

Московский государственный университет имени М.В. Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра вычислительных технологий и моделирования

Султонов Азамат Орзикулович

Малоранговая неотрицательная аппроксимация матриц и тензоров с использованием скетчинга

КУРСОВАЯ РАБОТА

Научные руководители:

к.ф.-м.н., доцент

С. А. Матвеев

к.ф.-м.н.

С. С. Будзинский

Москва, 2022

Содержание

1	Введение	2
2	Методы попеременных проекций для матриц	2
3	Скетчинг	3
4	Методы попеременных проекций для тензоров	3
4.1	Базовые понятия	3
4.2	Неотрицательное малоранговое разложение Таккера	4
4.3	Неотрицательное малоранговое ТТ-разложение	5
4.4	Оценки сложностей алгоритмов	6
5	Численные эксперименты	7
5.1	Случайная матрица из равномерного распределения	7
5.2	Чёрно-белое изображение	9
5.3	Случайный трёхмерный тензор из равномерного распределения	12
5.4	Цветное изображение	14
6	Заключение	19
	Список литературы	20

1 Введение

Малоранговые аппроксимации матриц играют важную роль во многих приложениях. Например, они снижают объёмы памяти, необходимые для хранения данных, а также позволяют быстро выполнять последующую обработку сжатых данных.

Во многих областях возникают неотрицательные функции и наборы данных, которые имеют самые различные формы: видео, изображения, плотности распределения, рейтинги. И обычно важно, чтобы аппроксимации таких данных тоже были неотрицательны.

Известно, что наилучшее малоранговое приближение матрицы в любой унитарно-инвариантной норме даёт усеченное сингулярное разложение. Однако нет никакой гарантии, что полученная матрица низкого ранга сохранит неотрицательность своих элементов. Поэтому для получения неотрицательных малоранговых аппроксимаций неотрицательных матриц применяют методы попеременных проекций.

В данной работе рассматриваются детерминированные и рандомизированные алгоритмы попеременных проекций для получения неотрицательных малоранговых аппроксимаций неотрицательных матриц, в частности, показывается, что рандомизированные версии лучше детерминированных. Также производится обобщение этих методов на тензоры.

2 Методы попеременных проекций для матриц

Неотрицательная малоранговая аппроксимация неотрицательной матрицы может быть получена [3] попеременным проецированием матрицы на множества $\mathcal{M}_{\leq r} = \{X \in \mathbb{R}^{m \times n} : \text{rank } X \leq r\}$ и $\mathbb{R}_+^{m \times n}$. На этой идее основан первый алгоритм [1, Algorithm 1]. Из теоремы Эккарта-Янга получаем

$$\Pi_{\mathcal{M}_{\leq r}}(X) = \operatorname{argmin}_{Y \in \mathcal{M}_{\leq r}} \|X - Y\|_F = SVD_r(X) = U_r \Sigma_r V_r^T,$$

где $U_r \in \mathbb{R}^{m \times r}$, $V_r \in \mathbb{R}^{n \times r}$, $\Sigma_r \in \mathbb{R}^{r \times r}$.

А проецирование на $\mathbb{R}_+^{m \times n}$ осуществляется обнулением отрицательных элементов матрицы

$$\Pi_{\mathbb{R}_+^{m \times n}}(X) = \operatorname{argmin}_{Y \in \mathbb{R}_+^{m \times n}} \|X - Y\|_F = \max(X, 0) = \{\max(x_{ij}, 0)\}.$$

Можно также проецировать матрицу на $\mathcal{M}_{\leq r}$ не сразу, а через касательное пространство [1, Algorithm 2]. Такой трюк позволяет заменить точную проекцию на более быстрые итерации [6].

Узким местом точных алгоритмов является вычисление сингулярного разложения. Поэтому можно попробовать от точных вычислений перейти к приближённым. Это приводит нас к скетчингу и рандомизированным алгоритмам [1, Section 2.2]. В [1, A.4] показано, что рандомизированные алгоритмы более быстрые, а [1, Section 3] и раздел 5 данной работы показывают, что сходятся они при этом не хуже детерминированных.

3 Скетчинг

Смысл использования скетчинга в методах попеременных проекций заключается в том, чтобы вычислять факторизацию не самой матрицы, а её приближения, полученного с помощью скетчинга. Это позволяет сэкономить на вычислениях за счёт гораздо меньших размеров скетча матрицы по сравнению с исходной матрицей. Опишем один из вариантов скетчинга с одной случайной матрицей [2]. Процесс можно разделить на два вычислительных этапа:

Этап 1: Приближённое вычисление образа входной матрицы $A \in \mathbb{R}^{m \times n}$. Другими словами, ищется матрица Q : $Q^*Q = I$, $\|A - QQ^*A\| < \varepsilon$ при некотором малом $\varepsilon > 0$. Поиск матрицы Q можно выполнить очень эффективно, используя случайные вектора. Для этого берутся k случайных векторов $\{w^{(i)}, i = \overline{1, k}\}$ из некоторого распределения и отображаются матрицей A :

$$y^{(i)} = Aw^{(i)}, \quad i = \overline{1, k}.$$

Матрица Q получается из QR -разложения матрицы $Y = [y^{(1)}, y^{(2)}, \dots, y^{(k)}]$.

Этап 2: Вычисление факторизации (SVD , QR и т.д.) с использованием матрицы Q . Делается это за 3 простых шага:

1. Вычисление $B := Q^*A$, $B \in \mathbb{R}^{k \times n}$.
2. Вычисление факторизации (например, SVD) маленькой матрицы: $B = \tilde{U}\Sigma V^*$.
3. $U = Q\tilde{U}$.

Обоснование тому, почему матрица Q , полученная таким способом, удовлетворяет указанным требованиям, может дать следующая теорема [2, Theorem 1.1].

Теорема. Пусть матрица $A \in \mathbb{R}^{m \times n}$ имеет сингулярные числа $\sigma_1 \geq \sigma_2 \geq \dots$. Пусть задан целевой ранг $k \geq 2$ и параметр $p \geq 2$, где $k + p \leq \min\{m, n\}$. Формируется тестовая матрица $\Omega \in \mathbb{R}^{n \times (k+p)}$ из нормального распределения, и $Y = A\Omega$. Матрица Q получается из QR -разложения матрицы Y . Тогда для матожидания ошибки аппроксимации справедливо неравенство

$$\mathbb{E}\|A - QQ^*A\| \leq \left(1 + \frac{4\sqrt{k+p}}{p-1} \cdot \sqrt{\min m, n}\right) \sigma_{k+1}.$$

4 Методы попеременных проекций для тензоров

4.1 Базовые понятия

Сначала введём некоторые понятия, которые понадобятся далее.

Под *тензором* размерности d будем понимать многомерный массив чисел. Элемент тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ в позиции (i_1, \dots, i_d) будем обозначать как $A(i_1, \dots, i_d)$.

Определение. Матрицей развёртки тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ по k -ой моде называется матрица $A_{(k)} \in \mathbb{R}^{n_k \times n_1 \dots n_{k-1} n_{k+1} \dots n_d}$, которая определяется как

$$A_{(k)}(i_k, i_1 \dots i_{k-1} i_{k+1} \dots i_d) = A(i_1, \dots, i_d).$$

Определение. Произведением тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_k \times \dots \times n_d}$ на матрицу $B \in \mathbb{R}^{n_k \times m_k}$ по k -ой моде называется тензор $C \in \mathbb{R}^{n_1 \times \dots \times m_k \times \dots \times n_d}$ с элементами

$$C(i_1, \dots, j_k, \dots, i_d) \equiv (A \times_k B)(i_1, \dots, j_k, \dots, i_d) := \sum_{i_k=1}^{n_k} A(i_1, \dots, i_k, \dots, i_d) B(i_k, j_k).$$

4.2 Неотрицательное малоранговое разложение Таккера

Разложением Таккера называется представление тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ в виде

$$A = G \times_1 U_1 \cdots \times_d U_d,$$

где $G \in \mathbb{R}^{r_1 \times \dots \times r_d}$, $U_1 \in \mathbb{R}^{r_1 \times n_1}, \dots, U_d \in \mathbb{R}^{r_d \times n_d}$.

Разложение Таккера, в котором каждый из рангов разложения r_k является минимально возможным, называется *минимальным разложением Таккера*. Тензор G называется *ядром разложения*, а матрицы U_k называются *Таккеровскими факторами*.

Метод вычисления минимального разложения Таккера может дать следующая теорема [7, Теорема 1.13.1].

Теорема. Для заданного тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ минимальное разложение Таккера можно вычислить, выполнив следующие действия:

1. Положить $V^{(0)} = A$
2. Для $k := 1, 2, \dots, d$ найти минимальное скелетное разложение k -й матрицы развёртки тензора $V^{(k-1)}$:

$$V_{(k)}^{(k-1)}(i_k, \alpha_1 \dots \alpha_{k-1} i_{k+1} \dots i_d) = U_k(i_k, \alpha_k) V_{(k)}^{(k)}(\alpha_k, \alpha_1 \dots \alpha_{k-1} i_{k+1} \dots i_d)$$

3. Искомое разложение Таккера определяется матрицами $U_1(i_1, \alpha_1), \dots, U_d(i_d, \alpha_d)$ и ядром $G(\alpha_1, \dots, \alpha_d) = V^{(d)}(\alpha_1, \dots, \alpha_d)$.

На основе этого разложения можно получить аппроксимации данного тензора в виде ортогональных разложений Таккера с меньшими рангами [9, Теорема 2.1.7].

Теорема. Пусть дан тензор $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$, задача минимизации ошибки в формате Таккера с рангами (r_1, \dots, r_d) имеет решение, т.е.

$$\exists X_{min} : \|X_{min} - A\|_F = \inf_X \|X - A\|_F = \min_X \|X - A\|_F = \varepsilon,$$

где X и X_{min} представляются в формате Таккера с рангами (r_1, \dots, r_d) .

Квази-оптимальное приближение X_* , дающее точность

$$\|X_* - A\| \leq \sqrt{d}\varepsilon,$$

вычисляется с помощью d сингулярных разложений матриц развёртки.

