

---

**ОБЩИЕ  
ЧИСЛЕННЫЕ МЕТОДЫ**

---

УДК 512.643.8

**ПОИСК РАЗРЕЖЕННЫХ РЕШЕНИЙ ДЛЯ СВЕРХБОЛЬШИХ СИСТЕМ, ОБЛАДАЮЩИХ ТЕНЗОРНОЙ СТРУКТУРОЙ<sup>1)</sup>**

© 2022 г. Д. А. Желтков<sup>1,\*</sup>, Н. Л. Замарашкин<sup>1,\*\*</sup>, С. В. Морозов<sup>1,\*\*\*</sup>

<sup>1</sup> 119333 Москва, ул. Губкина, 8, Институт вычислительной математики им. Г.И. Марчука РАН, Россия

\*e-mail: dmitry.zheltkov@gmail.com

\*\*e-mail: nikolai.zamarashkin@gmail.com

\*\*\*e-mail: stanis-morozov@yandex.ru

Поступила в редакцию 30.12.2021 г.  
Переработанный вариант 06.06.2022 г.  
Принята к публикации 07.07.2022 г.

Задача поиска разреженного решения для больших систем линейных уравнений возникает во многих приложениях, связанных с обработкой сигналов. В некоторых случаях размеры возникающих систем оказываются столь велики, что известные методы становятся неэффективными. Решение таких систем возможно только при наличии в них дополнительной структуры. В настоящей работе предлагается эффективный метод поиска разреженных решений сверхбольших систем линейных уравнений, обладающих тензорной структурой определенного вида. Приведенный теоретический анализ и экспериментальные результаты позволяют судить об эффективности предложенного метода. Библиография: 14. Фиг. 7.

**Ключевые слова:** метод наименьших квадратов, разреженное решение, тензорная структура оператора.

**DOI:** 10.31857/S0044466922110151

## 1. ВВЕДЕНИЕ

Восстановление разреженного сигнала по небольшому числу линейных измерений (возможно, искаженных аддитивным шумом) является фундаментальной задачей в обработке сигналов. Остановимся на следующей модели

$$y = X\beta + \varepsilon,$$

где  $\beta \in \mathbb{R}^n$  – сигнал,  $y \in \mathbb{R}^m$  – вектор наблюдений,  $X \in \mathbb{R}^{m \times n}$  – матрица измерений, а  $\varepsilon \in \mathbb{R}^m$  – вектор шума. В некоторых прикладных задачах интерес представляет ситуация, когда размерность сигнала  $n$  велика и существенно превосходит число измерений  $m$  (т.е.  $n \gg m$ ), а вектор сигнала  $\beta$  разреженный (имеет малое число ненулевых компонент).

Обозначим множество ненулевых компонент вектора  $\beta \in \mathbb{R}^n$  через

$$\text{supp } \beta = \{i : \beta_i \neq 0\}$$

и будем искать  $k$ -разреженное решение  $\beta$  задачи наименьших квадратов

$$\|y - X\beta\|_2 \rightarrow \min, \quad \|\beta\|_0 = |\text{supp } \beta| \leq k. \quad (1)$$

Среди методов, используемых для решения (1), наиболее популярным является алгоритм ОМР (Orthogonal Matching Pursuit [1]). Сложность ОМР оценивается как  $\mathcal{O}(mnk)$ . Это приемлемо при  $n$ , сравнимых с  $10^6$ , но уже недостаточно эффективно при  $n$  порядка  $10^9$ .

В данной работе основное внимание уделяется решению сверхбольших систем линейных уравнений, в которых  $n$  имеет порядок  $10^{13}$  и более. Очевидно, что при таких размерах линейной

<sup>1)</sup>Работа выполнена при финансовой поддержке Московского центра фундаментальной и прикладной математики (Соглашение с Минобрнауки РФ № 075-15-2019-1624).

системы требуются дополнительные предположения о структуре задачи. В работе таких предположений три:

1) пространство параметров тензоризовано  $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \dots \times \mathbb{R}^{n_d}$ ,  $n = n_1 \times n_2 \times \dots \times n_d$ , а действие матрицы  $X$  на пространстве  $\mathbb{R}^n$  соответствует действию линейного оператора  $\mathcal{A}$  на тензоризованном пространстве. Относительно структуры оператора  $\mathcal{A}$  будем предполагать, что он допускает возможность эффективного применения к тензорам ранга 1;

2) нормальное решение задачи наименьших квадратов

$$v = \arg \min_u \|\mathcal{A}u - y\|_2$$

может быть приближено в  $\mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_d}$  тензорами малого тензорного ранга;

3) строки матрицы измерений задают отображение ограниченной изометрии из пространства размерности  $m$  в пространство размерности  $n$

$$(1 - \delta)(h, l) < (\mathcal{A}^T h, \mathcal{A}^T l) < (1 + \delta)(h, l), \quad \text{где } h, l \in \mathbb{R}^m. \quad (2)$$

В работе изучается метод, позволяющий при изложенных выше предположениях находить разреженное решение для систем со сверхбольшой размерностью пространства параметров сигнала  $n$ . Новый метод является адаптацией метода ОМР, в которой учитывается структура задачи. По сути, предлагается новый алгоритм выбора существенных (ненулевых) компонент разреженного решения. Реализованная идея выглядит несколько парадоксально. Поиск существенных компонент разреженного решения (1) осуществляется с помощью анализа компонент нормального решения. На первый взгляд может показаться, что в алгоритме происходит замена задачи меньшей вычислительной сложности на задачу большей вычислительной сложности. Это, как будет показано далее, не так, если допустить существование тензорной структуры в нормальном решении и свойства квазиизометрии оператора измерений  $\mathcal{A}$ .

Оставшаяся часть текста организована следующим образом. В разд. 2 рассмотрен ОМР алгоритм поиска разреженных решений. В разд. 3 дается описание предлагаемого метода, который мы называем *тензоризованным ОМР*. В разд. 4 приведен теоретический анализ различных этапов нового метода. Разд. 5 содержит результаты численных экспериментов, сравнение работы классического и тензоризованного алгоритмов, а также ряд дополнительных эвристик, позволяющих повысить качество работы предлагаемого метода. Разд. 6 содержит заключительные замечания.

## 2. АЛГОРИТМ ОМР

Одним из наиболее эффективных методов поиска разреженных решений для задачи наименьших квадратов является Orthogonal Matching Pursuit (ОМР) алгоритм [1]. Это жадная итерационная процедура [2], получающая значительное внимание благодаря простоте записи алгоритма и его эффективности в большом числе реальных приложений [3]–[6]. На каждом шаге ОМР увеличивает список выбранных компонент на одну так, чтобы добиться наибольшего падения невязки. Стандартное описание ОМР метода приводится ниже в алгоритме 1.

---

### Алгоритм 1. Orthogonal Matching Pursuit

---

**Вход:** матрица  $A \in \mathbb{R}^{m \times n}$ , вектор правой части  $b \in \mathbb{R}^m$ , порог ошибки  $\varepsilon_0$ .

**Инициализация:**

- $k = 0$ ;
- $x^0 = 0$ ;
- $r^0 = b - Ax^0 = b$ ;
- множество выбранных столбцов  $S^0 = \emptyset$ ;

**while**  $\|r^k\|_2 \geq \varepsilon_0$  **do**

1. Вычислить ошибки для каждого из столбцов  $\epsilon(j) = \min_{z_j} \|a_j z_j - r^k\|_2^2$  для всех  $j$ . Решением задачи минимизации является  $z_j^* = \frac{a_j^\top r^k}{\|a_j\|_2^2}$ .
  2. Найти столбец  $j_0$ , сильнее всего уменьшающий невязку, а именно  $j_0 : \forall j \notin S^k, \epsilon(j_0) \leq \epsilon(j)$ . Затем обновляем множество выбранных элементов:  $S^{k+1} = S^k \cup \{j_0\}$ .
  3. Найти оптимальные коэффициенты при столбцах из  $S^{k+1}$ , т.е. вычислить  $x^{k+1}$ , минимизирующий  $\|Ax - b\|_2^2$  при условии  $\text{supp } x^{k+1} = S^{k+1}$ .
  4. Вычислить невязку  $r^{k+1} = b - Ax^{k+1}$ .
  5.  $k = k + 1$ .
- end while**
- return**  $S^k$

Сложность итерации ОМР оценивается через  $\mathcal{O}(mn)$  арифметических операций. При размерах  $n$ , по порядку больших  $10^9$ , алгоритм становится неэффективным. Цель данной работы адаптировать ОМР алгоритм на случай больших  $n$ , при условии, что пространство  $\mathbb{R}^n$  тензоризовано, а в операторе  $\mathcal{A}$  системы и нормальном решении задачи наименьших квадратов присутствует тензорная структура.

