z = I - 2vv^*, \quad \|u\| = \|v\| = 1.$$

Итак, для решения задачи (17) предложен некоторый метод, то алгоритм вычисления суперобобщенного разложения Шура записывается следующим образом (см. алг. 2).

Алгоритм 2: Суперобобщенное разложение Шура (14)

Инициализация: $m = n$, $Y = Z = I$.

1: while $m > 1$ do

2: Решить суперобобщенную задачу на собственные значения (17)

3: Найти матрицы вращений $Y', Z' \in \mathbb{C}^{m \times m}$, такие что $(Y')^* e_1 = uy^*$, $Z' e_1 = \zeta z$, где $|u| = |\zeta| = 1$.

4: Вычислить новые активные подматрицы H_k в матрицах $Y' G_k Z'$ согласно (16)

5: Обновить

$$Y := \begin{pmatrix} I & 0 \\ 0 & Y' \end{pmatrix} Y, \quad Z := Z \begin{pmatrix} I & 0 \\ 0 & Z' \end{pmatrix}.$$

6: Установить $m := m - 1$, $G_k = H_k$.

7: end while

3.3. Супербобщенная задача на собственные значения

Уравнения (16) в поэлементной записи $\sum_{j=1}^n g_{ijk} z_j \stackrel{\text{ls}}{=} x_k \bar{y}_i$ превращаются в систему

$$\sum_{j=1}^n z_j G_j \stackrel{\text{ls}}{=} x y^*, \quad (18)$$

где матрицы-срежки $G_j = [(g_{ki})_j] \in \mathbb{C}^{p \times n}$ составлены из элементов $(g_{ki})_j = g_{ijk}$. Перед нами стоит задача подобрать веса x_j линейной комбинации матриц G_j , которая даст нам матрицу, наиболее близкую к матрице единичного ранга. Уравнения (18) можно рассматривать как переопределенную систему нелинейных уравнений и применить для их решения метод Гаусса-Ньютона. Линеаризация уравнений (18) около некоторой точки (x, y, z) приводит к системе

$$\sum_{j=1}^n \hat{z}_j G_j \stackrel{\text{ls}}{=} \Delta x y^* + x \Delta y^* + x y^*, \quad (19)$$

где $\hat{z} = z + \Delta z$ — новое значение вектора z . Найдем такие матрицы вращения $X \in \mathbb{C}^{p \times p}$ и $Y \in \mathbb{C}^{n \times n}$, что

$$X e_1 = \chi x, \quad Y e_1 = \nu y, \quad |\chi| = |\nu| = 1. \quad (20)$$

Умножая (19) на X^* слева и на Y справа, получаем

$$\sum_{j=1}^n \hat{z}_j (X^* G_j Y) \stackrel{\text{ls}}{=} x' e_1^* + e_1 (y')^*,$$

где $x' = \nu X \Delta x + \bar{\chi} e_1 \in \mathbb{C}^p$, а $y' = \bar{\chi} Y^* \Delta y \in \mathbb{C}^n$. Поскольку X, Y унитарны, количество степеней свободы у векторов x', y' то же самое, что у Δx и Δy . Поэтому при любых значениях \hat{z} первый столбец и первая строка матрицы $F = \sum_{j=1}^n \hat{z}_j (X^* G_j Y)$ в левой части системы могут быть представлены правой частью *точно* за счет выбора $x' = f_{:,1}$ и $(y')^* = f_{1,:}$. Таким образом, коэффициенты \hat{z} надо выбирать так, чтобы минимизировать

$$\left\| \sum_{j=1}^n \hat{z}_j (X^* G_j Y)_{[2:p, 2:n]} \right\|_F^2 \rightarrow \min, \quad \|\hat{z}\| = 1, \quad (21)$$

то есть норму подматрицы, полученной из F отбрасыванием первой строки и столбца. Соберем вышеизложенное в алгоритме 3, параметром которого является желаемая точность ε решения задачи (17).