На основе этих теорем можно получить метод попеременных проекций для тензоров. Обозначим \mathcal{F} – множество детерминированных и рандомизированных алгоритмов, вычисляющих усечённое сингулярное разложение, и пусть функция $\text{fold}_k: \mathbb{R}^{n_k \times n_1 \dots n_{k-1} n_{k+1} \dots n_d} \rightarrow \mathbb{R}^{n_1 \times \dots \times n_d}$, восстанавливает тензор из её k -ой матрицы развёртки. Будем попеременно проецировать тензор $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ на множества $\mathbb{R}_+^{n_1 \times \dots \times n_d} =$

$\{X \in \mathbb{R}^{n_1 \times \dots \times n_d} : X(i_1, \dots, i_d) \geq 0\}$ и $\mathcal{M}_{r_1, \dots, r_d} = \{X \in \mathbb{R}^{n_1 \times \dots \times n_d} : \text{rank}(X_{(k)}) \leq r_k, k = 1, \dots, d\}$, причём проецирование на множество $\mathcal{M}_{r_1, \dots, r_d}$ осуществляется вычислением d сингулярных разложений матриц развертки. Детали реализации алгоритма приведены в Алг. (1).

Algorithm 1: Метод попеременных проекций на основе HOSVD

Input: Тензор $X \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, мультилинейный ранг (r_1, r_2, \dots, r_d) , число итераций s и функция $\text{svdr} \in \mathcal{F}$, вычисляющая усечённое сингулярное разложение матрицы

Output: Неотрицательное усечённое разложение Таккера с рангами (r_1, r_2, \dots, r_d) тензора X , т.е. ядро G и Таккеровские факторы U_1, U_2, \dots, U_d

```

1  $X^{(0)} \leftarrow X$ 
2 for  $i = 1, 2, \dots, s$  do
3    $X^{(i)} \leftarrow \max(X^{(i-1)}, 0)$ 
4    $G \leftarrow X^{(i)}$ 
5   for  $k = 1, 2, \dots, d$  do
6      $[U_{r_k}, \Sigma_{r_k}, V_{r_k}^*] \leftarrow \text{svdr}(G_{(k)}, r_k)$ 
7      $U_k \leftarrow U_{r_k}^T \in \mathbb{R}^{r_k \times n_k}$ 
8      $Y \leftarrow \Sigma_{r_k} V_{r_k}^* \in \mathbb{R}^{r_k \times r_1 \dots r_{k-1} n_{k+1} \dots n_d}$ 
9      $G \leftarrow \text{fold}_k(Y) \in \mathbb{R}^{r_1 \times \dots \times r_k \times n_{k+1} \times \dots \times n_d}$ 
10   $X^{(i)} \leftarrow G \times_1 U_1 \dots \times_d U_d$ 
11 return  $G, U_1, U_2, \dots, U_d$ 

```

4.3 Неотрицательное малоранговое ТТ-разложение

ТТ-разложением называется представление тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ в виде

$$A(i_1, \dots, i_d) = \sum_{\alpha_0, \dots, \alpha_d} G_1(\alpha_0, i_1, \alpha_1) G_2(\alpha_1, i_2, \alpha_2) \dots G_d(\alpha_{d-1}, i_d, \alpha_d),$$

где $G_k \in \mathbb{R}^{r_{k-1} \times n_k \times r_k}$ называются *ТТ-ядрами*, причём $r_0 = 1$ и $r_d = 1$. Минимально возможные числа r_k называются *ТТ-рангами*.

Метод вычисления неотрицательных малоранговых аппроксимаций тензоров можно также получить на основе алгоритма ТТ-SVD [10]. Для этого достаточно попеременно проецировать тензор на множество неотрицательных тензоров и на множество \mathcal{M} тензоров с ТТ-рангами, не превосходящими целевые ранги. Проецирование на \mathcal{M} осуществляется вычислением на каждом шаге алгоритма ТТ-SVD усечённого сингулярного разложения. Детали реализации алгоритма приведены в Алг. (2).

Algorithm 2: Метод попеременных проекций на основе TT-SVD

Input: Тензор $X \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, мультилинейный ранг $(r_1, r_2, \dots, r_{d-1})$, число итераций s и функция $\text{svdr} \in \mathcal{F}$, вычисляющая усечённое сингулярное разложение матрицы

Output: Неотрицательное усечённое TT-разложение с TT-рангами $(r_1, r_2, \dots, r_{d-1})$ тензора X , т.е. ядра G_1, \dots, G_d

```
1  $X^{(0)} \leftarrow X$ 
2 for  $i = 1, 2, \dots, s$  do
3    $X^{(i)} \leftarrow \max(X^{(i-1)}, 0)$ 
4    $G \leftarrow X^{(i)}$ 
5    $[U, \Sigma, V^*] \leftarrow \text{svdr}(G_{(1)}, r_1)$ ;
6    $G_1 \leftarrow U \in \mathbb{R}^{n_1 \times r_1}$ 
7    $G \leftarrow \Sigma V^* \in \mathbb{R}^{r_1 \times n_2 \dots n_d}$ 
8   for  $k = 2, \dots, d-1$  do
9      $G \leftarrow \text{reshape}(G, r_{k-1}n_k, n_{k+1} \dots n_d)$ 
10     $[U_{r_k}, \Sigma_{r_k}, V_{r_k}^*] \leftarrow \text{svdr}(G, r_k)$ 
11     $G_k \leftarrow \text{reshape}(U_{r_k}, r_{k-1}, n_k, r_k) \in \mathbb{R}^{r_{k-1} \times n_k \times r_k}$ 
12     $G \leftarrow \Sigma V_{r_k}^* \in \mathbb{R}^{r_k \times n_{k+1} \dots n_d}$ 
13   $G_d \leftarrow G \in \mathbb{R}^{r_{d-1} \times n_d}$ 
14   $X^{(i)} \leftarrow \sum_{\alpha_1, \dots, \alpha_{d-1}} G_1(i_1, \alpha_1) G_2(\alpha_1, i_2, \alpha_2) \dots G_{d-1}(\alpha_{d-2}, i_{d-1}, \alpha_{d-1}) G_d(\alpha_{d-1}, i_d)$ 
15 return  $G_1, \dots, G_d$ 
```

4.4 Оценки сложностей алгоритмов

Выведем на основе [1, A] асимптотические оценки сложностей алгоритмов 1 и 2 для одной итерации.

Начнём с первого алгоритма. С увеличением k размеры тензора G уменьшаются, поэтому для оценки сверху можно ограничиться первой итерацией. Пусть $n = \max\{n_1, \dots, n_d\}$, $\mathbf{r} = (r_1, \dots, r_d)$ – целевой ранг. Восстановление тензора из разложения Таккера занимает $O(n^d)$, и так как $G_{(1)} \in \mathbb{R}^{n_1 \times n_2 \dots n_d}$, при вычислении точных попеременных проекций получаем оценку

$$14n_2 \dots n_d n_1^2 + 8n_1^3 + 2n_1 \dots n_d r_1 + n_1 r_1 + r_1 n_2 \dots n_d + O(n^d) + n_1 \dots n_d = O(n^{d+1}).$$

При использовании рандомизации НМТ [2, Algs 4.4 and 5.1] с параметрами k и p и нормально распределённой случайной матрицей имеем оценку

$$\begin{aligned} & (4p + 4)n_1 \dots n_d k + 2n_1 \dots n_d r_1 + (4p + 6)(n_1 + n_2 \dots n_d)k^2 + n_2 \dots n_d r \\ & + 75n_2 \dots n_d k + \frac{8}{3}(7 - p)k^3 + r_1 n_2 \dots n_d + O(n^d) + n_1 \dots n_d = O(n^d). \end{aligned}$$

Такая же оценка справедлива и при всех остальных рассматриваемых нами рандомизированных факторизаторах.

Совершенно аналогично для второго алгоритма получаем оценку $O(n^{d+1})$ при вычислении точных проекций и оценку $O(n^d)$ при использовании рандомизации.

Таким образом, рандомизация снижает сложность с $O(n^{d+1})$ до $O(n^d)$.

5 Численные эксперименты

5.1 Случайная матрица из равномерного распределения

Возьмём случайную матрицу размера 256×256 с независимыми элементами из равномерного распределения на отрезке $[0, 1]$. Лучшая аппроксимация ранга 64, полученная усечённым сингулярным разложением, содержит много отрицательных элементов (см. Рис.1). Попробуем исправить это, используя методы попеременных проекций, т.е. вычислим аппроксимации ранга 64 с помощью алгоритмов **SVD** [1, Algorithm 1], **Tangent** [1, Algorithm 2], **HMT** [1, Algorithm 3], **Tropp** [1, Algorithm 4] и **GN** [1, Algorithm 5]. Убывание норм отрицательных элементов показано на Рис. 2. Видно, что рандомизированные и детерминированные алгоритмы обладают аналогичной сходимостью, и при этом рандомизированные алгоритмы, при правильном выборе параметров, быстрее детерминированных (см. Таб. 1).

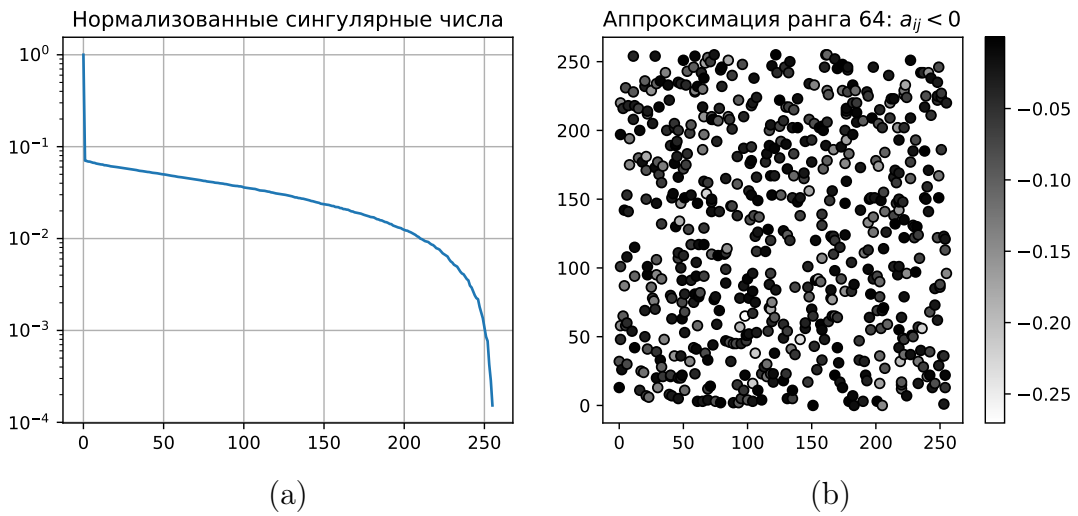
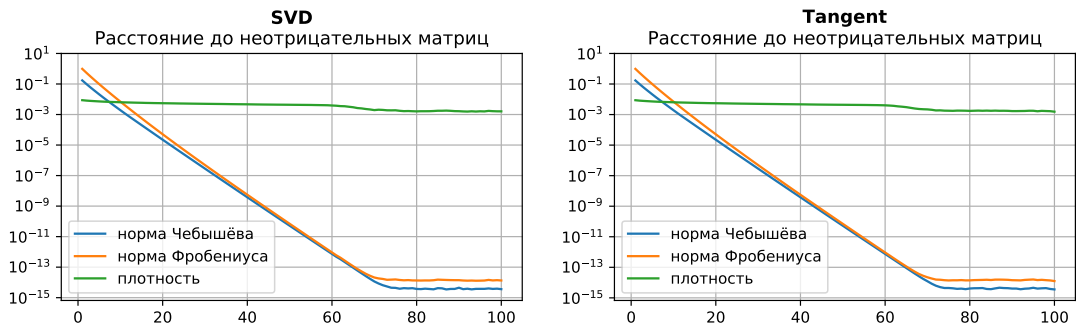


