

**ОБЩИЕ
ЧИСЛЕННЫЕ МЕТОДЫ**

УДК 519.65

**О ТТ-РАНГАХ ПРИБЛИЖЕННЫХ ТЕНЗОРИЗАЦИЙ НЕКОТОРЫХ
ГЛАДКИХ ФУНКЦИЙ¹⁾**

© 2021 г. Л. И. Высоцкий^{1,2}¹ 119333 Москва, ул. Губкина, 8, Ин-т вычисл. матем. им. Г.И. Марчука РАН, Россия² 119991 Москва, Ленинские горы, МГУ им. М.В. Ломоносова, ВМК, Россия

e-mail: vysotskylev@yandex.ru

Поступила в редакцию 24.11.2020 г.

Переработанный вариант 24.11.2020 г.

Принята к публикации 14.01.2021 г.

Исследуются “тензоризации” функций, т.е. тензоры с элементами $A(i_1, \dots, i_d) = f(x(i_1, \dots, i_d))$, где $f(x)$ — некоторая функция, заданная на отрезке, а $\{x(i_1, \dots, i_d)\}$ — сетка на этом отрезке. Для таких тензоров ставится задача приближения тензорами, допускающими ТТ (Tensor Train)-разложение с малыми ТТ-рангами. Для класса функций, являющихся следами аналитических в некоторых эллипсах на комплексной плоскости функций комплексного переменного, получены верхние и нижние оценки ТТ-рангов оптимальных приближений. Указанные оценки применены к тензоризациям полиномиальных функций. В частности, известная верхняя граница ТТ-рангов приближений таких функций улучшена до $O(\log n)$, где n — степень полинома. Библ. 10.

Ключевые слова: ТТ-разложение, tensor train, ТТ-ранги, тензоризации функций, приближения.

DOI: 10.31857/S0044466921050173

1. ВВЕДЕНИЕ

Относительно недавние успехи по преодолению “проклятия размерности” позволяют оперировать с дискретизациями функций на очень густых сетках. Например (в духе работ [1]–[5]), заданную на отрезке $[L, R]$ функцию $f(x)$ можно дискретизировать, т.е. превратить в вектор с элементами $f(x(i))$, где $\{x(i)\}_{i=0}^{N-1}$ — произвольная сетка на $[L, R]$. Если $N = n_1 \times \dots \times n_d$, то полученный вектор можно рассматривать (см. [6]) как тензор (многомерный массив) размером $n_1 \times \dots \times n_d$ с элементами:

$$A(i_1, \dots, i_d) = f(x(i_1 N_1 + \dots + i_d N_d)), \quad N_k = n_{k+1} \cdots n_d, \quad i_k \in \{0, \dots, n_k - 1\}.$$

Тензор A будем называть *тензоризацией* функции f на сетке $\{x(i)\}$. Этот тензор на практике часто допускает приближение другим тензором B , имеющим малопараметрическое представление. К примеру, можно использовать ТТ-разложение [2], [7] (tensor train, тензорный поезд):

$$B(i_1, \dots, i_d) = \sum_{\alpha_1, \dots, \alpha_{d-1}} H_1(i_1, \alpha_1) \cdots H_k(\alpha_{k-1}, i_k, \alpha_k) \cdots H_d(\alpha_{d-1}, i_d), \quad (1)$$

где H_k имеет размеры $r_{k-1} \times n_k \times r_k$, а индекс суммирования α_k пробегает значения от 1 до r_k . Числа r_k называются ТТ-рангами представления (1), причем для единообразия определения тензоров H_k считается, что $r_0 = r_d = 1$.

Стоит также отметить, что в случае тензоризаций высокой размерности, когда число N очень велико, вычисление и хранение в памяти элементов тензора A не представляется возможным. Тем не менее иногда есть возможность получить сразу представление (1), вычисляя $f(x(i))$ лишь для небольшого количества точек $x(i)$. Например, можно использовать *крестовое приближение* (TT-cross approximation) из [8].

¹⁾Работа выполнена при поддержке Московского центра фундаментальной и прикладной математики (соглашение 075-15-2019-1624 с Минобрнауки РФ).