### 3. ОБЩАЯ СХЕМА ТЕНЗОРИЗОВАННОГО ОМР

Рассмотрим линейный оператор

$$\mathcal{A} : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \dots \times \mathbb{R}^{n_d} \rightarrow \mathbb{R}^m.$$

Будем считать, что столбцы матрицы оператора  $\mathcal{A}$  имеют одинаковую длину. От оператора потребуем возможности быстрого умножения на векторы, представимые тензорами ранга 1, т.е. на векторы вида  $v_1 \otimes v_2 \otimes \dots \otimes v_d$ , где  $v_i \in \mathbb{R}^{n_i}$  (здесь  $\otimes$  обозначает кронекерово произведение).

Кроме того, предположим, что нормальное решение задачи наименьших квадратов

$$v = \arg \min_u \|\mathcal{A}u - y\|_2$$

приближается (возможно довольно грубо) тензором малого ранга. Воспользуемся этим предположением, чтобы снизить сложность стандартного ОМР.

Начнем со следующего простого наблюдения: наиболее трудоемкий этап ОМР относится к вычислению невязок

$$\epsilon(j) = \min_{z_j} \|a_j z_j - r\|_2 \quad (3)$$

для всех столбцов  $\mathcal{A}$ . Действительно, для  $j$ -го столбца оптимальное значение  $z_j^*$  дается выражением  $z_j^* = \frac{a_j^\top r}{\|a_j\|_2^2}$ , где  $r$  обозначает текущий вектор невязки. Прямое вычисление  $z_j^*$  имеет сложность  $\mathcal{O}(m)$ . Всего таких вычислений  $n$ . Таким образом, только прямое вычисление всех  $z_j^*$  уже имеет сложность  $\mathcal{O}(mn)$ .

Поскольку  $a_j z_j^*$  и  $a_j z_j^* - r$  ортогональны, то решение задачи (3) эквивалентно поиску максимума среди величин

$$c_j = \frac{(a_j, y)}{\|a_j\|_2}$$

с последующим выбором столбца с максимальным  $c_j$ . Учитывая, что все длины столбцов  $\|a_j\|_2$  совпадают, наилучшая компонента на текущей итерации ОМР дается максимумом модуля скалярного произведения  $|(a_j, y)|$ . Таким образом, снизить алгоритмическую сложность ОМР возможно, предъявив алгоритм малой сложности для вычисления величин  $|(a_j, y)|$ .

Сделаем элементарное, но важное наблюдение. Для оператора  $\mathcal{A}^T$ , обладающего свойством ограниченной изометрии (см. предположение 3), оценки на величины  $|(a_j, y)|$  можно получить, рассматривая нормальное решение задачи наименьших квадратов. Действительно, нормальное решение  $\tilde{y}$  представляется в виде

$$\tilde{y} = \mathcal{A}^\dagger y = \mathcal{A}^T (\mathcal{A}\mathcal{A}^T)^{-1} y. \tag{4}$$

Тогда для компоненты  $\tilde{y}_i$

$$|\tilde{y}_i| = |a_i^T (\mathcal{A}\mathcal{A}^T)^{-1} y| = |(a_i, (\mathcal{A}\mathcal{A}^T)^{-1} y)| \leq \frac{1}{1-\delta} |(a_i, \mathcal{A}^T (\mathcal{A}\mathcal{A}^T)^{-1} y)| = \frac{1}{1-\delta} |(a_i, y)|.$$

Аналогично получается оценка снизу

$$|\tilde{y}_i| \geq \frac{1}{1+\delta} |(a_i, y)|.$$

Имея в виду полученные оценки, будем выбирать существенные компоненты в ОМР на основе модуля компонент нормального решения задачи наименьших квадратов.

На первый взгляд делается парадоксальный шаг. Действительно, поиск нормального решения задачи наименьших квадратов превосходит по сложности ОМР алгоритм. Напомним, однако, что нами сделано предположение о тензорной структуре нормального решения. В этом случае, как будет показано далее, существует эффективный алгоритм, который является незначительной модификацией алгоритма ALS.

Остается еще вопрос о быстром поиске больших компонент в полученном приближении к нормальному решению. Для тензорных представлений ответ хорошо известен, а соответствующие алгоритмы описаны, например в [7], [8]. Однако в наших численных экспериментах тензорный ранг  $T$  всегда выбирался равным 1. В этом случае выбор  $C$  максимальных элементов тензора может быть еще упрощен. А именно, пусть после решения задачи наименьших квадратов для тензора  $v_T$  получается представление в виде

$$v_T = u_1 \otimes u_2 \otimes \dots \otimes u_d.$$

Выберем из  $u_1$  не более  $C$  максимальных по модулю элементов и обозначим получившееся множество через  $G_1$ . Затем перемножим попарно все элементы из  $G_1$  с элементами из  $u_2$  и выберем не более  $C$  максимальных по модулю элементов, а полученное множество обозначим через  $G_2$ . Продолжая процедуру, в конце получим множество  $G_d$ , содержащее  $C$  максимальных по модулю элементов тензора  $v_T$ . Сложность такого поиска при простейшей реализации составляет  $\mathcal{O}(C(n_1 + \dots + n_d + d \log C))$ .

Основываясь на сказанном, предлагается следующий порядок вычислений в тензоризованном алгоритме 2.

**Algorithm 2.** Тензоризованный ОМР

**Вход:** тензоризованный линейный оператор  $\mathcal{A}$ , вектор правой части  $b$ , количество компонент разреженного решения  $K$ , ранг тензора для малоранговой аппроксимации  $T$ , количество кандидатов  $C$ .

**Инициализация:**

- $x^0 = 0$ ;
- $r^0 = b - \mathcal{A}x^0 = b$ ;
- множество выбранных столбцов  $S^0 = \emptyset$ ;
- множество столбцов для ортогонализации  $Q^0 = \emptyset$ ;

**for**  $k = 0, \dots, K - 1$  **do**

1. Вычислить решение ранга  $T$  для решения задачи наименьших квадратов

$$v_T = \arg \min_{\substack{u \in \mathbb{R}^{n_1 \times \dots \times n_d} \\ \text{rank } u \leq T}} \|Q^k \mathcal{A}u - Q^k b\|_2.$$

Для простоты мы обозначили  $Q^k$  и множество столбцов, и матрицу ортогонализации к этим столбцам.

2. Найти  $S$  элементов с максимальными по модулю элементами в тензоре  $v_T$ . Обозначим полученное множество  $\hat{S}$ .

3. Построим матрицу  $\hat{A}$ , состоящую из столбцов с номерами из  $\hat{S}$ ;  $i$ -й столбец может быть получен действием оператора  $\mathcal{A}$  на единичный вектор.

4. Применим один шаг метода ОМР к задаче с матрицей  $Q^k \hat{A}$  и правой частью  $Q^k b$ . В результате получим некоторый номер столбца  $j_0$  для матрицы исходного оператора  $\mathcal{A}$ .

5. Добавим столбец  $j_0$  в множество выбранных столбцов:  $S^{k+1} = S^k \cup \{j_0\}$ , а также столбец  $Q^k a_{j_0}$  в множество столбцов для ортогонализации:  $Q^{k+1} = Q^k \cup \{Q^k a_{j_0}\}$ .

**end for**

**return**  $S^K$

Остается вопрос об эффективном поиске решения малого тензорного ранга для ортогонализованной задачи наименьших квадратов

$$v_T = \arg \min_{\substack{u \in \mathbb{R}^{n_1 \times \dots \times n_d} \\ \text{rank } u \leq T}} \|Q^k \mathcal{A}u - Q^k b\|_2.$$

Для этого предлагается использовать простую модификацию алгоритма ALS. Действительно, для решаемой задачи

$$\left\| Q \mathcal{A} \left( \sum_{i=1}^T u_{1i} \otimes u_{2i} \otimes \dots \otimes u_{di} \right) - Qb \right\|_2 \rightarrow \min. \tag{5}$$

выберем некоторое  $i$  от 1 до  $d$  и фиксируем все  $u_{ji}$  при  $j \neq i$ . Решим задачу относительно  $u_{ii}$ . Пусть  $U_i = [u_{i1}^T \ u_{i2}^T \ \dots \ u_{iT}^T]^T$ , перепишем (5) в виде

$$\|QM_i U_i - Qb\|_2 \rightarrow \min,$$

где  $M_i \in \mathbb{R}^{m \times T n_i}$  размерность новой задачи наименьших квадратов. Для получения решения алгоритм ALS многократно итерирует по всем  $i$  от 1 до  $d$ .