Рис. 1: Свойства случайной матрицы размера 256×256 с независимыми элементами из равномерного распределения на отрезке $[0, 1]$: нормализованные сингулярные числа (a) и отрицательные элементы наилучшего приближения ранга 64 (b).



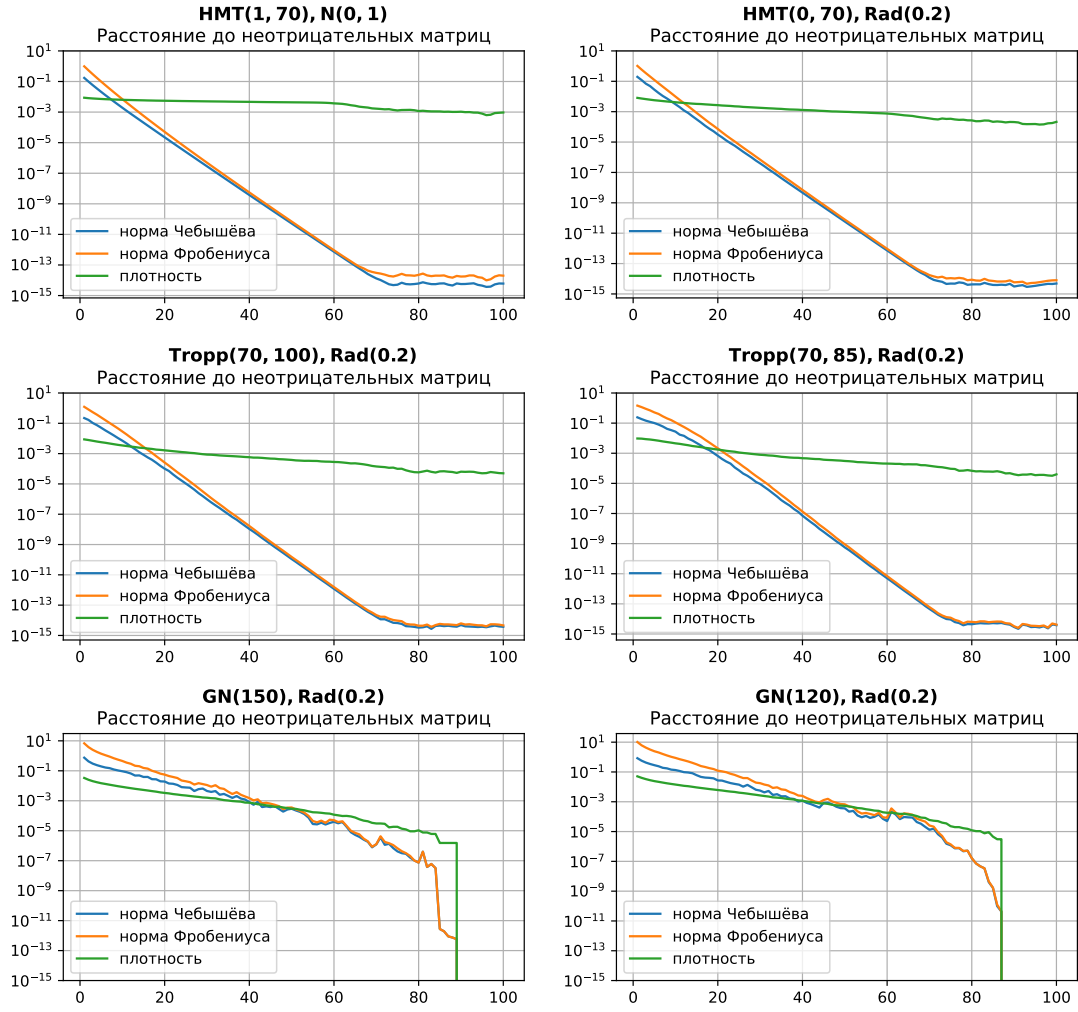


Рис. 2: Сравнение методов попеременных проекций для неотрицательной аппроксимации ранга 64 случайной матрицы размера 256×256 с независимыми элементами из равномерного распределения на отрезке $[0, 1]$: нормы Фробениуса и нормы Чебышёва отрицательных элементов и плотности отрицательных элементов за 100 итераций. Результаты усреднены по 10 экспериментам.

Метод	Скетч	Время (с.)	н. Фробениуса	н. Чебышёва
SVD_r	-	0.02	$3.09 \cdot 10^{-1}$	$7.34 \cdot 10^{-1}$
SVD	-	1.96	$3.09 \cdot 10^{-1}$	$7.34 \cdot 10^{-1}$
Tangent	-	1.06	$3.09 \cdot 10^{-1}$	$7.34 \cdot 10^{-1}$
HMT(1, 70)	$\mathcal{N}(0, 1)$	1.14	$3.09 \cdot 10^{-1}$	$7.34 \cdot 10^{-1}$
HMT(0, 70)	Rad(0.2)	0.82	$3.11 \cdot 10^{-1}$	$7.34 \cdot 10^{-1}$
Tropp(70, 100)	Rad(0.2)	1.05	$3.19 \cdot 10^{-1}$	$7.78 \cdot 10^{-1}$
Tropp(70, 85)	Rad(0.2)	0.96	$3.24 \cdot 10^{-1}$	$8.05 \cdot 10^{-1}$
GN(150)	Rad(0.2)	0.78	$5.06 \cdot 10^{-1}$	$11.13 \cdot 10^{-1}$
GN(120)	Rad(0.2)	0.63	$5.62 \cdot 10^{-1}$	$13.44 \cdot 10^{-1}$

Таблица 1: Сравнение методов попеременных проекций для неотрицательной аппроксимации ранга 64 случайной матрицы размера 256×256 с независимыми элементами из равномерного распределения на отрезке $[0, 1]$: времена вычислений и относительные ошибки в нормах Фробениуса и Чебышёва после 100 итераций.

5.2 Чёрно-белое изображение

Методы попеременных проекций могут быть легко модифицированы для вычисления матриц низкого ранга с элементами, лежащими в заданном диапазоне. В качестве примера возьмём черно-белое изображение размера 512×512 и вычислим аппроксимации ранга 50, элементы которых лежат на отрезке $[0, 1]$. Наилучшее приближение ранга 50, полученное усечённым сингулярным разложением, содержит элементы, как меньшие 0, так и большие 1 (см. Рис 3). Видно, что рандомизированные и детерминированные алгоритмы обладают аналогичной сходимостью (см. Рис 4), но при этом рандомизированные алгоритмы быстрее детерминированных (см. Таб. 2). На Рис. 5 можно визуальнo сравнить полученные аппроксимации.