Величины ТТ-рангов играют ключевую роль для применимости ТТ-разложения, ведь требуемая для хранения тензоров H_k память и сложность распространенных операций над тензорами в этом формате растут пропорционально полиномам малой степени от r_k [7]. Поэтому представляется интересным исследовать, насколько малы могут быть ТТ-ранги приближений тензоризаций функций.

Для некоторых классов функций известны компактные представления для тензоризаций, например, для экспоненты, синуса и некоторых других [1], [3]. Для тензоризации полинома степени n на равномерной сетке известно [2], [3], что все ее ТТ-ранги ограничены $n + 1$. В [2] было показано, что ТТ-ранги приближений тензоризаций так называемых *асимптотически гладких* (т.е. бесконечно дифференцируемых с “не слишком быстро” растущими при приближении к сингулярностям производными) функций ограничены числом, растущим при уменьшении требуемой погрешности ϵ как $-\log_2(\epsilon)$. В [5] были получены оценки ТТ-рангов приближений тензоризаций полиномов на равномерных сетках, улучшающие известную оценку $n + 1$.

В данной работе для тензоризаций вещественнозначных функций, являющихся следами аналитических в некотором эллипсе функций комплексного переменного, доказаны верхние и нижние оценки ТТ-рангов приближений. Полученные оценки были применены к тензоризациям полиномов, существенно улучшив известные результаты: вместо $O(\sqrt[3]{n})$ из [5] доказана оценка $O(\ln n)$.

2. НЕОБХОДИМЫЕ ОПРЕДЕЛЕНИЯ

Определение 1. Матрицами развертки тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ называются матрицы

$$A_k \in \mathbb{R}^{(n_1 \dots n_k) \times (n_{k+1} \dots n_d)}, \quad A_k(i_1, \dots, i_k; i_{k+1}, \dots, i_d) = A(i_1, \dots, i_d),$$

где группы индексов до и после точки с запятой образуют так называемые *мультииндексы*, отождествляемые со своим номером (нумерация с нуля) в лексикографическом (словарном) порядке.

Для тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ определим нормы

$$\|A\|_\infty = \max |A(i_1, \dots, i_d)| \quad \text{и} \quad \|A\|_F = \sqrt{\sum (A(i_1, \dots, i_d))^2}.$$

В основном мы будем рассматривать “кубические” тензоризации на равномерных сетках, т.е. тензоры, все размеры которого равны одному и тому же числу b , а элементы заданы значениями функции $f(x)$ в равноудаленных узлах. Такие d -мерные тензоризации с шагом $(R - L)b^{-d}$ на отрезке $[L, R]$ будем обозначать через $T_{b,d,[L,R]}(f)$.

Для ТТ-разложения вида (1) d -мерного тензора $B \in \mathbb{R}^{b \times \dots \times b}$ определим $\mathcal{R}(H_1, \dots, H_d)$ как максимальный из его ТТ-рангов. Далее определим $\mathcal{R}(B) = \min \mathcal{R}(H_1, \dots, H_d)$, где минимум берется по всевозможным ТТ-разложениям вида (1) тензора B . Для оценки ТТ-рангов наилучших приближений нам понадобится функция

$$\mathcal{R}_\epsilon(A) = \min_{B: \|B-A\|_F \leq \epsilon} \mathcal{R}(B).$$

Определение 2. Эллипс на комплексной плоскости с центром в нуле и осями, параллельными осям координат, чьи полуоси равны $(\rho + \rho^{-1})/2$ и $(\rho - \rho^{-1})/2$ для некоторого $\rho > 1$, будем называть *эллипсом Бернштейна* и обозначать через Γ_ρ . Также эллипс Бернштейна можно рассматривать (см. [9]) как образ окружности $\{\rho e^{i\varphi} : \varphi \in [0, 2\pi)\}$ под действием отображения Жуковского $w \rightarrow (w + w^{-1})/2$. Такой же эллипс, но с центром в точке $z \in \mathbb{C}$, будем обозначать через $\Gamma_\rho(z)$.

Определение 3. Чебышёвской сеткой на $[-1, 1]$ с числом узлов n будем называть упорядоченное множество точек $\{x_0, \dots, x_{n-1}\}$, где $x_j = \cos(\pi/(2n) + \pi j/n)$. Известно [9], что числа x_j являются корнями полинома Чебышёва степени n , т.е. $T_n(x) = \cos(\arccos(nx))$.