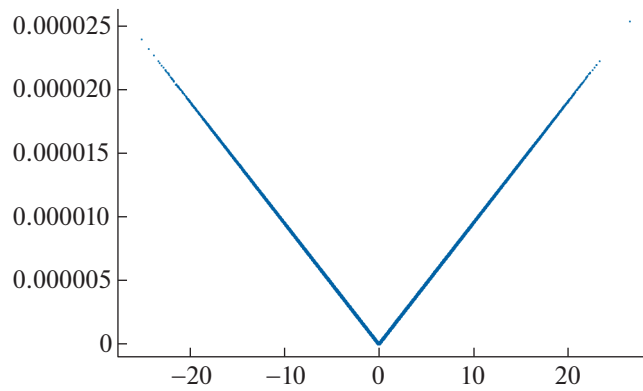
Заметим также, что наличие матрицы ортогонализации  $Q$  не существенно влияет на оценку сложности. Действительно,  $Q$  имеет вид

$$Q = (I - w_1 w_1^T)(I - w_2 w_2^T) \dots (I - w_k w_k^T),$$

где  $k$  соответствует количеству векторов, к которым происходит ортогонализация. Таким образом, умножение на  $Q$  соответствует последовательности из  $k$  одноранговых преобразований и имеет сложность  $\mathcal{O}(k m T n_i)$ , которая сопоставима со сложностью решения задачи наименьших квадратов  $\mathcal{O}(m T^2 n_i^2)$ . Таким образом, решение ортогонализованной задачи оказывается ненамного сложнее классического алгоритма ALS для нахождения малорангового решения.

Пусть алгоритм совершает  $N_{ALS}$  полных итераций (т.е. при которых  $i$  пробегает от 1 до  $d$ ). Тогда полная сложность поиска решения малого ранга равна

$$\mathcal{O} \left( m N_{ALS} \left( k T \sum_{i=1}^d n_i + T^2 \sum_{i=1}^d n_i^2 \right) \right).$$



**Фиг. 1.** График зависимости модуля компонент решения  $|v_j|$  от скалярного произведения  $(a_j, y)$  для случайной матрицы из нормального распределения. Размер матрицы  $20 \times 2^{20}$ .

Из приведенных рассуждений легко видеть, что алгоритм имеет полиномиальную сложность относительно параметров  $m, n_1, \dots, n_d$ . Более того, сложность линейна по  $m$ .

#### 4. ТЕОРЕТИЧЕСКИЙ АНАЛИЗ

В этом разделе мы не ставим целью доказать строгие результаты о сходимости метода тензоризованного ОМР. Мы лишь предполагаем дать более глубокое понимание различных шагов алгоритма.

Тензоризованное решение задачи наименьших квадратов малого тензорного ранга представляет собой довольно сложный для анализа объект. Кроме того, согласно одному из базовых предположений, нормальное решение задачи наименьших квадратов приближается тензором малого ранга. Поэтому вместо анализа решений малого ранга начнем с анализа нормального решения общей задачи наименьших квадратов для случайных матриц. Для достаточно богатого класса случайных матриц, в частности, матриц из нормального распределения, можно показать наличие свойства ограниченной изометрии [9]. Как было показано выше, для матриц, обладающих свойством ограниченной изометрии, нормальное решение задачи наименьших квадратов

$$v = \arg \min_u \|Au - y\|_2$$

является хорошей оценкой на величину

$$c_j = \frac{(a_j, y)}{\|a_j\|_2}.$$

В качестве подтверждения этому был проведен численный эксперимент. Были сгенерированы случайная матрица  $A$  из нормального распределения размера  $m \times 2^m$  при  $m = 20$  и случайный вектор правой части  $y$ . Для каждого  $j = 1, \dots, 2^m$  были вычислены значение  $v_j$  компоненты решения задачи наименьших квадратов и величина  $(a_j, y)$ . На фиг. 1 изображены пары точек  $((a_j, y), |v_j|)$ . Отсюда легко видеть, что компоненты  $|v_j|$  являются хорошими оценками на величины  $|c_j|$  при малых значениях  $\delta$ .

##### 4.1. Влияние ортогонализации

Опишем, что изменится, если вместо задачи наименьших квадратов для  $Ax = y$  решается задача

$$\|QAx - Qy\|_2 \rightarrow \min.$$

Рассуждения этого пункта верны для любых матриц. Пусть имеются матрица  $A \in \mathbb{R}^{m \times n}$  и вектор правой части  $y$ , и на первом шаге метода мы некоторым образом нашли вектор  $a^{(1)}$  и ортогонализуем к нему матрицу. Матрица ортогонализации может быть представлена в виде

$$Q_1 = I - a^{(1)}(a^{(1)})^T / \|a^{(1)}\|_2^2.$$

Таким образом, решается система  $Q_1 A = Q_1 y$ . Пусть далее мы нашли столбец  $a^{(2)}$  и ортогонализуем систему к нему. Поскольку  $a^{(2)}$  выбирался из столбцов матрицы  $Q_1 A$ , векторы  $a^{(1)}$  и  $a^{(2)}$  будут ортогональны. Матрица  $Q_2$  строится аналогичным образом:

$$Q_2 = I - a^{(2)}(a^{(2)})^T / \|a^{(2)}\|_2^2$$

и решается система  $Q_2 Q_1 A = Q_2 Q_1 y$ . Верна следующая

**Теорема 1.** Пусть  $A \in \mathbb{R}^{m \times n}$  и  $y \in \mathbb{R}^m$ . Пусть  $Q_1, Q_2, \dots, Q_k$  – набор матриц ортогонализации, порожденных попарно ортогональными векторами. В этом случае решения задач наименьших квадратов систем

$$Q_n Q_{n-1} \dots Q_1 A = Q_n Q_{n-1} \dots Q_1 y \quad (6)$$

и

$$A = Q_n Q_{n-1} \dots Q_1 y \quad (7)$$

совпадают.

**Доказательство.** Докажем, что псевдообратная матрица к матрице ортогонализации  $Q_i$  совпадает с ней самой, т.е.  $Q_i^\dagger = Q_i$ . Легко проверить, что

$$Q_i^2 = Q_i, \quad Q_i^T = Q_i,$$

откуда следуют все аксиомы псевдообратной матрицы. Далее покажем, что матрицы  $Q_i$  и  $Q_j$  коммутируют. Действительно,

$$\begin{aligned} Q_i Q_j &= \left( I - a^{(i)}(a^{(i)})^T / \|a^{(i)}\|_2^2 \right) \left( I - a^{(j)}(a^{(j)})^T / \|a^{(j)}\|_2^2 \right) = \\ &= I - a^{(i)}(a^{(i)})^T / \|a^{(i)}\|_2^2 - a^{(j)}(a^{(j)})^T / \|a^{(j)}\|_2^2 + \frac{1}{\|a^{(i)}\|_2^2 \|a^{(j)}\|_2^2} a^{(i)}(a^{(i)})^T a^{(j)}(a^{(j)})^T. \end{aligned}$$

Но последнее слагаемое равно 0 в силу того, что  $(a^{(i)})^T a^{(j)} = 0$ , откуда видно, что матрицы  $Q_i$  и  $Q_j$  коммутируют.

Решение системы (6) представляется в виде

$$\begin{aligned} (Q_n Q_{n-1} \dots Q_1 A)^\dagger Q_n Q_{n-1} \dots Q_1 y &= A^\dagger Q_1^\dagger Q_2^\dagger \dots Q_n^\dagger Q_n Q_{n-1} \dots Q_1 y = \\ &= A^\dagger Q_1 Q_2 \dots Q_n Q_n Q_{n-1} \dots Q_1 y = A^\dagger Q_n^2 Q_{n-1}^2 \dots Q_1^2 y = A^\dagger Q_n Q_{n-1} \dots Q_1 y, \end{aligned} \quad (8)$$

что в точности совпадает с решением задачи наименьших квадратов (7). Теорема доказана.

Из этого видно, что решения задач

$$\|QAx - Qy\|_2 \rightarrow \min$$

и

$$\|Ax - Qy\|_2 \rightarrow \min$$

совпадают, т.е. достаточно ортогонализировать только правую часть системы. Отсюда следует, что рассуждения о ранжировании по величине компонент нормального решения верны на любом шаге метода, поскольку опираются на свойство ограниченной изометрии матрицы.