Алгоритм 3: Суперобобщенная задача на собственные значения в виде (18)

- 1: **repeat**
 - 2: Вычислить SVD матрицы $G := \sum_{j=1}^n z_j G_j$. Найти наилучшее приближение единичного ранга $G = xy^* + E$, $\|y\| = 1$ и погрешность $\text{err} := \|E\|_F$.
 - 3: Вычислить матрицы вращения X, Y согласно (20).
 - 4: Построить массив $\mathcal{F} = [f_{kij}]$ с элементами $f_{kij} := (X^* G_j Y)_{ki}$, причем $k = 2, \dots, p$, $i = 2, \dots, n$, $j = 1, \dots, n$.
 - 5: Рассмотреть $\mathcal{F} = [f_{kij}]$ как матрицу $[f_{ki,j}] \in \mathbb{C}^{(p-1)(n-1) \times n}$ и найти ее младший правый сингулярный вектор \hat{z} . Положить $z := \hat{z}$.
 - 6: **until** $\text{err} \leq \varepsilon$
-

3.4. Одновременное приведение треугольных матриц к диагональным

Укажем теперь для замкнутости изложения способ решения задачи (15), что завершит описание метода вычисления канонического разложения на основе суперобобщенного разложения Шура. Система (15) в поэлементном виде равносильна

$$x_{ijk} \stackrel{\text{ls}}{=} \sum_{i \leq l \leq j} s_{il} d_{lk} t_{lj}, \quad (22)$$

причем не ограничивая общности можно положить $s_{ii} = t_{ii} = 1$ для всех $i = 1, \dots, n$. Тогда рассматривая (22) как последовательность переопределенных линейных систем при $j = i$, $j = i + 1$, $j = i + 2, \dots$, получаем

$$\begin{array}{lll} [j = i] & b_{iik} = s_{ii} \underline{d_{ik}} t_{kj} & \Rightarrow d_{ik} = b_{iik} \\ [j = i + 1] & b_{ij+1k} \stackrel{\text{ls}}{=} s_{ii} \underline{d_{ik}} \underline{t_{ii+1}} + \underline{s_{ii+1}} d_{i+1k} t_{i+1i+1} \\ [j = i + 2] & b_{ij+2k} \stackrel{\text{ls}}{=} s_{ii} \underline{d_{ik}} \underline{t_{ii+2}} + \underline{s_{ii+1}} d_{i+1k} t_{i+1i+2} + \underline{s_{ii+2}} d_{i+2k} t_{i+2i+2} \end{array}$$

В каждой строке подчеркнуты те неизвестные, которые будут определены на текущем шаге, остальные величины определены на предыдущих шагах.

4. Численные эксперименты

Рассмотрим сонограмму сигнала, принятого на антенной системе, состоящей из трех взаимно ортогональных элементов (см. рис. 2). На ней указано распределение спектральной плотности энергии (яркие точки — много энергии, темные — мало) в спектральной полосе шириной порядка 10КHz в зависимости от времени (сверху вниз). Вид сонограммы характерен для цифровой передачи стандарта DRM с модуляцией ODFM. Однако под цифровым сигналом «спрятан» аналоговый сигнал с амплитудной модуляцией — о нем свидетельствует тонкая «несущая» полоса в центре спектра. Выделить АМ-составляющую из-под цифровой компоненты и распознать речь невооруженным ухом не представляется реальным. На помощь приходит методика слепого разделения сигналов.