Круг радиуса r с центром в точке z на комплексной плоскости будем обозначать через $U(r, z)$.

3. ВЕРХНИЕ ОЦЕНКИ

Докажем некоторые верхние оценки ГТ-рангов и связанных с ними величин для приближений тензоризаций функций на равномерных сетках. Основная лемма 1 уточняет известную (см., например, [9]) оценку точности приближения функции, являющейся следом аналитической в эллипсе Бернштейна. Теорема 1 демонстрирует возможность низкорангового приближения матриц развертки тензоризаций функций, аналитических в круге, чей диаметр является отрезком дискретизации. Следствие 1 дает оценку непосредственно для величины $\mathcal{R}_\epsilon(T_{b,d,[L,R]}(f))$ для функций f с указанным свойством. Наконец, теорема 2 рассматривает более простой случай, когда функция является аналитической в достаточно большом эллипсе. Заметим, что теорему 1 может иметь смысл применять даже для функций, являющихся аналитическими в произвольно больших множествах, так как в оценку теоремы 2 входит максимум модуля $f(z)$ на некотором эллипсе – число, которое может быть нежелательно велико для интересующей нас функции $f(z)$. Эта особенность будет использована ниже при применении полученных результатов к полиномам.

Лемма 1. Пусть $f(x)$ – след аналитической внутри эллипса Бернштейна Γ_ρ , $\rho > 1$, функции $f(z)$, причем $|f(z)| \leq M$ на Γ_ρ . Пусть также $P_n(x)$ – полином Лагранжа степени n для $f(x)$ на чебышёвской сетке на $[-1, 1]$ с $(n + 1)$ узлом $\{x_0, \dots, x_n\}$. Тогда для всех $x \in [-1, 1]$ выполнено следующее:

$$|f(x) - P_n(x)| \leq \frac{M}{\rho^{n+1} - \rho^{-n-1}} \frac{\rho + \rho^{-1}}{\frac{1}{2}(\rho + \rho^{-1} - 1)}.$$

Доказательство. Зафиксируем произвольное $x \in [-1, 1] \setminus \{x_0, \dots, x_n\}$. Далее рассмотрим функцию

$$F(z) = \frac{f(z)}{(z - x) \prod_{j=0}^n (z - x_j)}$$

и, применяя теорему о вычетах и домножая на $\prod_j (x - x_j)$ (т.е. рассуждая аналогично [9, Теорема 13.6]), придем к представлению:

$$\prod_{j=0}^n (x - x_j) \int_{\Gamma_\rho} F(z) dz = 2\pi i \left(f(x) - \underbrace{\sum_{\ell=0}^n f(x_\ell) \frac{\prod_{j \neq \ell} (x - x_j)}{\prod_{j \neq \ell} (x_\ell - x_j)}}_{P_n(x)} \right).$$

Так как x_j суть в точности корни полинома Чебышёва $T_{n+1}(x)$ степени $n + 1$, то можно записать

$$T_{n+1}(x) \int_{\Gamma_\rho} \frac{f(z) dz}{(z - x) T_{n+1}(z)} = 2\pi i (f(x) - P_n(x)).$$

Используя доказанное в [9, Теорема 13.6] для $z \in \Gamma_\rho$ неравенство $|T_{n+1}(z)| \geq (\rho^{n+1} - \rho^{-n-1})/2$, а также очевидное для $x \in [-1, 1]$ неравенство $|T_{n+1}(x)| \leq 1$, при $x \in [-1, 1]$ получаем

$$\begin{aligned} |f(x) - P_n(x)| &\leq \frac{1}{2\pi} |T_{n+1}(x)| \left| \int_{\Gamma_\rho} \frac{f(z) dz}{(z - x) T_{n+1}(z)} \right| \leq \frac{1}{2\pi} \int_{\Gamma_\rho} \frac{M |dz|}{|z - x| |T_{n+1}(z)|} \leq \\ &\leq \frac{M}{2\pi \rho^{n+1} - \rho^{-n-1}} \int_{\Gamma_\rho} \frac{|dz|}{|z - x|} \leq \frac{M}{\pi(\rho^{n+1} - \rho^{-n-1})} \frac{|\Gamma_\rho|}{\min_{z \in \Gamma_\rho} |z - x|}. \end{aligned}$$