Заметим, что это утверждение верно лишь для нормального решения задачи наименьших квадратов и может оказаться неверным для решения задачи наименьших квадратов малого ранга.

Проведенный анализ показывает, что выбор кандидатов на основе модулей компонент решения задачи наименьших квадратов не ограничивает классического алгоритма ОМР в том смысле,

что для широкого класса случайных матриц алгоритм тензоризованного ОМР строит множество столбцов, которое не может сильно отличаться от ранжирования по величинам

$$c_j = \frac{(a_j, r)}{\|a_j\|_2}.$$

Однако на практике мы не можем найти точное решение задачи наименьших квадратов для больших систем, поэтому алгоритм тензоризованного ОМР выбирает  $C$  кандидатов на основе компонент малорангового решения. С учетом предположения, что точное решение задачи наименьших квадратов может быть хорошо приближено тензором малого ранга, это дает возможность надеяться, что в множество из  $C$  кандидатов попадет столбец, который бы выбрал алгоритм ОМР.

4.2. О падении невязки на различных шагах метода

Изучим вопрос о скорости падения невязки для матрицы  $A$ , имеющей вид

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & -1 & -1 \\ 1 & 1 & 1 & 1 & \dots & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & \dots & -1 & -1 \\ 1 & 1 & -1 & -1 & \dots & -1 & -1 \\ 1 & -1 & 1 & -1 & \dots & 1 & -1 \end{bmatrix}. \tag{9}$$

Эта матрица содержит всевозможные столбцы из  $\pm 1$ , взятые в лексикографическом порядке. Легко показать, что эта матрица обладает свойством ограниченной изометрии с  $\delta = 0$ . Пусть правая часть задана в виде суммы  $k$  столбцов матрицы  $A$

$$b = \sum_{j=1}^k \alpha_j a_{i_j} + \varepsilon.$$

Пусть произошла ортогонализация к вектору  $p$ , где  $\|p\|_2 = 1$ . Разложим столбцы матрицы  $A$ , входящие в правую часть, на компоненты, коллинеарные  $p$  и ортогональные ему:

$$a_{i_j} = a_{i_j}^{\parallel} + a_{i_j}^{\perp},$$

откуда  $a_{i_j}^{\parallel} = \beta_j p$ . Тогда

$$b = \sum_{j=1}^k \alpha_j a_{i_j}^{\parallel} + \varepsilon^{\parallel} + \sum_{j=1}^k \alpha_j a_{i_j}^{\perp} + \varepsilon^{\perp}.$$

Тогда после ортогонализации

$$b_1 = (I - pp^T)b = \sum_{j=1}^k \alpha_j a_{i_j}^{\parallel} + \varepsilon^{\parallel} + \sum_{j=1}^k \alpha_j a_{i_j}^{\perp} + \varepsilon^{\perp} - \left( \sum_{j=1}^k \alpha_j (p, a_{i_j}^{\parallel}) + (p, \varepsilon^{\parallel}) \right) p.$$

Откуда, используя  $a_{i_j}^{\parallel} = \beta_j p$  и  $\varepsilon^{\parallel} = \gamma p$ , получаем

$$b_1 = \sum_{j=1}^k \alpha_j a_{i_j}^{\perp} + \varepsilon^{\perp}.$$

Добавим и вычтем коллинеарную составляющую

$$b_1 = \sum_{j=1}^k \alpha_j a_{i_j}^{\perp} + \varepsilon^{\perp} + \sum_{j=1}^k \alpha_j a_{i_j}^{\parallel} + \varepsilon^{\parallel} - \sum_{j=1}^k \alpha_j \beta_j p - \gamma p = \sum_{j=1}^k \alpha_j a_{i_j} + \tau p + \varepsilon.$$

Заметим, что  $p$  с точностью до нормировки является одним из столбцов матрицы  $A$ . Таким образом, после шага метода в правой части становится на 1 столбец больше.



Оценим норму вектора правой части  $b_1$  после ортогонализации. По теореме Пифагора имеем

$$\|b\|_2^2 = \|b_1\|_2^2 + \frac{(p, b)^2}{\|p\|_2^2},$$

откуда

$$\frac{\|b_1\|_2^2}{\|b\|_2^2} = 1 - \frac{(p, b)^2}{\|p\|_2^2 \|b\|_2^2},$$

что равносильно

$$\frac{\|b_1\|_2}{\|b\|_2} = |\sin \angle(p, b)|.$$

Оценим для начала худший случай. Выберем нормированный вектор  $b$  так, что  $\max_p (p, b)^2$  минимально. Поскольку  $p$  состоит из  $\pm 1$ , можно, не ограничивая общности, считать, что все компоненты вектора  $b$  неотрицательны и вектор  $p$  состоит из всех 1. Тогда нужно найти вектор  $b$  такой, что  $\|b\|_2 = 1$ , все элементы неотрицательны и их сумма минимальна. Очевидно, это вектор  $b = e_1$ , имеющий единицу в 1 позиции и нули в остальных.

Тогда в худшем случае имеем

$$\frac{\|b_1\|_2^2}{\|b\|_2^2} = 1 - \frac{1}{m}.$$

Заметим, что этот случай вовсе не является нереалистичным:

$$b = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Однако в среднем все не так плохо. Были проведены следующие численные эксперименты. Для правой части выбиралось  $k$  случайных столбцов матрицы и  $k$  коэффициентов из нормального распределения. Находилась падение квадрата нормы правой части и результат усреднялся по  $10^6$  экспериментов. На фиг. 2 приведен график зависимости  $1 - \|b_1\|_2^2 / \|b\|_2^2$  от количества столбцов  $k$ . Легко видеть, что среднее падение всегда больше 0.5 и очень слабо зависит от числа строк матрицы  $m$ .

Оценим величину

$$\max_p \frac{(p, b)^2}{\|p\|_2^2 \|b\|_2^2}. \quad (10)$$

Во-первых, ясно, что  $\|p\|_2^2 = m$ . Пусть

$$b = \sum_{j=1}^k \alpha_j a_{ij}.$$

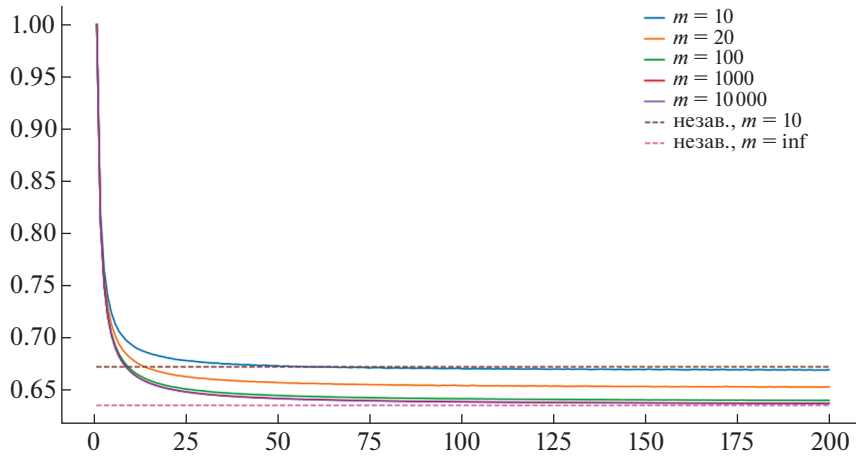
Тогда все компоненты имеют вид

$$b_i = \pm \alpha_1 \pm \alpha_2 \pm \dots \pm \alpha_k,$$

где  $\alpha_j \sim \mathcal{N}(0, 1)$ , и знаки  $+$  и  $-$  выбираются с равной вероятностью. Легко видеть, что  $\pm \alpha_j \sim \mathcal{N}(0, 1)$ , следовательно,

$$b_i \sim \mathcal{N}(0, k).$$

Однако компоненты вектора  $b$  не являются независимыми.



Фиг. 2. График зависимости квадрата падения нормы невязки от количества компонент в правой части.

Далее, из (10) легко посчитать максимум

$$\max_p \frac{(p, b)^2}{\|p\|^2 \|b\|^2} = \frac{\|b\|_1^2}{m \|b\|_2^2}.$$

Нетрудно понять, что, чем больше слагаемых в  $b$  (т.е. чем больше  $k$ ), тем более независимы между собой будут компоненты  $b_i$ , и наоборот – чем меньше  $k$ , тем более зависимы компоненты. Так, если  $k = 1$ , то будет всего один вариант  $|\alpha_1|$ , если  $k = 2$ , то 2 варианта:  $|\alpha_1 + \alpha_2|$  и  $|\alpha_1 - \alpha_2|$ . В

общем случае  $2^{k-1}$  вариантов. Величина  $\frac{\|b\|_1^2}{\|b\|_2^2}$  будет максимальна, если все компоненты равны, т.е.