Рис. 2. Сонограмма исходного сигнала

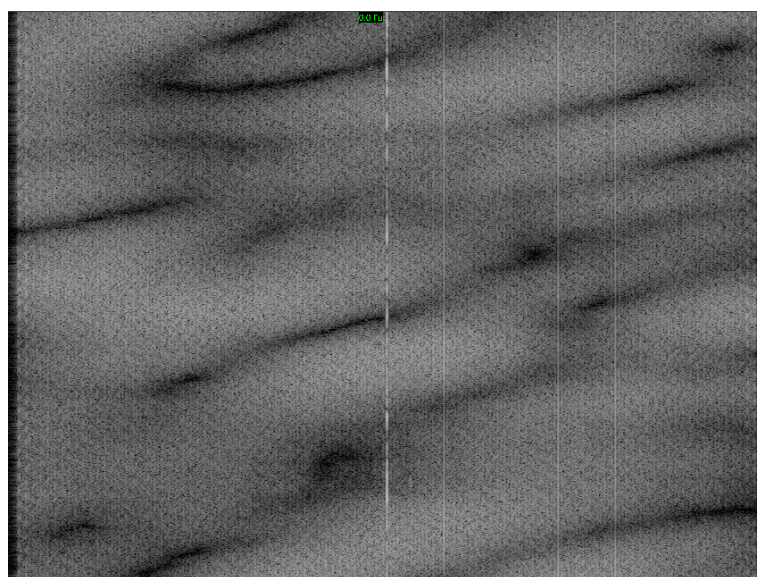
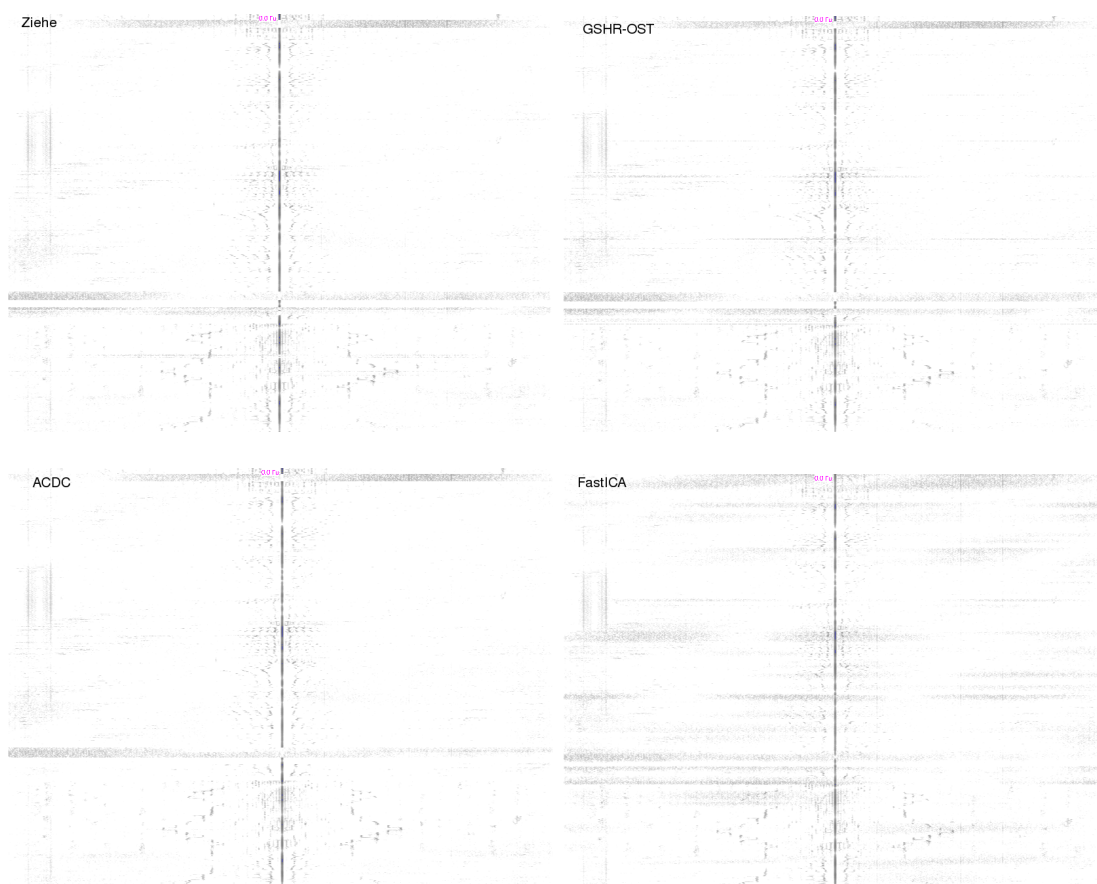