Для оценки длины эллипса Γ_ρ воспользуемся упомянутым фактом, что первый является образом окружности $S_\rho = \{\rho e^{i\varphi} : \varphi \in [0, 2\pi)\}$ под действием отображения $w \rightarrow (w + w^{-1})/2$:

$$|\Gamma_\rho| = \int_{\Gamma_\rho} |dz| = \int_{S_\rho} \left| \frac{1}{2} dw - \frac{1}{2} w^{-2} dw \right| = \int_0^{2\pi} \frac{1}{2} |1 - (\rho e^{i\varphi})^{-2}| \rho d\varphi \leq \frac{1}{2} (1 + \rho^{-2}) \rho 2\pi.$$

Далее вычислим $\min |z - x|^2$ для $z \in \Gamma_\rho$ и $x \in [-1, 1]$. Обозначим через r_1 и r_2 большую и меньшую полуоси эллипса Γ_ρ соответственно. Тогда точки этого эллипса параметризуются углом φ : $z = r_1 \cos \varphi + i r_2 \sin \varphi$. Преобразуем минимизируемое выражение

$$|z - x|^2 = (r_1 \cos \varphi - x)^2 + (r_2 \sin \varphi)^2 = (r_1^2 - r_2^2) \cos^2 \varphi - 2r_1 x \cos \varphi + x^2 + r_2^2. \tag{2}$$

В силу симметрии достаточно рассмотреть $x \geq 0$. При фиксированном x выражение (2) – это квадратный трехчлен относительно $\cos \varphi$, причем вершина соответствующей параболы имеет абсциссу $r_1 x / (r_1^2 - r_2^2) = r_1 x$. При $r_1 x \leq 1$ минимум по φ достигается при $\cos \varphi = r_1 x$ и равен r_2^2 . При $r_1 x > 1$ минимум достигается при $\cos \varphi = 1$ и равен $(r_1 - x)^2$, причем это выражение достигает минимума по x при $x = 1$. Так как $r_1 - 1 < r_2$, то и $\min |z - x| = r_1 - 1 = (\rho + \rho^{-1})/2 - 1$.

В итоге имеем

$$|f(x) - P_n(x)| \leq \frac{M \rho \pi (1 + \rho^{-2})}{\pi (\rho^{n+1} - \rho^{-n-1}) \left(\frac{1}{2} (\rho + \rho^{-1}) - 1 \right)} = \frac{M}{\rho^{n+1} - \rho^{-n-1}} \frac{\rho + \rho^{-1}}{\frac{1}{2} (\rho + \rho^{-1}) - 1}.$$

Теорема 1. Пусть $f: [L, R] \rightarrow \mathbb{R}$ является следом аналитической в круге с диаметром $[L, R]$ функции $f(z)$, причем $|f(z)| \leq M$ для всех z из этого круга.

Тогда для произвольного $\varepsilon > 0$, каждой матрицы развертки $A_k \in \mathbb{R}^{b^k \times b^{d-k}}$ тензора $T_{b,d,[L,R]}(f)$ и любого натурального $\mu \in [1, b^k - 1]$ существует матрица $B_k \in \mathbb{R}^{b^k \times b^{d-k}}$ такая, что

$$\text{rank } B_k \leq 2\mu + s + 1 \quad \text{и} \quad \|A_k - B_k\|_\infty \leq \varepsilon,$$

где $s = \lfloor \log_\rho(3M/\varepsilon + 1) \rfloor$, $\rho = 2\mu + 1 + 2\sqrt{\mu^2 + \mu}$.

Доказательство. Обозначим через h шаг дискретизации: $h = (R - L)b^{-d}$. Строка матрицы A_k с индексом j соответствует отрезку $[L + jh, L + (j + 1)h]$.

Рассмотрим функцию

$$g_j(y) = f\left(L + jh + \frac{1+y}{2}h\right), \quad y \in [-1, 1],$$

а также соответствующую функцию комплексного аргумента $g_j(w)$, $w \in U((\rho + \rho^{-1})/2, 0)$.