наиболее зависимы. Из этого наблюдения и из фиг. 2 можно видеть, что с ростом  $k$  (по оси  $x$ ) величина (10) убывает. Отсюда вытекает гипотеза, что оценкой снизу на (10) будет ситуация, когда все компоненты независимы. Доказательством этой гипотезы мы не обладаем, однако, численные эксперименты хорошо согласуются с ней. В случае если все компоненты  $b_i$  независимые и одинаково распределенные с  $b_i \sim \mathcal{N}(0, 1)$  (в случае  $b_i \sim \mathcal{N}(0, k)$  можно отнормировать числитель и знаменатель), имеем

$$\mathbb{E} \left( \frac{\|b\|_1^2}{\|b\|_2^2} \right) = 1 + \frac{2}{\pi} (m - 1).$$

Для сравнения эти оценки изображены на фиг. 2 штриховыми линиями. Эту оценку можно доказать аналитически.

Действительно,

$$\|b\|_1^2 = (|b_1| + |b_2| + \dots + |b_m|)^2 = \sum_{j=1}^m b_j^2 + \sum_{i,j=1, i \neq j}^k |b_i b_j|.$$

Тогда

$$\frac{\|b\|_1^2}{\|b\|_2^2} = 1 + \sum_{i,j=1, i \neq j}^k \frac{|b_i b_j|}{b_1^2 + \dots + b_m^2},$$

откуда имеем, что

$$\mathbb{E} \left( \frac{\|b\|_1^2}{\|b\|_2^2} \right) = 1 + m(m - 1) \mathbb{E} \left( \frac{|b_1 b_2|}{b_1^2 + \dots + b_m^2} \right).$$

Последнее матожидание равно

$$\begin{aligned} \mathbb{E}\left(\frac{|b_1 b_2|}{b_1^2 + \dots + b_m^2}\right) &= \frac{1}{(2\pi)^{m/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{|x_1 x_2|}{x_1^2 + \dots + x_m^2} \exp\left(-\frac{x_1^2 + \dots + x_m^2}{2}\right) dx_1 \dots dx_m = \\ &= \frac{2^{m/2}}{\pi^{m/2}} \int_0^{\infty} \dots \int_0^{\infty} \frac{x_1 x_2}{x_1^2 + \dots + x_m^2} \exp\left(-\frac{x_1^2 + \dots + x_m^2}{2}\right) dx_1 \dots dx_m. \end{aligned}$$

Вычислим при  $\varepsilon > 0$  интеграл

$$\int_0^{\infty} \dots \int_0^{\infty} \int_{\varepsilon}^{\infty} \int_{\varepsilon}^{\infty} \frac{x_1 x_2}{x_1^2 + \dots + x_m^2} \exp\left(-\frac{x_1^2 + \dots + x_m^2}{2}\right) dx_1 \dots dx_m.$$

Сделаем замену переменных

$$\begin{aligned} y_1 &= x_1^2, \\ y_2 &= x_2^2, \\ y_3 &= x_3^2, \\ &\vdots \\ y_m &= x_m^2. \end{aligned}$$

Легко видеть, что якобиан такого преобразования равен  $\frac{1}{4x_1 x_2}$ , откуда последний интеграл равен

$$\frac{1}{4} \int_0^{\infty} \dots \int_0^{\infty} \int_{\varepsilon^2}^{\infty} \int_{\varepsilon^2}^{\infty} \frac{\exp\left(-\frac{y_1 + y_2 + y_3 + \dots + y_m}{2}\right)}{y_1 + y_2 + y_3 + \dots + y_m} dy_1 \dots dy_m.$$

Сделаем еще одну замену

$$\begin{aligned} t_1 &= \frac{y_1 + y_2 + y_3 + \dots + y_m}{2}, \\ t_2 &= y_2, \\ t_3 &= y_3, \\ &\vdots \\ t_m &= y_m. \end{aligned}$$

В этом случае якобиан равен 2 и интеграл принимает вид

$$\frac{1}{4} \int_0^{\infty} \dots \int_0^{\infty} \int_{\frac{\varepsilon^2 \varepsilon^2 + t_2 + t_3 + \dots + t_m}{2}}^{\infty} \frac{e^{-t_1}}{t_1} dt_1 \dots dt_m. \quad (11)$$

Введем обозначение для экспоненциального интеграла

$$E_n(x) = \int_1^{\infty} \frac{e^{-xt}}{t^n} dt.$$

Заметим, что

$$E_1(x) = \int_1^{\infty} \frac{e^{-xt}}{t} dt$$

и сделаем замену  $y = xt$  при  $x > 0$ . Тогда получим

$$E_1(x) = \int_x^{\infty} \frac{e^{-y}}{y} dy.$$

В этих обозначениях интеграл (11) равен

$$\frac{1}{4} \int_0^\infty \dots \int_0^\infty E_1 \left( \frac{\varepsilon^2 + t_2 + t_3^2 + \dots + t_m^2}{2} \right) dt_2 \dots dt_m.$$

Вычислим производную

$$\begin{aligned} \left( E_n \left( \frac{x}{2} + c \right) \right)'_x &= \int_1^\infty \frac{e^{-\left(\frac{x}{2}+c\right)t} \left( -\frac{t}{2} \right)}{t^n} dt = -\frac{1}{2} E_{n-1} \left( \frac{x}{2} + c \right). \\ E_n \left( \frac{x}{2} + c \right) &= -2 \left( E_{n+1} \left( \frac{x}{2} + c \right) \right)'_x. \end{aligned}$$

Посчитаем неопределенный интеграл

$$\begin{aligned} \int E_1 \left( \frac{\varepsilon^2 + t_2 + t_3^2 + \dots + t_m^2}{2} \right) dt_2 &= -2 \int \left( E_2 \left( \frac{\varepsilon^2 + t_2 + t_3^2 + \dots + t_m^2}{2} \right) \right)'_{t_2} dt_2 = \\ &= -2 \int_1^\infty \frac{e^{-\frac{\varepsilon^2 + t_2 + t_3^2 + \dots + t_m^2}{2}}}{t_1^2} dt_1 + C. \end{aligned} \tag{12}$$

Откуда имеем, что

$$\int_{\varepsilon^2}^\infty E_1 \left( \frac{\varepsilon^2 + t_2 + t_3^2 + \dots + t_m^2}{2} \right) dt_2 = 2 \int_1^\infty \frac{e^{-\left(\frac{\varepsilon^2 + t_3^2 + \dots + t_m^2}{2}\right)t_1}}{t_1^2} dt_1.$$

Переходя к пределу при  $\varepsilon \rightarrow 0$ , получаем, что

$$\begin{aligned} \int_0^\infty \dots \int_0^\infty \frac{x_1 x_2}{x_1^2 + \dots + x_m^2} \exp \left( -\frac{x_1^2 + \dots + x_m^2}{2} \right) dx_1 \dots dx_m &= \\ = \frac{1}{2} \int_0^\infty \dots \int_0^\infty \frac{e^{-\left(\frac{t_3^2 + \dots + t_m^2}{2}\right)t_1}}{t_1^2} dt_1 dt_3 \dots dt_m &= \frac{1}{2} \int_1^\infty \frac{1}{t_1^2} \left( \int_0^\infty e^{-\frac{t_3^2}{2} dt_3} \right) \dots \left( \int_0^\infty e^{-\frac{t_m^2}{2} dt_m} \right) dt_1. \end{aligned}$$

Простой заменой получаем, что

$$\int_0^\infty e^{-\frac{t_3^2}{2}} dt_3 = \sqrt{2t_1}.$$

Тогда имеем

$$\int_0^\infty \dots \int_0^\infty \frac{x_1 x_2}{x_1^2 + \dots + x_m^2} \exp \left( -\frac{x_1^2 + \dots + x_m^2}{2} \right) dx_1 \dots dx_m = \frac{\pi^{\frac{m-2}{2}}}{2^{\frac{m}{2}}} \int_1^\infty \frac{dt_1}{t^{\frac{m+2}{2}}} = \frac{\pi^{\frac{m-2}{2}}}{2^{\frac{m-2}{2}} m}.$$

И в итоге получаем

$$\begin{aligned} \mathbb{E} \frac{|b_1 b_2|}{b_1^2 + \dots + b_m^2} &= \frac{2^{\frac{m}{2}} \pi^{\frac{m-2}{2}}}{\pi^2 2^{\frac{m-2}{2}} m} = \frac{2}{\pi m}, \\ \mathbb{E} \left( \frac{\|b\|_1^2}{\|b\|_2^2} \right) &= 1 + m(m-1) \frac{2}{\pi m} = 1 + \frac{2}{\pi} (m-1), \end{aligned}$$

$$\mathbb{E} \left( \max_p \frac{(p, b)^2}{\|b\|^2 \|b\|^2} \right) = \frac{2}{\pi} + \left(1 - \frac{2}{\pi}\right) \frac{1}{m}.$$

Тогда матожидание падения нормы правой части оказывается равно

$$\mathbb{E} \frac{\|b_1\|_2^2}{\|b\|_2^2} = \left(1 - \frac{2}{\pi}\right) \left(1 - \frac{1}{m}\right),$$

где  $b_1$  – правая часть после первого шага метода.