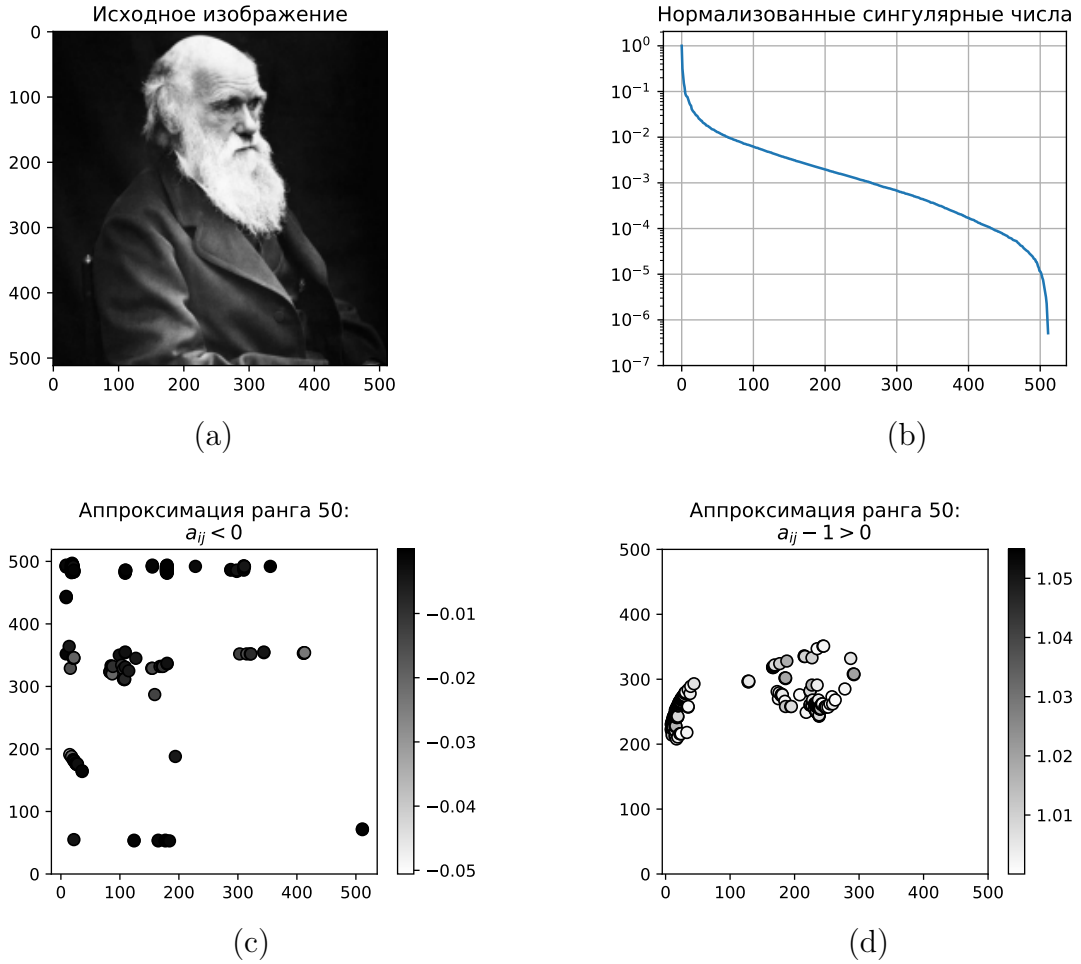
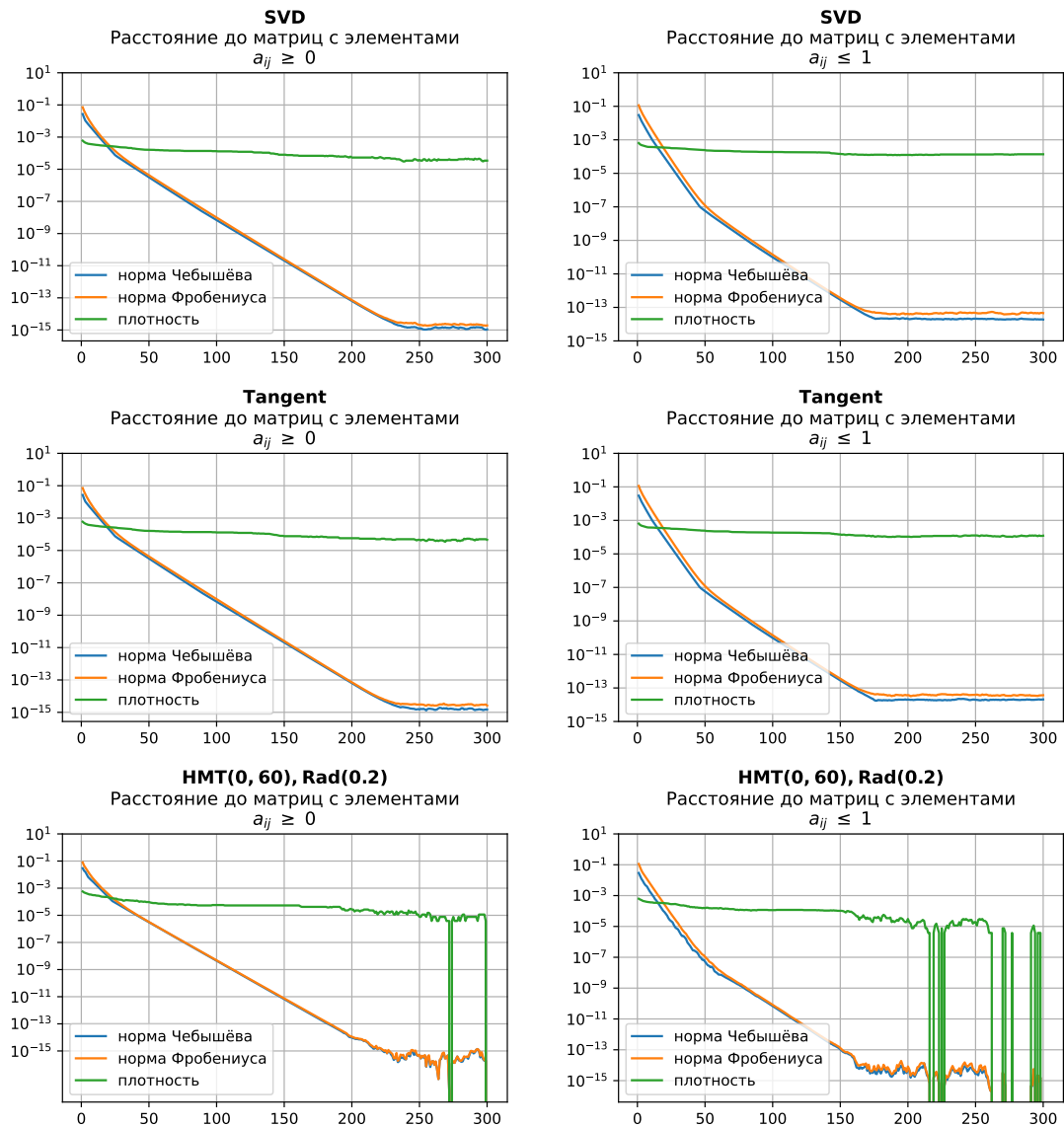


Рис. 3: Свойства изображения размера 512×512 : исходное изображение (a), нормализованные сингулярные числа (b), отрицательные элементы наилучшей аппроксимации ранга 50 (c) и элементы, большие 1, в наилучшей аппроксимации ранга 50 (d).

Метод	Скетч	Время (с.)	н. Фробениуса	н. Чебышёва
SVD_r	-	0.11	$6.82 \cdot 10^{-2}$	$3.76 \cdot 10^{-1}$
SVD	-	21.1	$6.83 \cdot 10^{-2}$	$3.76 \cdot 10^{-1}$
Tangent	-	6.81	$6.83 \cdot 10^{-2}$	$3.77 \cdot 10^{-1}$
HMT(0, 60)	Rad(0.2)	3.15	$3.09 \cdot 10^{-1}$	$7.34 \cdot 10^{-1}$
Tropp(60, 120)	Rad(0.2)	5.52	$6.84 \cdot 10^{-2}$	$3.76 \cdot 10^{-1}$
GN(340)	Rad(0.2)	6.38	$1.62 \cdot 10^{-1}$	$5.16 \cdot 10^{-1}$

Таблица 2: Сравнение методов попеременных проекций для неотрицательной аппроксимации ранга 50 чёрно-белого изображения размера 512×512 : время вычислений и относительные ошибки в нормах Фробениуса и Чебышёва после 300 итераций.



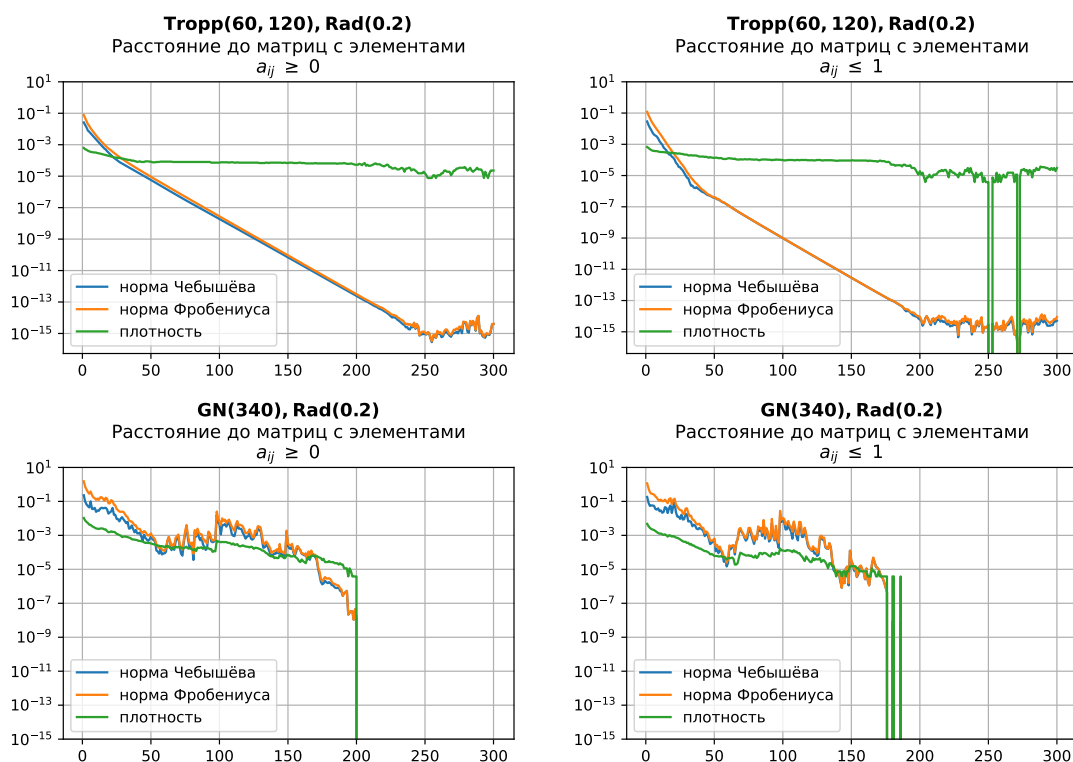


Рис. 4: Сравнение методов попеременных проекций для неотрицательной аппроксимации ранга 50 чёрно-белого изображения размера 512×512 : нормы Фробениуса и нормы Чебышёва отрицательных элементов и плотности отрицательных элементов за 300 итераций (слева), то же самое для элементов, больших 1 (справа).