Отображение $w \mapsto \alpha + \beta w$ переводит круг $U((\rho + \rho^{-1})/2, 0)$ в $U(\beta(\rho + \rho^{-1})/2, \alpha)$. В нашем случае $\alpha = L + jh + h/2$, $\beta = h/2$.

Покажем, что при $j = \mu, \mu + 1, \dots, b^k - \mu - 1$ круг $U(\beta(\rho + \rho^{-1})/2, \alpha)$ лежит внутри круга с диаметром $[L, R]$, т.е.

$$\begin{aligned} L + jh + \frac{h}{2} - \frac{h}{4}(\rho + \rho^{-1}) &\geq L, \\ L + jh + \frac{h}{2} + \frac{h}{4}(\rho + \rho^{-1}) &\leq R. \end{aligned}$$

Действительно, для заданного в условии значения ρ верно равенство $\rho + \rho^{-1} = 4\mu + 2$, поэтому указанная пара условий переписывается в виде

$$\begin{cases} (j - \mu)h \geq 0, \\ L + (j + \mu + 1)h \leq R, \end{cases} \Leftrightarrow \begin{cases} j \geq \mu, \\ j \leq b^k - \mu - 1. \end{cases}$$

Внутри круга с диаметром $[L, R]$ по условию выполнено неравенство $|f(z)| \leq M$, поэтому и $|g_j(w)| \leq M$ для $w \in U((\rho + \rho^{-1})/2, 0)$, а значит, и для $w \in \Gamma_\rho$. Поэтому из леммы 1 следует, что для полинома Лагранжа $\hat{P}_{s,j}(y)$ степени s на чебышёвской сетке для $g(y)$ верно неравенство:

$$\begin{aligned} |g_j(y) - \hat{P}_{s,j}(y)| &\leq \frac{\rho + \rho^{-1}}{2} \frac{M}{(\rho + \rho^{-1}) - 1\rho^{s+1} - \rho^{-s-1}} = \frac{4\mu + 2}{2\mu} \frac{M}{\rho^{s+1} - \rho^{-s-1}} \leq \\ &\leq \frac{3M}{\rho^{s+1} - 1} \leq \frac{3M}{3M/\varepsilon + 1 - 1} = \varepsilon. \end{aligned} \tag{3}$$

Пусть $P_{s,j}(x) = \hat{P}_{s,j}\left(\frac{2}{h}(x - L - jh) - 1\right)$. Из неравенства (3) получаем, что $|f(x) - P_{s,j}(x)| \leq \varepsilon$ на отрезке $[L + jh, L + (j + 1)h]$ для $j = \mu, \mu + 1, \dots, b^k - \mu - 1$. Поэтому строки a_j^T с этими индексами приблизим строками:

$$\mathbf{P}_{s,j}^T = [P_{s,j}(L + h(j + \ell b^{k-d}))]_{\ell=0}^{b^{d-k}-1}.$$

Так как степень полиномов $P_{s,j}(x)$ не превосходит s , все строки $\mathbf{P}_{s,j}^T$ лежат в линейной оболочке векторов $p_0, \dots, p_s \in \mathbb{R}^{d-k}$, $p_t(\ell) = \ell^t$.

Строки a_j^T матрицы A_k с индексами $j \in \{0, \dots, \mu - 1\} \cup \{b^k - \mu, \dots, b^k - 1\}$ “приблизим” ими самими, т.е. матрицу B_k построим в виде

$$\left[a_0^T, \dots, a_{\mu-1}^T, \mathbf{P}_{s,\mu}^T, \dots, \mathbf{P}_{s,b^k-\mu-1}^T, a_{b^k-\mu}^T, \dots, a_{b^k-1}^T \right]^T.$$

Очевидно, $\text{rank } B_k \leq 2\mu + s + 1$ и $\|A_k - B_k\|_\infty \leq \varepsilon$.