Таким образом, для матриц вида (9) на первом шаге невязка будет падать в среднем более чем в 2 раза. Это позволяет надеяться на высокую скорость сходимости метода. Однако такое быстрое падение невязки является особенностью того, что столбцы матрицы (9) очень плотно покрывают все пространство. Для случайных матриц, не обладающих такой особенностью, падение невязки ожидается более медленным.

## 5. ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

В этом разделе приведены результаты численных экспериментов для предложенного метода.

### 5.1. Описание модели

Модель, рассматриваемая в этом разделе, носит название модель ЕМР-Вольтерра. Введем обозначения. Пусть  $X, Y \in \mathbb{C}^N$  и обозначают входы и выходы модели соответственно. Здесь  $N$  обозначает длину сигнала,  $P = \left\lceil \frac{\max X_i}{2^K} \right\rceil$ , где  $K$  является нормировочным параметром модели.

Матрица  $F \in \mathbb{C}^{N \times (P+1)}$  является некоторой функцией от входа  $X$  и считается известной,  $e$  обозначает вектор из единиц. Параметрами модели, подлежащими определению, являются  $u_{rd} \in \mathbb{C}^{S_x}$ ,  $v_{rd} \in \mathbb{C}^{S_x}$ ,  $w_{rd} \in \mathbb{C}^{S_L}$ ,  $a_{rd} \in \mathbb{C}^{P+1}$ ,  $\alpha_{rd}, \beta_{rd} \in \mathbb{C}$ . Тогда модель ЕМР-Вольтерра может быть записана следующим образом:

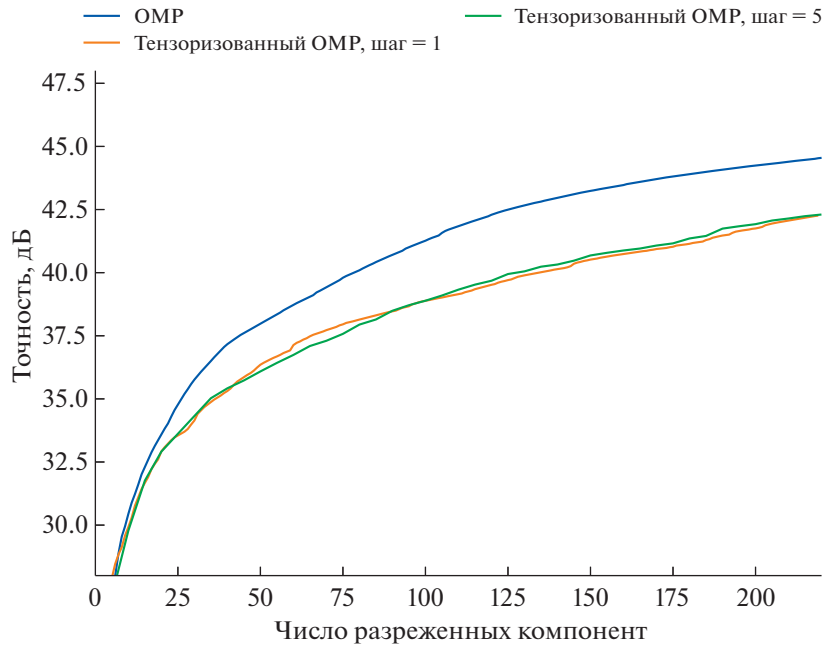
$$Y \approx \sum_{r=1}^R \left( \bigodot_{d=1}^{\left\lceil \frac{D_X}{2} \right\rceil} (X \otimes u_{rd} + \alpha_{rd} e) \right) \odot \left( \bigodot_{d=1}^{\left\lceil \frac{D_X}{2} \right\rceil} (\bar{X} \otimes v_{rd} + \beta_{rd} e) \right) \odot \left( \bigodot_{d=1}^{D_L} ((F \alpha_{rd}) \otimes w_{rd}) \right).$$

Здесь символом  $\odot$  обозначено адамарово произведение, а символом  $\otimes$  свертка между двумя векторами.

Несложно видеть, что эту модель можно записать как тензоризованный линейный оператор относительно  $[u_{rd}, \alpha_{rd}]$ ,  $[v_{rd}, \beta_{rd}]$ ,  $w_{rd}$  и  $a_{rd}$ . Число столбцов матрицы такого оператора может быть вычислено по формуле

$$N_{COL} = (S_X + 1)^{D_X} \times S_L^{D_L} \times (P + 1)^{D_L}.$$

В качестве входного сигнала  $X$  и выходного сигнала  $Y$  были использованы реальные данные для решения задачи цифрового предискажения (Digital Predistortion, DPD). Длина сигнала  $N \sim 5 \times 10^4$ . Наименьшие размеры модели, которая обеспечивает достаточное для практического применения качество цифрового предискажения для данного сигнала, равны  $S_X = 7$ ,  $S_L = 7$ ,  $D_X = 3$ ,  $D_L = 1$ . В этом случае число столбцов получающейся матрицы равно 39,424. Чуть более сложная модель получается выбором  $S_X = 7$ ,  $S_L = 7$ ,  $D_X = 3$ ,  $D_L = 3$ , что соответствует 233,744,896 столбцам. Максимальный размер задачи, для которой проводились эксперименты, соответствует параметрам  $S_X = 13$ ,  $S_L = 13$ ,  $D_X = 3$ ,  $D_L = 3$ , где матрица содержит приблизительно  $8 \times 10^9$  столбцов и  $5 \times 10^4$  строчек.



**Фиг. 3.** График сравнения методов ОМР и тензоризованного ОМР с различным числом шагов в пункте 4 Алгоритма 2. Модель  $S_X = 7, S_L = 7, D_X = 3, D_L = 1$ . Для метода тензоризованного ОМР отбиралось  $C = 400$  столбцов.

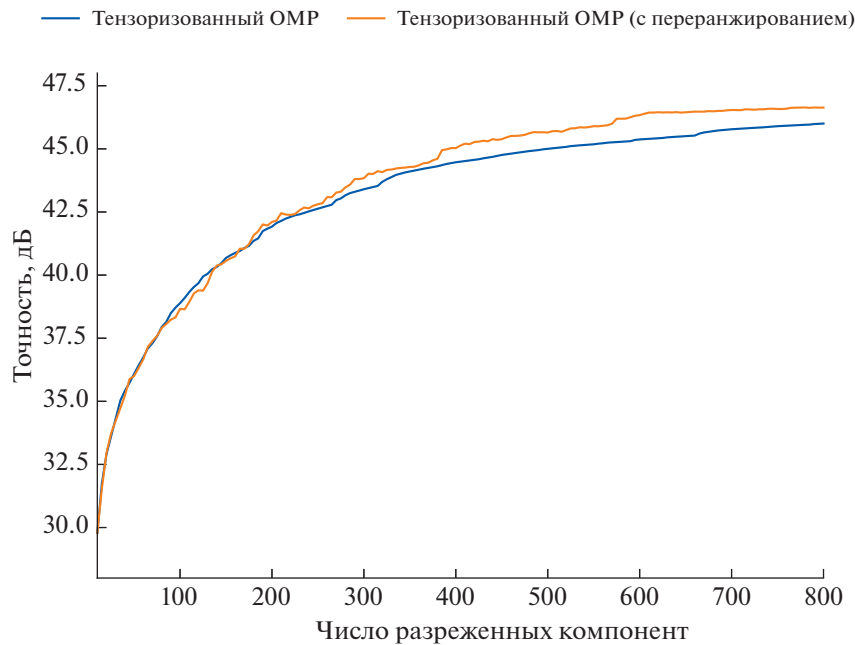
В качестве метрики качества работы была взята классическая для теории обработки сигналов единица измерения – децибелы. Таким образом, точность вычисляется по формуле

$$-20 \lg_{10} \frac{\|Su - Y\|_2}{\|Y\|_2},$$

где  $u$  – найденное разреженное решение.