Рис. 3. Сонограмма отделенной АМ-компоненты (негатив)



Результат разделения представлен на рис. 3 в негативе (белые точки — мало энергии, черные — много). Мы использовали три тензорных алгоритма: комплексные версии алгоритма мультипликативного обновления (Ziehe) и алгоритма суперобобщенного разложения Шура (GSHR-OST), а также известный метод ACDC [40], предложенный в 2002 году А. Yeredor. С помощью этих алгоритмов выполнялось одновременное приведение к диагональному виду матриц-статистик вида (6). Кроме того, мы использовали популярный среди инженеров вычислителей алгоритм FastICA [33], основанный на максимизации меры негауссовости независимых компонент.

По сонограммам на рис. 3 мы видим, что все методы показали неплохое качество разделения — АМ-сигнал после их применения можно легко прослушать, кроме того, становится различима слабая, вероятно цифровая, компонента в составе сигнала от той же вещательной станции (две полосы в левой части сонограммы). Замечательным является то, что разделение, полученное с помощью тензорных методов, оказалось несколько более «чистым», чем разделение с помощью FastICA.

Следует отметить также, что тензорные методы разделения, обрабатывающие массивы сравнительно небольшого размера (зависящего только от числа антенн, но не зависящего от количества отсчетов по времени) лучше подходят для решения задач разделения и пеленгации электромагнитных сигналов в следующих случаях.

- В условиях пакетной обработки данных — когда сигнал обрабатывается небольшими пачками, размер которых ограничен оперативной памятью вычислительного устройства и/или временем устойчивости канала между источником и приемником сигнала.
- При использовании специальных процессоров (ПЛИС, GPU и др.) для предобработки данных — в этом случае можно насчитывать статистики (6) и аналогичные им на специализированной ПЛИС-плате, размещенной после АЦП, обмениваясь со стандартным процессором только массивами маленького размера (см. рис. 4).

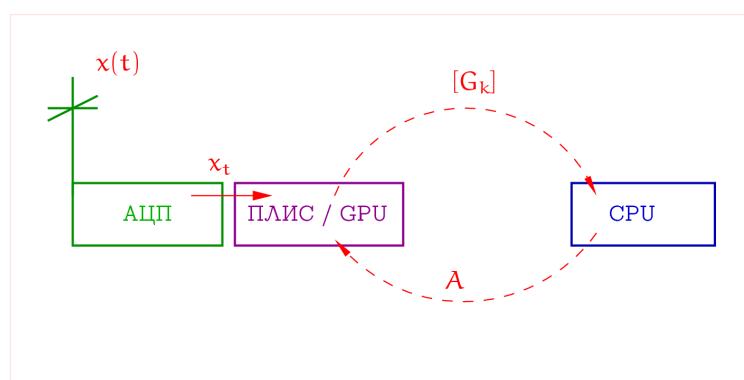
Это дает основание продолжать работу над развитием тензорных методов и их применением к слепому разделению сигналов.

Благодарности

Работа поддержана грантом фонда «Научный потенциал» (Human Capital Foundation) и ФЦП «Кадры» (гос. контракт П940).

Автор благодарен А. А. Кучумову (МТУСИ) за сигналы, предоставленные для проведения численных экспериментов, а также множество полезных советов в ходе работы и помощь в интерпретации результатов.