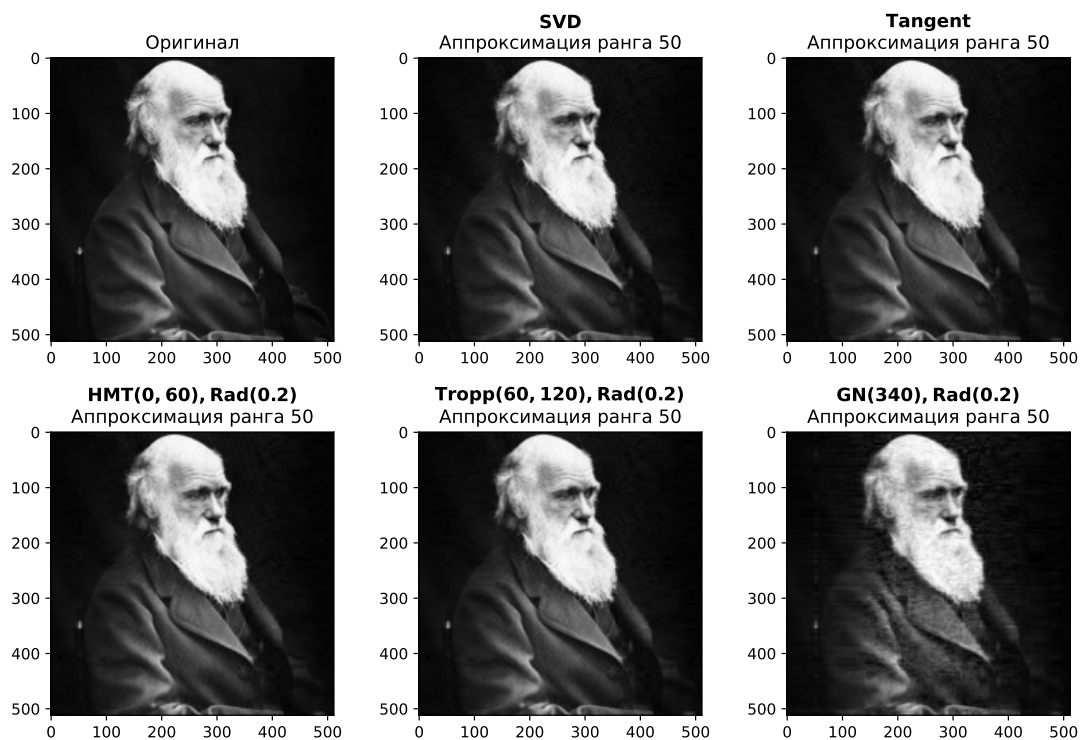


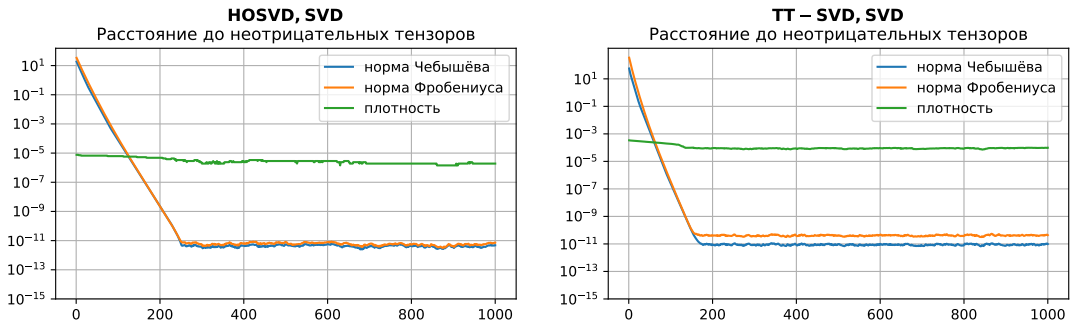
Рис. 5: Сравнение неотрицательных аппроксимаций ранга 50.

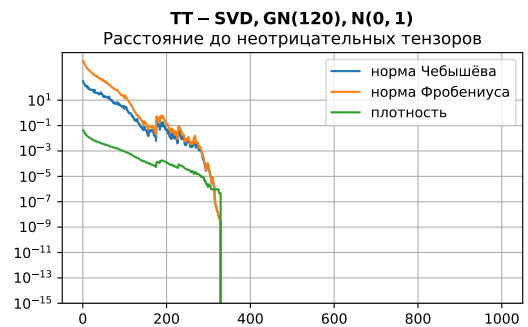
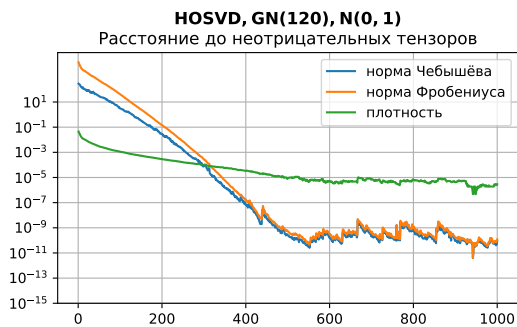
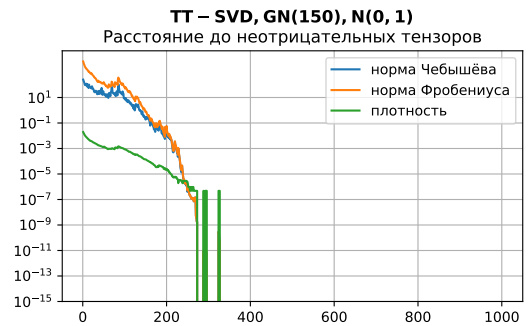
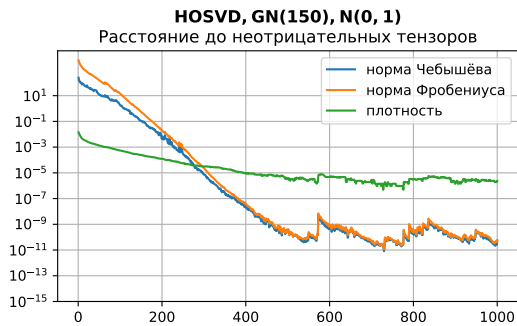
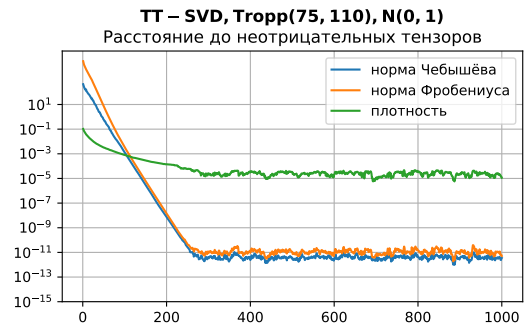
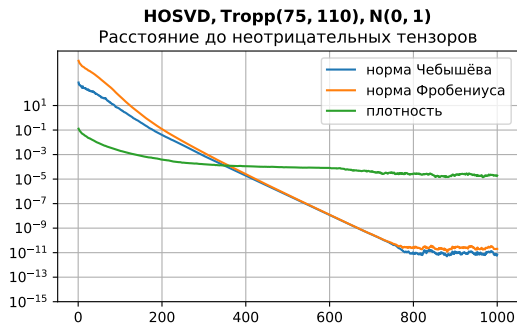
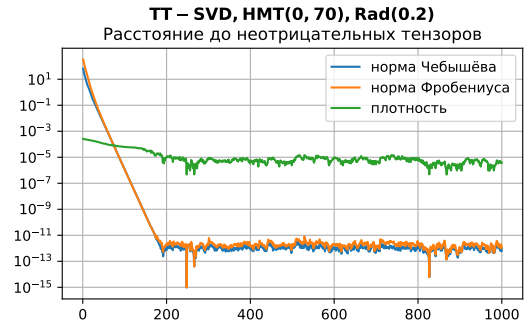
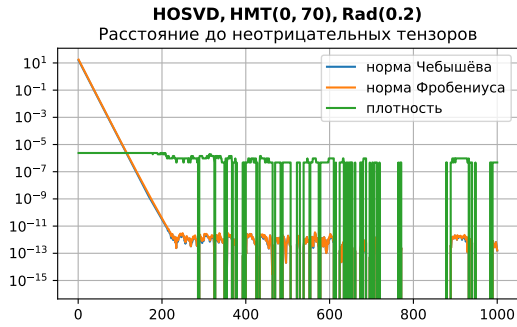
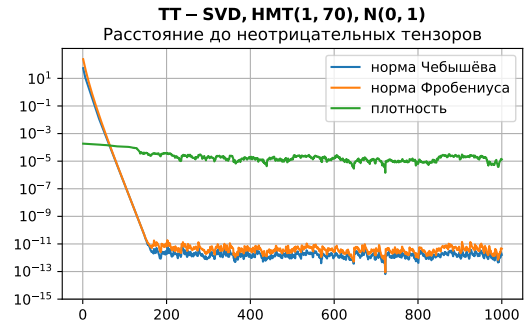
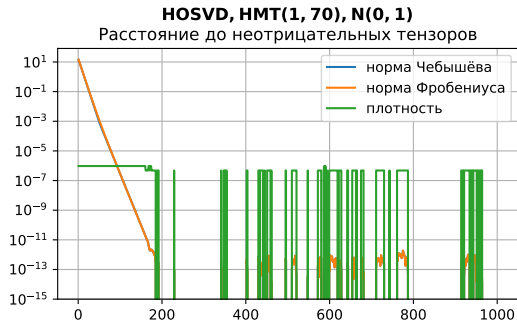
5.3 Случайный трёхмерный тензор из равномерного распределения

Возьмём случайный трёхмерный тензор размеров $128 \times 128 \times 128$ с независимыми элементами из равномерного распределения на отрезке $[0, 255]$ и вычислим неотрицательные аппроксимации мультилинейного ранга $(60, 60, 60)$ с помощью алгоритма 1 и аппроксимации мультилинейного ранга $(60, 60)$ с помощью алгоритма 2. В Таблице 3 можно видеть, что использование рандомизации даёт значительный прирост производительности. Интересно отметить, что лучшая производительность достигается при использовании рандомизированного факторизатора GN. Убывание норм отрицательных элементов показано на Рис. 6.