### 5.2. Результаты

Ниже приведены результаты применения метода тензоризованного ОМР к описанной выше модели. На фиг. 3 приведен график сравнения алгоритма ОМР и метода тензоризованного ОМР для модели с параметрами  $S_X = 7, S_L = 7, D_X = 3, D_L = 1$ . Здесь и далее алгоритм тензоризованного ОМР отбирает  $C = 400$  столбцов-кандидатов для множества  $\hat{S}$ . В качестве метода поиска малорангового решения вместо алгоритма ALS использовался метод Левенберга-Марквардта [10], [11]. Как отмечалось выше, на каждой итерации метода тензоризованного ОМР в пункте 4 Алгоритма 2 можно выбирать не 1 столбец, а несколько. Поскольку нахождение малорангового решения занимает большую часть времени работы, такой прием позволяет существенно ускорить время вычислений. Для сравнения на фиг. 3 приведен также график для метода тензоризованного ОМР с выбором 5 столбцов на каждой итерации. Из графиков, во-первых, можно видеть, что выбор 5 столбцов вместо одного практически не ухудшает точности построенного решения, но позволяет почти в 5 раз ускорить работу программы. Поэтому во всех дальнейших экспериментах в методе тензоризованного ОМР выбирает 5 столбцов за итерацию. Во-вторых, из графиков видно, что метод тензоризованного ОМР работает ненамного хуже классического ОМР, хотя имеет полиномиальную сложность, в отличие от ОМР, имеющего экспоненциальную сложность. На данном этапе качество и не могло получиться лучше, чем у ОМР, поскольку метод тензоризованного ОМР в некотором смысле является его аппроксимацией, предназначенной для работы со сверхбольшими тензоризованными матрицами.

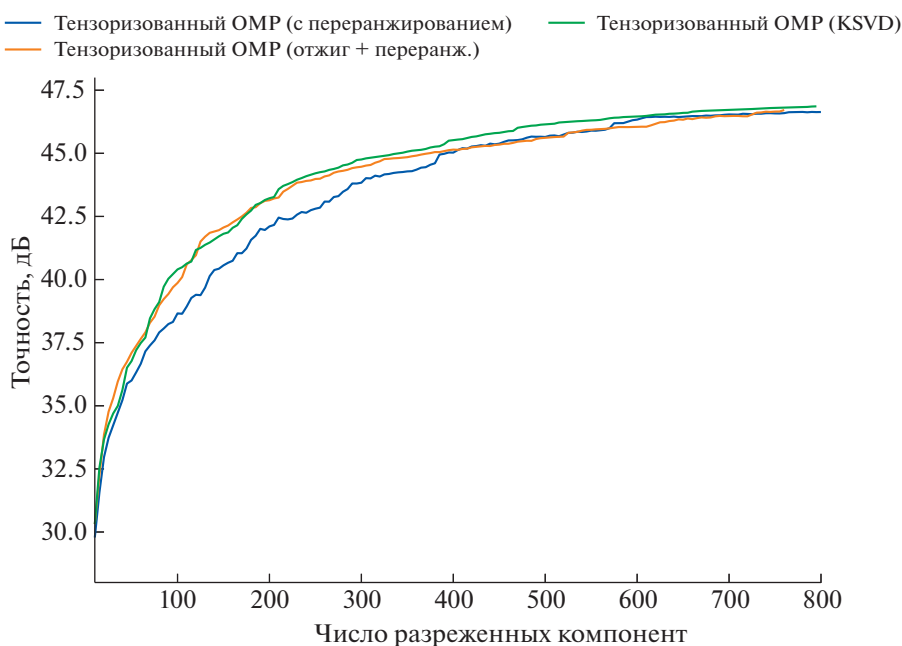


**Фиг. 4.** График сравнения методов простого тензоризованного ОМР и тензоризованного ОМР с переранжированием на каждом шаге. Модель  $S_X = 7$ ,  $S_L = 7$ ,  $D_X = 3$ ,  $D_L = 1$ . Для метода тензоризованного ОМР отбиралось  $C = 400$  столбцов.

Метод тензоризованного ОМР может быть дополнительно улучшен с помощью следующего подхода. На шагах 3–5 Алгоритма 2 будем строить матрицу  $\hat{A}$  содержащей не только столбцы из  $\hat{S}$ , но также и все столбцы из множества уже отобранных ранее столбцов  $S$ . Пусть на некоторой итерации Алгоритма 2 уже выбрано множество  $S$ . Получим множество новых столбцов-кандидатов  $\hat{S}$ . Построим матрицу  $\hat{A}$ , содержащую столбцы и  $S \cup \hat{S}$ , и отберем из нее (после ортогонализации)  $|S| + 5$  столбцов с помощью классического алгоритма ОМР. Такой подход гарантированно не ухудшит качества работы, но может помочь найти хорошее решение с меньшим числом разреженных компонент, что можно видеть из фиг. 4.

Заметим, что матрица  $Q^k \hat{A}$ , для которой применяется ОМР в пункте 4 Алгоритма 2 имеет всего  $C = 400$  столбцов, что позволяет применять здесь не только жадный алгоритм ОМР, но и более продвинутые стратегии оптимизации. В качестве примера были использованы метод имитации отжига и алгоритм K-SVD [12], [13]. Метод имитации отжига представляет собой метод глобальной оптимизации, который на каждом шаге переходит либо в состояние, уменьшающее значение функционала, либо, с некоторой вероятностью, в состояние, ухудшающее значение функционала, причем вероятность последнего зависит от температурного коэффициента, значение которого убывает в процессе работы алгоритма [14]. Алгоритм K-SVD отличается от ОМР своей возможностью не только жадно добавлять столбцы, но и выбрасывать ненужные. На фиг. 5 приведены результаты сравнения методов. Алгоритм K-SVD по построению всегда пересматривает множество ранее выбранных столбцов, поэтому для ОМР и метода имитации отжига было использовано переранжирование всех столбцов на каждой итерации. Из графиков можно видеть, что на первых итерациях метод имитации отжига выбирает столбцы лучше, поскольку является менее жадным методом, однако затем лидерство переходит к K-SVD, поскольку задача оптимизации становится слишком сложной для метода имитации отжига.

Одной из проблем метода тензоризованного ОМР является борьба с влиянием уже выбранных столбцов. Если столбец  $j$  уже был выбран ранее, то после ортогонализации он будет нулевым, поэтому коэффициент на  $j$ -й позиции может быть выбран любым, и это не скажется на невязке. В случае тензоризованного ОМР ситуация несколько менее критична, поскольку малоранговость решения представляет собой своего рода регуляризацию. Однако возможность ставить произвольные значения в некоторые позиции может позволить находить лучшее с точки



**Фиг. 5.** График сравнения различных методов отбора столбцов в пункте 4 Алгоритма 2. Использованные методы: OMP (с переранжированием после каждого шага), метод имитации отжига (с переранжированием после каждого шага) и K-SVD [12], [13]. Модель  $S_X = 7, S_L = 7, D_X = 3, D_L = 1$ . Для метода тензоризованного OMP отбиралось  $C = 400$  столбцов.