Рис. 4. Схема вычислительной платформы



Список литературы

- [1] Chen B., Petropulu A. P. Frequency domain MIMO system identification based on second and higher-order statistics // *IEEE Trans. Signal Processing*. 2001. V. 49, № 8. P. 1677-1688.
- [2] de Lathauwer L., Comon P., de Moor B., Vandewalle J. ICA algorithms for 3 sources and 2 sensors // *Proc. of the IEEE Signal Processing Workshop on Higher-Order Statistics (HOS'99)*. — 1999. — P. 116-120.
- [3] de Lathauwer L., de Moor B., Vandewalle J. An algebraic approach to blind MIMO identification // *Proc. of the 2nd Int. Workshop on independent component analysis and blind source separation (ICA)*. — 2000. — P. 211-214.
- [4] Schmidt R. O. Multiple emitter location and signal parameter estimation // *IEEE Transactions on Antennas and Propagation*. 1986. V. 34, № 3. P. 276-280.
- [5] Pisarenko V. F. The retrieval of harmonics from a covariance function // *Geophysics, J. Roy. Astron. Soc.* 1973. V. 33. P. 347-366.
- [6] Roy R., Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques // *IEEE Trans. on Acoustics Speech and Signal Processing*. 1989. V. 37, № 7. P. 984-995.
- [7] Pearson K. On lines and planes of closest fit to systems of points in space // *Philosophical Magazine*. 1901. V. 2, № 6. P. 559-572.
- [8] Hotelling H. Analysis of a Complex of Statistical Variables with Principal Components // *Journal of Educational Psychology*. 1933. V. 24. P. 417-441.
- [9] Sato T. Application of principal-component analysis on near-infrared spectroscopic data of vegetable oils for their classification // *Journal of the American Oil Chemists' Society*. 1994. V. 71, № 3. P. 293-298.

- [10] Aiyi L., Ying Z., Gehan E., Clarke R. Block principal component analysis with application to gene microarray data classification // *Statistics in Medicine*. 2002. V. 21. P. 3465-3474.
- [11] Parra L., Sajda P. Blind Source Separation via Generalized Eigenvalue Decomposition // *J. of Machine Learning Research*. 2003. V. 4. P. 1261-1269.
- [12] Hämäläinen M., Hari R., Ilmoniemi R. et al. Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of signal processing in the human brain // *Reviews of Modern Physics*. 1993. V. 65. P. 413-497.
- [13] de Lathauwer L., de Moor B., Vandewalle J. Fetal electrocardiogram extraction by source subspace separation // *Proc. of IEEE Signal Processing / Athos Workshop on Higher-Order Statistics*. — 1995. — P. 134-138.
- [14] Vigário R. N. Extraction of ocular artifacts from EEG using independent component analysis // *Electroenceph. clin. Neurophysiol.* 1997. V. 103. P. 395-404.
- [15] Vigário R., Särelä J., Oja E. Independent component analysis in wave decomposition of auditory evoked fields // *Proc. Int. Conf. on Artificial Neural Networks (ICANN'98)*. — 1998. — P. 287-292.
- [16] Xie L., Jun W. Global optimal ICA and its application in MEG data analysis // *Neurocomputing*. 2006. V. 69. P. 2438-2442.
- [17] Bro. R. Review on Multiway Analysis in Chemistry — 2000–2005 // *Analytical Chemistry*. 2006. V. 36. P. 279.
- [18] Back A. D., Weigend A. S. What drives stock returns? — An independent component analysis // *Proc. of the IEEE/IAFE/INFORMS Conf. on Comp. Intelligence for Financial Engineering (CIFEr)*. 1998. P. 141-156.
- [19] Moody J., Howard Y. Term structure of interactions of foreign exchange rates // *Computational Finance — Proc. of the 6th Int'l Conf.* — 1999. — P. 24-35.
- [20] Draper B. A., Baek K., Bartlett M. S., Beveridge J. R. Recognizing faces with PCA and ICA // *CVIU*. 2003. V. 91, №1-2. P. 115-137.
- [21] Jongsun K., Jongmoo C., Yuneho Y. Biometric Authentication. — Berlin: Springer, 2004.
- [22] Hitchcock F. L. The expression of a tensor or a polyadic as a sum of products // *J. Math. Phys.* 1927. V. 6, № 1. P. 164-189.
- [23] Harshman R. A. Foundations of the Parafac procedure: models and conditions for an explanatory multimodal factor analysis // *UCLA Working Papers in Phonetics*. 1970. V. 16. P. 1-84.