Метод	Скетч	Время (с.)	н. Фробениуса	н. Чебышёва
HOSVD	-	0.63	$4.60 \cdot 10^{-1}$	$8.98 \cdot 10^{-1}$
TT-SVD	-	0.49	$4.29 \cdot 10^{-1}$	$9.98 \cdot 10^{-1}$
HOSVD, SVD	-	641	$4.62 \cdot 10^{-1}$	$8.98 \cdot 10^{-1}$
HOSVD, HMT(1, 70)	$\mathcal{N}(0, 1)$	522	$4.69 \cdot 10^{-1}$	$8.66 \cdot 10^{-1}$
HOSVD, HMT(0, 70)	Rad(0.2)	326	$4.73 \cdot 10^{-1}$	$8.94 \cdot 10^{-1}$
HOSVD, Tropp(75, 110)	$\mathcal{N}(0, 1)$	376	$9.16 \cdot 10^{-1}$	3.35
HOSVD, GN(150)	$\mathcal{N}(0, 1)$	272	$5.74 \cdot 10^{-1}$	1.41
HOSVD, GN(120)	$\mathcal{N}(0, 1)$	243	$6.22 \cdot 10^{-1}$	1.67
HOSVD, GN(150)	Rad(0.2)	200	$5.75 \cdot 10^{-1}$	1.48
HOSVD, GN(120)	Rad(0.2)	181	$6.21 \cdot 10^{-1}$	1.78
TT-SVD, SVD	-	622	$4.29 \cdot 10^{-1}$	$9.68 \cdot 10^{-1}$
TT-SVD, HMT(1, 70)	$\mathcal{N}(0, 1)$	365	$4.37 \cdot 10^{-1}$	$9.21 \cdot 10^{-1}$
TT-SVD, HMT(0, 70)	Rad(0.2)	245	$4.41 \cdot 10^{-1}$	$1.02 \cdot 10^{-1}$
TT-SVD, Tropp(70, 100)	Rad(0.2)	275	$9.01 \cdot 10^{-1}$	3.62
TT-SVD, GN(150)	$\mathcal{N}(0, 1)$	281	$5.63 \cdot 10^{-1}$	1.39
TT-SVD, GN(120)	$\mathcal{N}(0, 1)$	253	$6.16 \cdot 10^{-1}$	1.69

Таблица 3: Сравнение методов попеременных проекций для неотрицательной аппроксимации мультилинейного ранга $(60, 60, 60)$ для HOSVD и мультилинейного ранга $(60, 60)$ для TT-SVD случайного трёхмерного тензора размеров $128 \times 128 \times 128$ с независимыми элементами из равномерного распределения на отрезке $[0, 255]$: времени вычислений и относительные ошибки в нормах Фробениуса и Чебышёва после 1000 итераций.





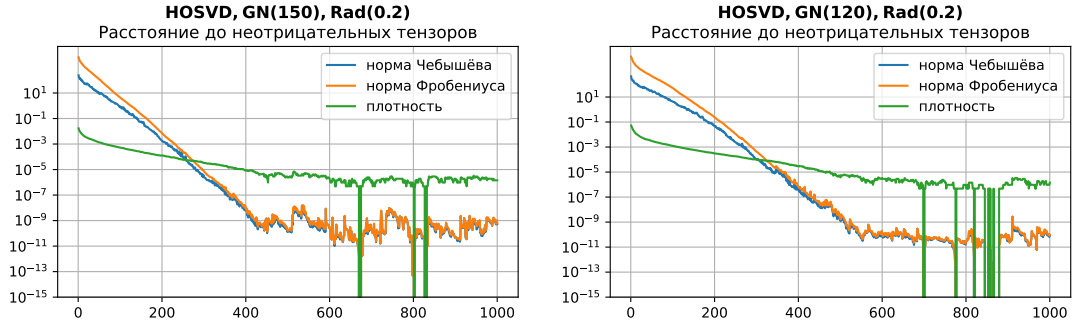


Рис. 6: Сравнение методов попеременных проекций для неотрицательной аппроксимации мультилинейного ранга $(60, 60, 60)$ для HOSVD и мультилинейного ранга $(60, 60)$ для TT-SVD случайного трёхмерного тензора размеров $128 \times 128 \times 128$ с независимыми элементами из равномерного распределения на отрезке $[0, 255]$: нормы Фробениуса и нормы Чебышева отрицательных элементов и плотности отрицательных элементов за 1000 итераций.

5.4 Цветное изображение

Как и в случае матриц, алгоритмы попеременных проекций 1 и 2 могут быть легко модифицированы для вычисления тензорных аппроксимаций с элементами, лежащими в заданном диапазоне. Возьмём в качестве примера цветное изображение размера 512×512 . В RGB модели изображению будет соответствовать трёхмерный тензор размера $512 \times 512 \times 3$ с элементами, лежащими на отрезке $[0, 1]$. Вычислим с помощью точных неитерационных алгоритмов HOSVD и TT-SVD аппроксимации мультилинейных рангов $(60, 60, 3)$ и $(60, 60)$ соответственно. Полученные аппроксимации содержат, как отрицательные элементы, так и элементы, большие 1 (см. Рис. 7). Исправим это, используя модифицированные алгоритмы 1 и 2. Результаты представлены в Таблице 4 и на Рис. 8. На Рис. 9 можно визуальнo сравнить полученные аппроксимации.