зрения невязки решение, что в корне противоречит тому, как дальше будут использованы компоненты решения. Большая по модулю компонента должна соответствовать более полезному для решения столбцу. Для того чтобы добиться этого, можно добавить простейший  $l_2$ -регуляризатор с коэффициентом  $\lambda$ . Минимизируемая функция в этом случае записывается как

$$\left\| Q \mathcal{A} \left( \sum_{t=1}^T u_{1t} \otimes u_{2t} \otimes \dots \otimes u_{dt} \right) - Qb \right\|_2^2 + \lambda \sum_{k,t} \|u_{kt}\|_2^2 \rightarrow \min. \tag{13}$$

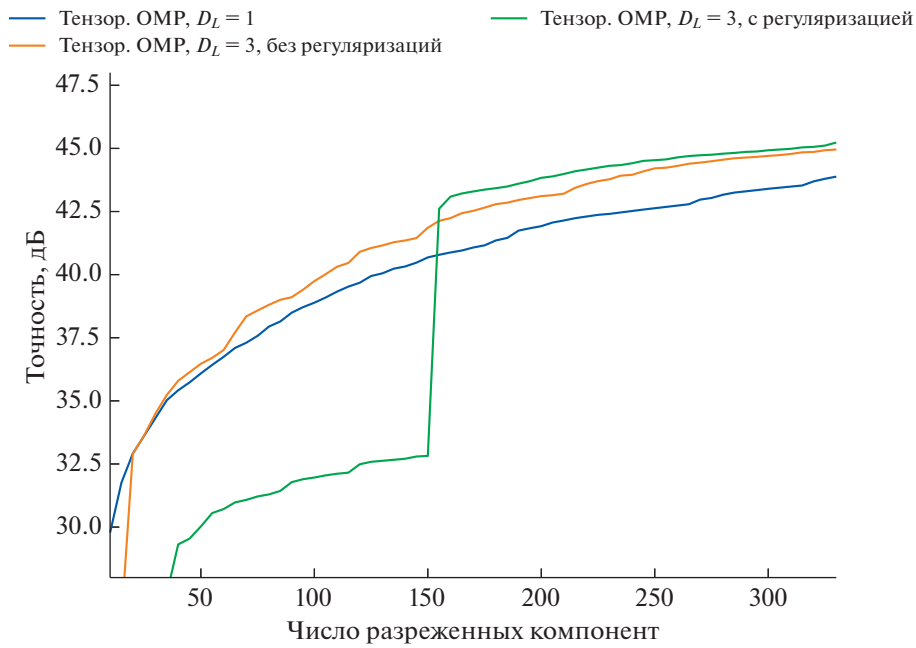
На фиг. 6 приведен график сравнения метода с нерегуляризованной моделью при  $D_L = 1$  и  $D_L = 3$  и регуляризованной модели с  $D_L = 3$ . Из графика можно видеть, что увеличение размера модели закономерно приводит к улучшению качества. Из уравнения (13) несложно заметить, что коэффициент регуляризации  $\lambda$  необходимо выбирать адаптивно. Действительно, на первых итерациях метода невязка будет большой и необходима большая  $\lambda$  для получения эффекта от регуляризации. В то же время слишком большая  $\lambda$  будет приводить к нахождению нулевого решения в качестве оптимального. Поэтому  $\lambda$  выбирается адаптивно: изначально выбирается достаточно большое значение коэффициента регуляризации, но когда поиск малорангового решения начинает давать большую невязку, коэффициент  $\lambda$  уменьшается в несколько раз. Результат такого уменьшения можно видеть на графике в виде скачка.

Классический метод OMP может быть применен к модели с  $D_L = 1$ , но его использование для модели с  $D_L = 3$  уже затруднительно. На фиг. 7 приведен график сравнения OMP для модели с  $D_L = 1$  и тензоризованного OMP с  $D_L = 3$ . Можно видеть, что при большом числе компонент тензоризованный OMP начинает обходить классический OMP, что доказывает его эффективность для поиска разреженных решений сверхбольших систем линейных уравнений.

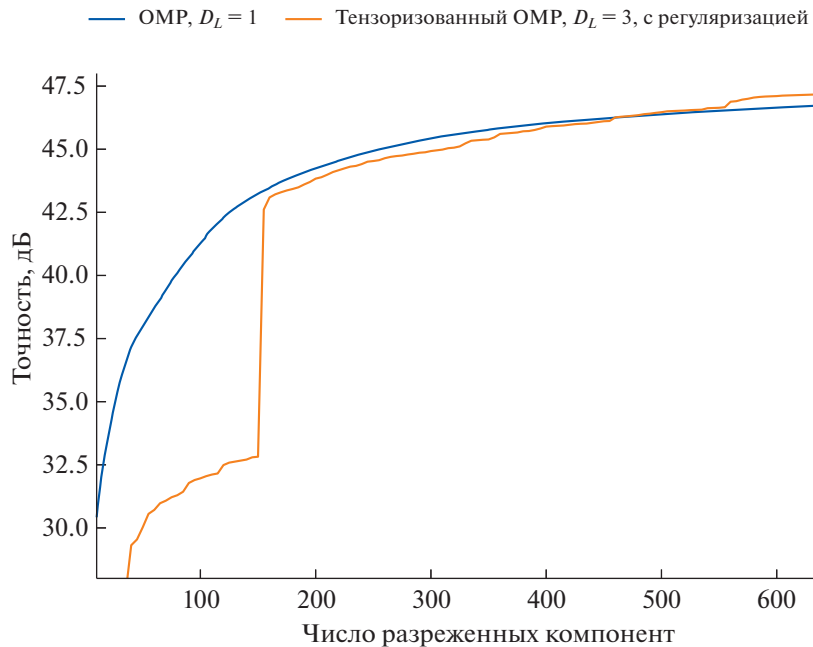
### 6. ЗАКЛЮЧЕНИЕ

В работе предложен метод нахождения разреженного решения для сверхбольших систем линейных уравнений, обладающих тензорной структурой. Проведенный теоретический анализ позволяет понять, почему метод тензоризованного OMP способен успешно находить решения





**Фиг. 6.** График сравнения качества работы на различных моделях, а также эффект от применения регуляризации. Во всех моделях  $S_X = 7$ ,  $S_L = 7$ ,  $D_X = 3$ . Для метода тензоризованного OMP отбиралось  $C = 400$  столбцов.



**Фиг. 7.** График сравнения методов OMP и тензоризованного OMP. Во всех моделях  $S_X = 7$ ,  $S_L = 7$ ,  $D_X = 3$ . Для метода тензоризованного OMP отбиралось  $C = 400$  столбцов.

систем линейных уравнений, а также проследить отдельные этапы работы алгоритма. Представленные экспериментальные результаты демонстрируют хорошую масштабируемость метода тензоризованного OMP, и его способность эффективно находить разреженные решения сверх-больших систем линейных уравнений.

## СПИСОК ЛИТЕРАТУРЫ

1. *Mallat S.G., Zhang Z.* Matching pursuits with time-frequency dictionaries // IEEE Transactions on Signal Proc. 1993. V. 41. № 12. P. 3397–3415.
2. *Elad M.* Sparse and redundant representations: from theory to applications in signal and image processing. Springer Science & Business Media. 2010.
3. *Cai T.T., Wang L.* Orthogonal matching pursuit for sparse signal recovery with noise // IEEE Transactions on Information Theory. 2011. V. 57. № 7. P. 4680–4688.
4. *Candes E.J., Tao T.* Decoding by linear programming // IEEE Transactions on Information Theory. 2005. V. 51. № 12. P. 4203–4215.
5. *Davenport M.A., Wakin M.B.* Analysis of orthogonal matching pursuit using the restricted isometry property // IEEE Transactions on Information Theory. 2010. V. 56. № 9. P. 4395–4401.
6. *Tropp J.A.* Greed is good: Algorithmic results for sparse approximation // IEEE Transactions on Information theory. 2004. V. 50. № 10. P. 2231–2242.
7. *Lebedeva O.* Tensor conjugate-gradient-type method for rayleigh quotient minimization in block qtt-format // Russian Journal of Numerical Analysis and Math. Modelling. 2011. V. 26. № 5. P. 465–489.
8. *Zheltkov D.A., Osinsky A.* Global optimization algorithms using tensor trains. Internat. Conference on Large-Scale Scientific Computing. Springer. 2019. P. 197–202.
9. *Vershynin R.* Introduction to the non-asymptotic analysis of random matrices // arXiv preprint arXiv:1011.3027. 2010.
10. *Levenberg K.* A method for the solution of certain non-linear problems in least squares // Quarterly of Applied Math. 1944. V. 2. № 2. P. 164–168.
11. *Marquardt D.W.* An algorithm for least-squares estimation of nonlinear parameters // J. of the Society for Industrial and Applied Math. 1963. V. 11. № 2. P. 431–441.
12. *Aharon M., Elad M., Bruckstein A.* K-svd: An algorithm for designing overcomplete dictionaries for sparse representation // IEEE Transactions on Signal Proc. 2006. V. 54. № 11. P. 4311–4322.
13. *Rubinstein R., Bruckstein A.M., Elad M.* Dictionaries for sparse representation modeling // Proc. of the IEEE. 2010. V. 98. № 6. P. 1045–1057.
14. *Kirkpatrick S., Gelatt Jr, C., Vecchi M.* Optimization by Simulated Annealing // Science. 1983. V. 220. № 4598. P. 671–680.