Следствие 1. В условиях теоремы 1 для любого $\hat{\varepsilon} > 0$ выполнено

$$\mathcal{R}_{\hat{\varepsilon}}(T_{b,d,[L,R]}(f)) \leq \hat{s} + 3, \quad \text{где} \quad \hat{s} = \left\lceil \log_\rho(3Mb^{d/2}\sqrt{d-1}\hat{\varepsilon}^{-1} + 1) \right\rceil, \quad \rho = 3 + 2\sqrt{2}.$$

Доказательство. Положим в теореме 1 $\mu = 1$ (чтобы удовлетворить условию $\mu \in [1, b^k - 1]$ для всех $k = 1, \dots, d - 1$) и

$$\varepsilon = \frac{\hat{\varepsilon}}{b^{d/2}\sqrt{d-1}}$$

и рассмотрим приближения B_k к матрицам развертки A_k со свойствами $\text{rank } B_k \leq \hat{s} + 3$ и $\|A_k - B_k\|_\infty \leq \varepsilon$. Из последнего неравенства следует, что $\|A_k - B_k\|_F \leq \varepsilon b^{d/2}$.

Теорема 2.2 из [8] гарантирует существование тензора \hat{B} с ГТ-рангами r_k , приближающего тензор A в норме Фробениуса с точностью:

$$\sqrt{\sum_{k=1}^{d-1} \varepsilon_k^2},$$

где r_k суть ε_k -ранги матриц развертки A_k .

Положим $\varepsilon_k = \varepsilon b^{d/2}$, тогда $r_k \leq \text{rank } B_k \leq \hat{s} + 3$ и при этом

$$\|A - \hat{B}\|_F \leq \sqrt{d-1} \varepsilon b^{d/2} = \hat{\varepsilon}.$$

Теорема 2. Пусть $\hat{\Gamma}$ – образ эллипса Γ_ρ под действием линейного отображения

$$\xi(w) = \frac{R+L}{2} + \frac{R-L}{2} w,$$

а $f(z)$ – аналитическая внутри $\hat{\Gamma}$ функция, для которой $|f(z)| \leq M$ для $z \in \hat{\Gamma}$ и имеющая вещественнозначный след на $[L, R]$.

Тогда для произвольного $\varepsilon > 0$ и натуральных b и d существует тензор B такой, что

$$\mathfrak{R}(B) \leq \lfloor \log_p(3M/\varepsilon + 1) \rfloor, \quad \|T_{b,d,[L,R]}(f) - B\|_\infty \leq \varepsilon.$$

Доказательство. Будем обозначать через $(f \circ \xi)(x)$ композицию функций f и ξ , т.е. $(f \circ \xi)(x) = f(\xi(x))$. Заметим, что $(f \circ \xi)(z)$ аналитична в Γ_p как композиция аналитических функций f и ξ . Приближим $(f \circ \xi)(x)$ полиномом Лагранжа $\hat{P}_s(x)$ степени s на чебышёвской сетке на $[-1, 1]$, $(f \circ \xi)(x)$ – след аналитической в Γ_p функции $(f \circ \xi)(z)$. Аналогично доказательству теоремы 1 можно показать, что для параметра s из условия выполнено неравенство:

$$\begin{aligned} |(f \circ \xi)(x) - \hat{P}_s(x)| \leq \varepsilon \quad \forall x \in [-1, 1], &\Rightarrow |(f \circ \xi)(\xi^{-1}(y)) - \hat{P}_s(\xi^{-1}(y))| \leq \varepsilon \quad \forall y \in [L, R], \Rightarrow \\ &\Rightarrow |f(y) - (\hat{P}_s \circ \xi^{-1})(y)| \leq \varepsilon \quad \forall y \in [L, R], \end{aligned}$$

где $(\hat{P}_s \circ \xi^{-1})(y)$ есть полином степени не более s , поэтому согласно [3], [2] его тензоризация B допускает ТТ-разложение с ТТ-рангами, не превосходящими $s + 1$.

4. НИЖНИЕ ОЦЕНКИ

Рассмотрим пример функции $f(x)$, являющейся следом аналитической в \mathbb{C} функции. Для приближенных ТТ-рангов тензоризации этой функции будет доказана нижняя оценка (теорема 3).

Зафиксируем натуральные $b \geq 2$, $d \geq 2$ и $k \geq d/3$. Определим $f(z) = \sin(2\pi b^{2k} z^2)$, $z \in \mathbb{C}$, и обозначим через $f(x)$ ее (вещественнозначный) след на \mathbb{R} . Пусть $A = T_{b,d,[0,1]}(f)$.