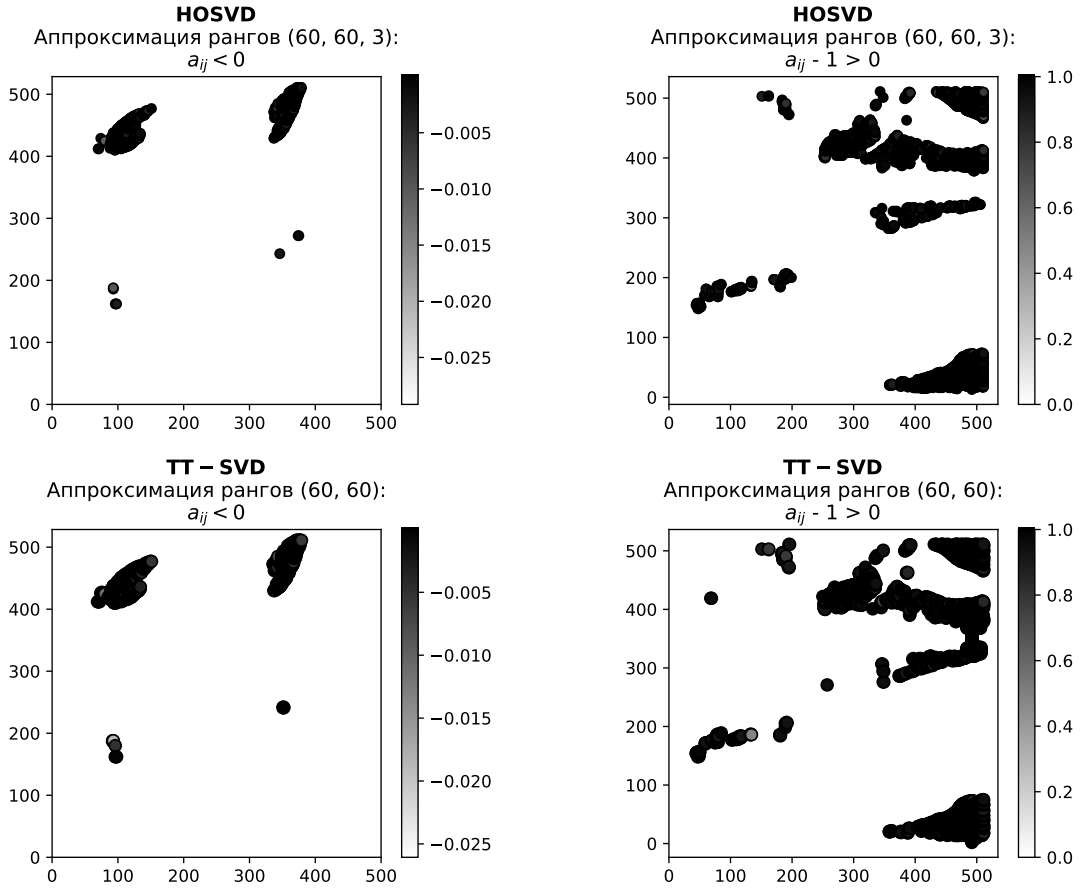
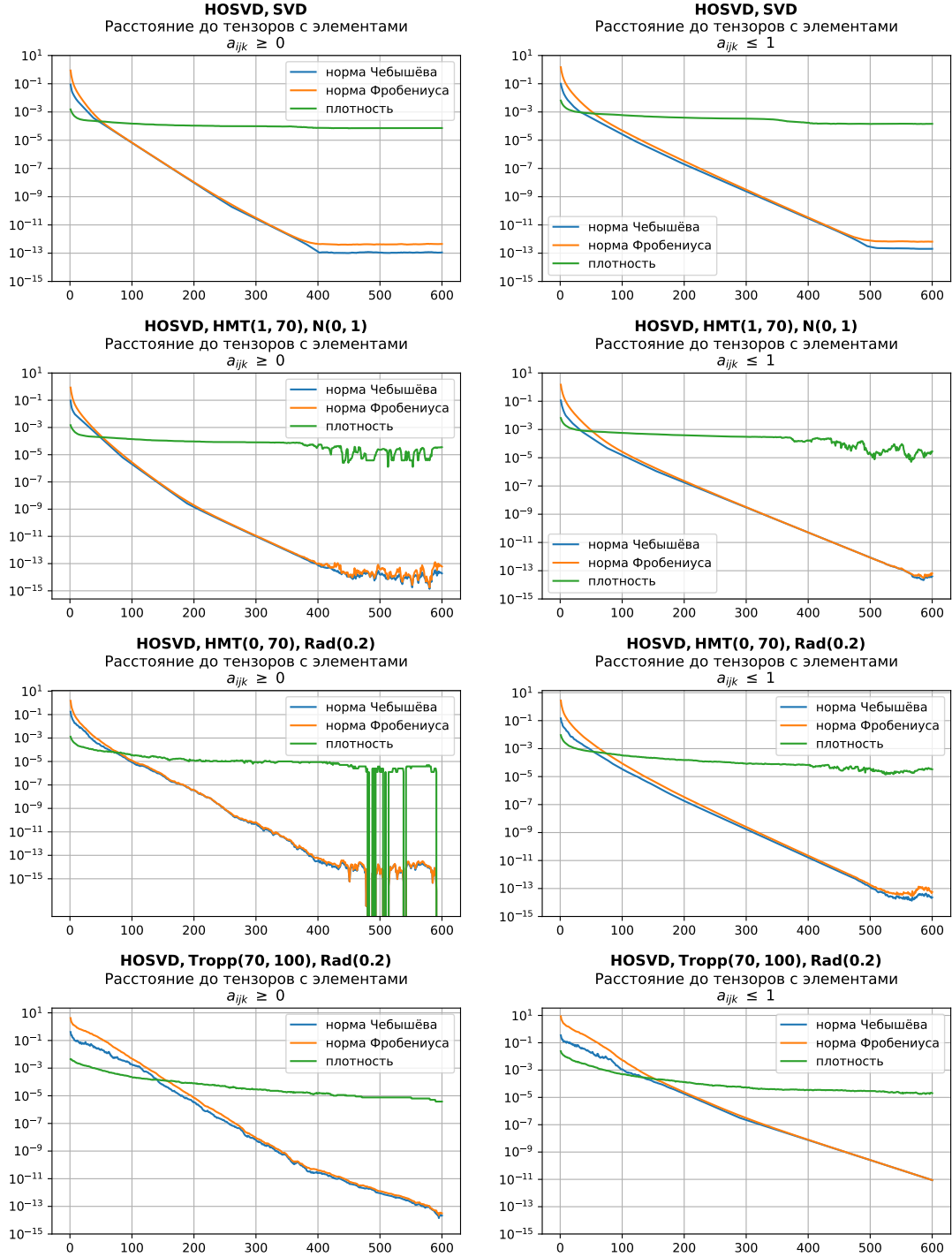
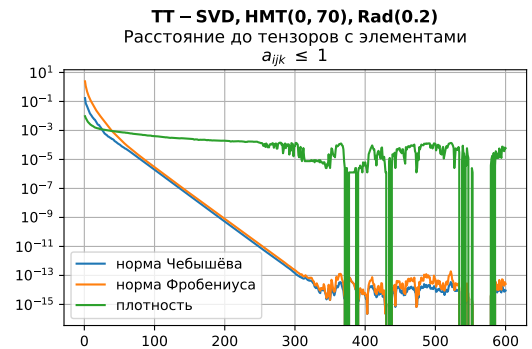
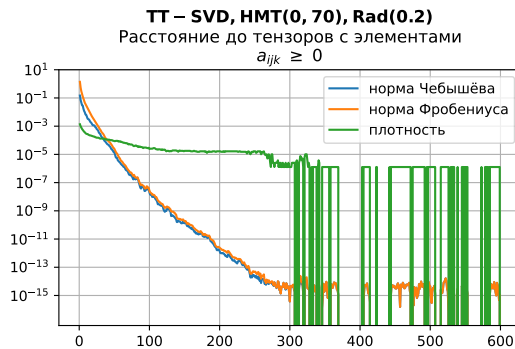
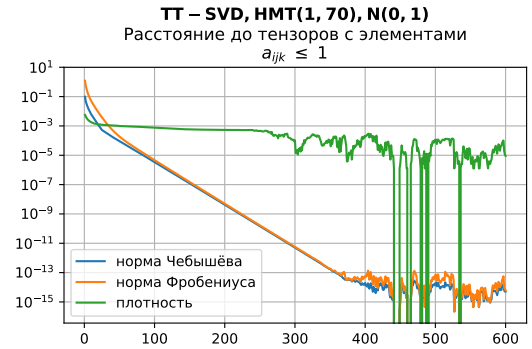
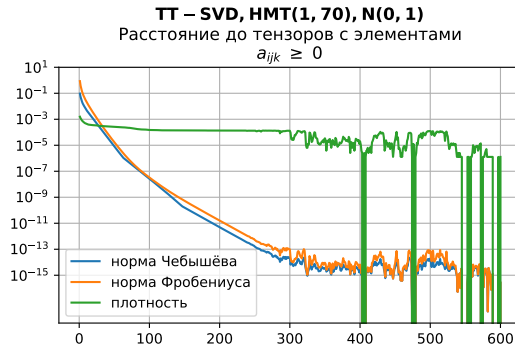
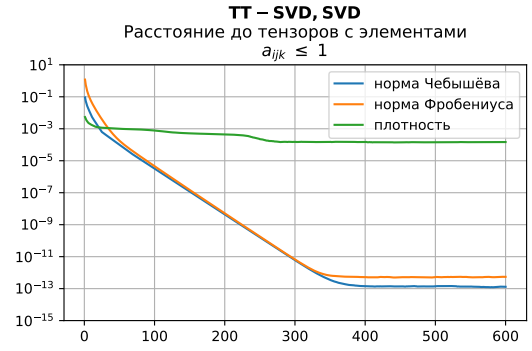
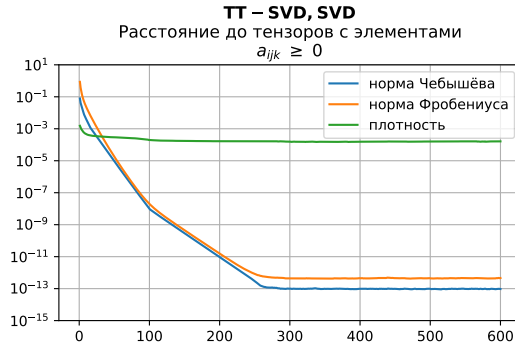
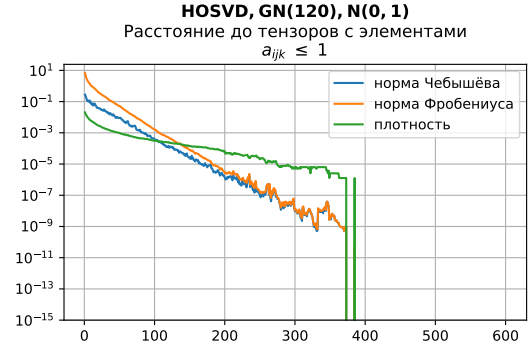
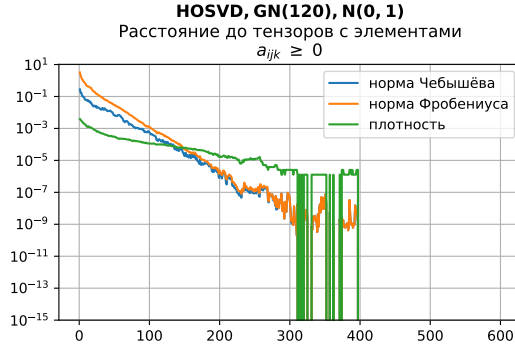
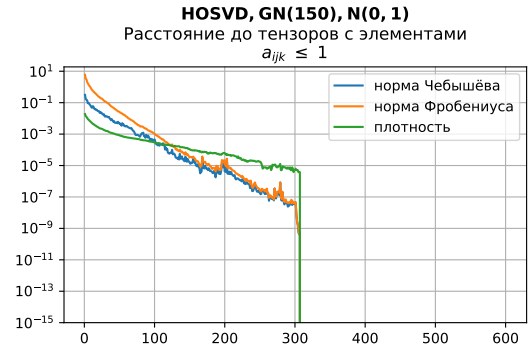
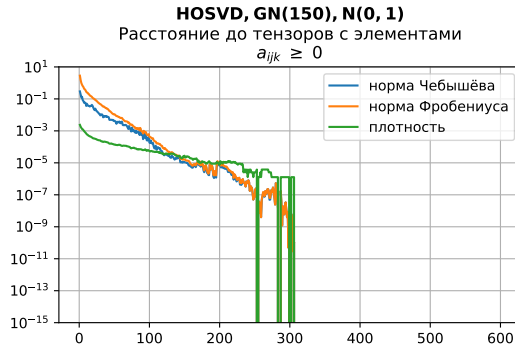


Рис. 7: Усреднённые значения отрицательных элементов и элементов, больших 1, в каждом пикселе аппроксимации, полученной алгоритмом HOSVD (сверху), то же самое для аппроксимации, полученной алгоритмом TT-SVD (снизу).

Метод	Скетч	Время (с.)	н. Фробениуса	н. Чебышёва
HOSVD	-	0.40	$3.75 \cdot 10^{-2}$	$2.95 \cdot 10^{-1}$
TT-SVD	-	0.29	$3.36 \cdot 10^{-2}$	$2.73 \cdot 10^{-1}$
HOSVD, SVD	-	189	$3.78 \cdot 10^{-2}$	$3.02 \cdot 10^{-1}$
HOSVD, HMT(1, 70)	$\mathcal{N}(0, 1)$	36	$3.87 \cdot 10^{-2}$	$2.95 \cdot 10^{-1}$
HOSVD, HMT(0, 70)	Rad(0.2)	23	$5.73 \cdot 10^{-2}$	$3.91 \cdot 10^{-1}$
HOSVD, Tropp(70, 100)	Rad(0.2)	26	$1.32 \cdot 10^{-1}$	$6.42 \cdot 10^{-1}$
HOSVD, GN(150)	$\mathcal{N}(0, 1)$	30	$9.62 \cdot 10^{-2}$	$5.79 \cdot 10^{-1}$
HOSVD, GN(120)	$\mathcal{N}(0, 1)$	22	$1.10 \cdot 10^{-1}$	$6.23 \cdot 10^{-1}$
HOSVD, GN(150)	Rad(0.2)	20	$9.25 \cdot 10^{-2}$	$6.15 \cdot 10^{-1}$
HOSVD, GN(120)	Rad(0.2)	19	$1.01 \cdot 10^{-1}$	$5.64 \cdot 10^{-1}$
TT-SVD, SVD	-	179	$3.39 \cdot 10^{-2}$	$2.88 \cdot 10^{-1}$
TT-SVD, HMT(1, 70)	$\mathcal{N}(0, 1)$	84	$3.50 \cdot 10^{-2}$	$2.72 \cdot 10^{-1}$
TT-SVD, HMT(0, 70)	Rad(0.2)	74	$5.26 \cdot 10^{-2}$	$3.48 \cdot 10^{-1}$
TT-SVD, Tropp(70, 100)	Rad(0.2)	138	$9.23 \cdot 10^{-2}$	$4.87 \cdot 10^{-1}$
TT-SVD, GN(150)	$\mathcal{N}(0, 1)$	223	$8.02 \cdot 10^{-2}$	$4.23 \cdot 10^{-1}$
TT-SVD, GN(120)	$\mathcal{N}(0, 1)$	228	$9.06 \cdot 10^{-2}$	$5.52 \cdot 10^{-1}$

Таблица 4: Сравнение методов попеременных проекций для неотрицательной аппроксимации мультилинейного ранга $(60, 60, 3)$ для HOSVD и мультилинейного ранга $(60, 60)$ для TT-SVD цветного изображения размера 512×512 : время вычислений и относительные ошибки в нормах Фробениуса и Чебышёва после 600 итераций.