- [24] *Caroll J. D., Chang J. J.* Analysis of individual differences in multidimensional scaling via n-way generalization of Eckart-Young decomposition // *Psychometrika*. 1970. V. 35. P. 283-319.
- [25] *Bro R.* PARAFAC: Tutorial and applications // *Chemometrics and Intelligent Lab. Syst.* 1997. V. 38, № 2. P. 149-171.
- [26] *Comon P.* Tensor decomposition: state of the art and applications // IMA Conf. Math. in Sig. Proc., Warwick, UK. — 2000.
- [27] *Grasedyck L.* Existence and computation of low kronecker-rank approximations for large systems in tensor product structure // *Computing*. 2004. V. 72. P. 247-265.
- [28] *de Lathauwer L., de Moor B., Vandewalle J.* Computing of Canonical decomposition by means of a simultaneous generalized Schur decomposition // *SIAM J. Matrix Anal. Appl.* 2004. V. 26. P. 295-327.
- [29] *Oseledets I. V., Savostyanov D. V., Tyrtysnikov E. E.* Fast simultaneous orthogonal reduction to triangular matrices // *SIAM J. Matrix Anal. Appl.* 2009. V. 31, № 2. P. 316-330.
- [30] *Оселедец И. В., Савостьянов Д. В.* Минимизационные методы аппроксимации тензоров и их сравнение // *ЖВМиМФ*. 2006. Т. 46, № 10. С. 1752-1734.
- [31] *de Silva V., Lim L.* Tensor Rank and the Ill-Posedness of the Best Low-Rank Approximation Problem // *SIAM Journal on Matrix Analysis and Applications*. 2008. V. 30, № 3. P. 1084-1127.
- [32] *Hyvärinen A.* Fast and robust fixed-point algorithm for independent component analysis // *IEEE Trans. on Neural Network*. 1999. V. 10, № 3. P. 626-634.
- [33] *Bingham E., Hyvärinen A.* A fast fixed-point algorithm for independent component analysis of complex-valued signals // *Int. J. of Neural Systems*. 2000. V. 10, № 1. P. 1-8.
- [34] *Ziehe A., Laskov P., Müller K.-R., Nolte G.* A linear least-squares algorithm for joint diagonalization // Proc. 4th Intern. Symp. on Independent Component Analysis and Blind Signal Separation (ICA2003). — 2003. — P. 469-474.
- [35] *Ziehe A., Kawanabe M., Hamerling S., Müller K.-R.* A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation // *Journal of Machine Learning Research*. 2004. V. 5. P. 801-818.
- [36] *de Lathauwer L., de Moor B., Vanderwalle J.* Computation of the canonical decomposition by means of a simultaneous generalized Schur decomposition // *SIAM J. Matrix Anal. Appl.* 2004. V. 26. P. 295-227.

- [37] *Chu M.* A continues Jacobi-like approach to the simultaneous reduction of real matrices // *Lin. Alg. Appl.* 1991. V. 147. P. 75-96.
- [38] *Van der Veen A.-J., Paulraj A.* An analytical constant modulus algorithm // *IEEE Trans. Signal Process.* 1996. V. 44. P. 1136-1155.
- [39] *Mueller D.* Householder's method for complex matrices and eigensystems of hermitian matrices // *Numerische Mathematik.* 1966. V. 8, № 1. P. 72–92.
- [40] *Yeredor A.* Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation // *IEEE Trans. on Signal Proc.* 2002. V. 50, № 7. P. 1545–1553.