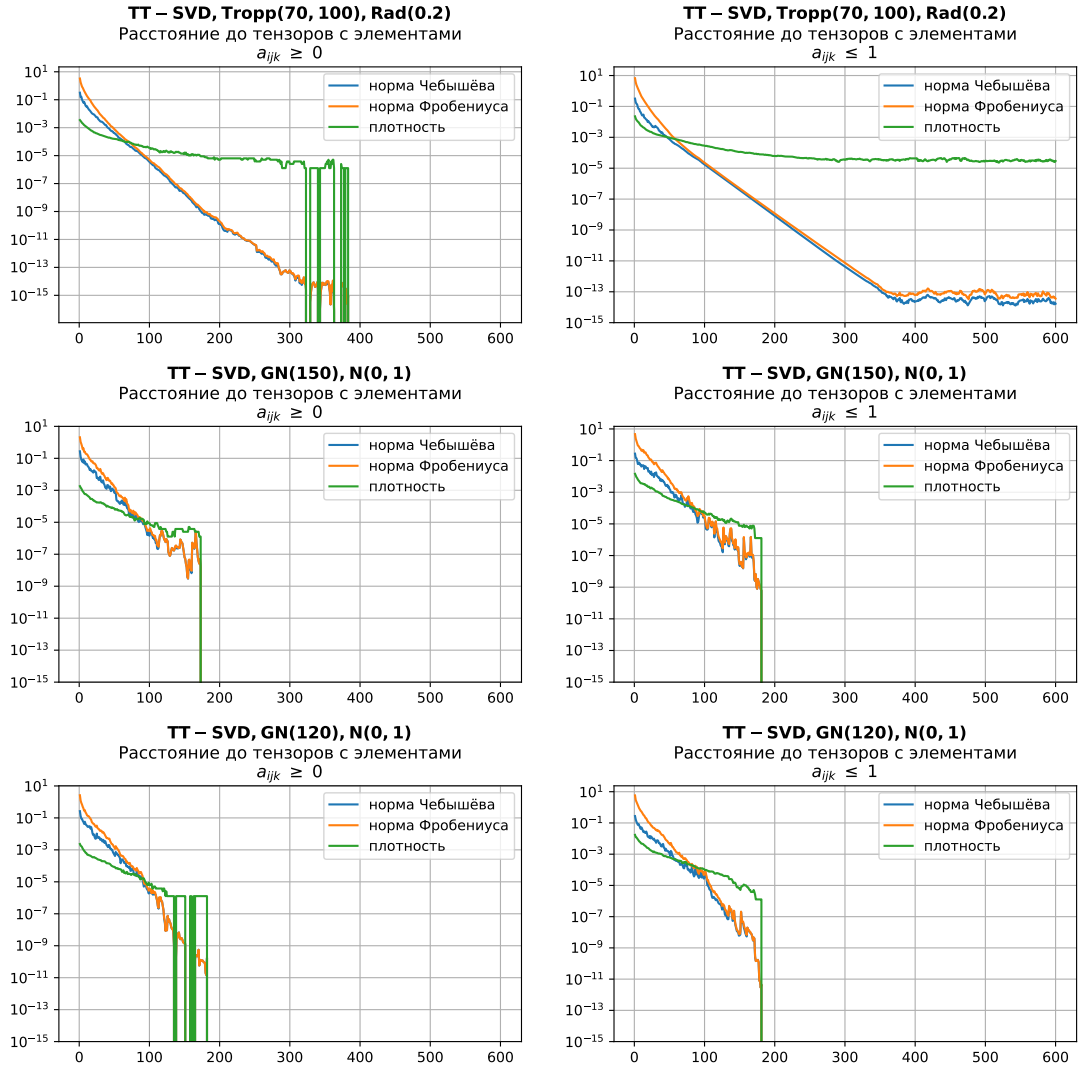


Рис. 8: Сравнение методов попеременных проекций для неотрицательной аппроксимации мультилинейного ранга $(60, 60, 3)$ для HOSVD и мультилинейного ранга $(60, 60)$ для TT-SVD цветного изображения размера 512×512 : нормы Фробениуса и нормы Чебышёва отрицательных элементов и плотности отрицательных элементов за 600 итераций.

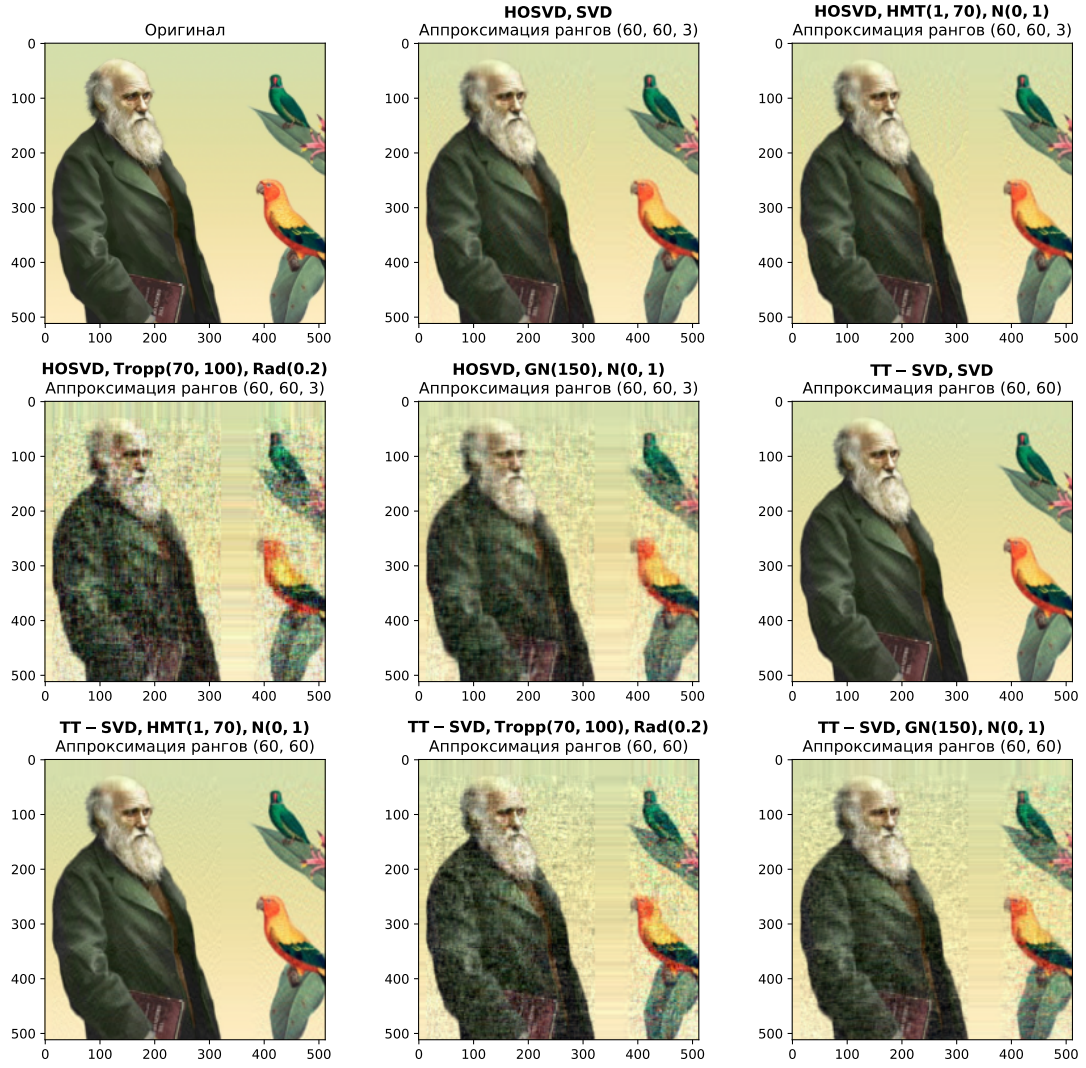


Рис. 9: Сравнение неотрицательных аппроксимаций мультилинейного ранга (60, 60, 3) для HOSVD и мультилинейного ранга (60, 60) для TT-SVD.

6 Заключение

Итак, численные эксперименты из статьи [1], повторенные в разделах 5.1 и 5.2, показывают, что использование рандомизации в методе попеременных проекций для матриц даёт прирост производительности, сохраняя при этом скорости сходимостей алгоритмов. Поэтому в данной работе было произведено обобщение метода и на тензоры. Оценки, полученные в разделе 4.4, показывают, что использование рандомизации в случае тензоров даёт более существенный прирост производительности, чем в случае матриц, так как сложность снижается с $O(n^{d+1})$ до $O(n^d)$. При этом в разделах 5.3 и 5.4 можно видеть, что правильный выбор методов скетчинга позволяет сохранить скорости сходимостей алгоритмов.

В дальнейшем планируется оптимизировать алгоритмы 1 и 2. Также для матриц представляется интересным поиск неточных более быстрых проекторов на \mathbb{R}_+ . Из экспериментов видно, что все алгоритмы сходятся, поэтому следует также получить и теоретические оценки сходимостей.

Список литературы

- [1] Sergey A. Matveev, Stanislav Budzinskiy, «Sketching for low-rank nonnegative matrix approximation: a numerical study».
- [2] N. Halko, P. G. Martinsson, and J. A. Tropp, «Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions,» SIAM Review, vol. 53, pp. 217–288, Jan. 2011
- [3] G.-J. Song and M. K. Ng, «Nonnegative low rank matrix approximation for nonnegative matrices,» Applied Mathematics Letters, vol. 105, p. 106300, July 2020.
- [4] J. A. Tropp, A. Yurtsever, M. Udell, and V. Cevher, «Practical sketching algorithms for lowrank matrix approximation,» SIAM Journal on Matrix Analysis and Applications, vol. 38, pp. 1454–1485, Jan. 2017.
- [5] Y. Nakatsukasa, «Fast and stable randomized low-rank matrix approximation,» arXiv:2009.11392 [cs, math], Sept. 2020.
- [6] G. Song, M. K. Ng, and T.-X. Jiang, «Tangent Space Based Alternating Projections for Nonnegative Low Rank Matrix Approximation,» arXiv:2009.03998 [cs, stat], Sept. 2020.
- [7] Е.Е.Тыртышников, «Матрицы, тензоры, вычисления».
- [8] S.Ahmadi-Asl, S.Abukovich, M.G.Asante-Mensah, A.Cichocki, A.H.Phan, T.Tanaka and I.Oseledets, «Randomized Algorithms for Computation of Tucker Decomposition and Higher Order SVD (HOSVD)».
- [9] С.В.Долгов, «Алгоритмы и применения тензорных разложений для численного решения многомерных нестационарных задач».
- [10] Ivan Oseledets, «Tensor-Train Decomposition».