Введем в рассмотрение функции:

$$f_s(\xi) = f\left(\left(s + \frac{\xi}{2\pi}\right)b^{-k}\right), \quad \xi \in [0, 2\pi], \quad s \in \{0, \dots, b^k - 1\}.$$

Строка матрицы развертки A_k с индексом s есть дискретизация функции $f_s(\xi)$ на равномерной сетке на $[0, 2\pi]$.

Для всех $s, t \in \{0, \dots, b^k - 1\}$ определим интегралы

$$d_{s,t} = \frac{1}{\pi} \int_0^{2\pi} f_s(\xi) f_t(\xi) d\xi.$$

Для тех же s, t введем величины (здесь $a_{s,i}$ – элементы матрицы A_k):

$$\hat{d}_{s,t} = \frac{2\pi}{b^{d-k}} \frac{1}{\pi} \sum_{i=0}^{b^{d-k}-1} a_{s,i} a_{t,i}. \tag{4}$$

Матрицу из элементов $\hat{d}_{s,t}$ обозначим через \hat{D} .

Лемма 2. Для всех $s, t \in \{0, \dots, b^k - 1\}$ верно неравенство

$$|d_{s,t} - \delta_{s,t}| \leq \frac{2}{(s+t)^2},$$

где $\delta_{s,t}$ – символ Кронекера.

Доказательство. Преобразуем выражения для $f_s(\xi)$:

$$f_s(\xi) = \sin\left(2\pi b^{2k} \left(s^2 + \frac{s\xi}{\pi} + \frac{\xi^2}{4\pi^2}\right) b^{-2k}\right) = \sin\left(2s\xi + \frac{1}{2\pi} \xi^2\right).$$

Отсюда получим

$$d_{s,t} = \frac{1}{\pi} \int_0^{2\pi} f_s(\xi) f_t(\xi) d\xi = \frac{1}{2\pi} \int_0^{2\pi} \cos(2(s-t)\xi) - \cos\left(2(s+t)\xi + \frac{1}{\pi}\xi^2\right) d\xi =$$

$$= \frac{1}{2\pi} \begin{cases} 2\pi - \int_0^{2\pi} \cos\left(4s\xi + \frac{1}{\pi}\xi^2\right) d\xi, & \text{если } s = t, \\ -\int_0^{2\pi} \cos\left(2(s+t)\xi + \frac{1}{\pi}\xi^2\right) d\xi, & \text{если } s \neq t. \end{cases} \quad (5)$$

Пусть $\beta = 1/\pi$, $p = 2(s+t)$. Исследуем интеграл $\int_0^{2\pi} e^{ip\xi + i\beta\xi^2} d\xi$:

$$\int_0^{2\pi} e^{ip\xi + i\beta\xi^2} d\xi = \int_0^{2\pi} \frac{1}{ip} e^{i\beta\xi^2} de^{ip\xi} = \frac{1}{ip} \left(e^{i\beta\xi^2} e^{ip\xi} \Big|_0^{2\pi} - \int_0^{2\pi} i\beta 2\xi e^{ip\xi} e^{i\beta\xi^2} d\xi \right).$$

Первый член в скобках равен 0, так как $e^{i\beta(2\pi)^2} = e^{i4\pi} = 1 = e^{ip2\pi}$. Повторяя интегрирование по частям, получаем

$$\int_0^{2\pi} e^{ip\xi + i\beta\xi^2} d\xi = \frac{-2i\beta}{ip} \int_0^{2\pi} \xi e^{ip\xi} e^{i\beta\xi^2} d\xi = \frac{-2\beta}{p} \frac{1}{ip} \int_0^{2\pi} \xi e^{i\beta\xi^2} de^{ip\xi} =$$

$$= \frac{-2\beta}{ip^2} \left(\xi e^{i\beta\xi^2} e^{ip\xi} \Big|_0^{2\pi} - \int_0^{2\pi} e^{ip\xi} e^{i\beta\xi^2} (1 + i\beta 2\xi^2) d\xi \right) = \frac{-2\beta}{ip^2} \left(2\pi - \int_0^{2\pi} e^{ip\xi} e^{i\beta\xi^2} (1 + i\beta 2\xi^2) d\xi \right).$$

Взяв вещественную часть от обеих частей полученного равенства, имеем

$$\int_0^{2\pi} \cos(p\xi + \beta\xi^2) d\xi = -\frac{2\beta}{p^2} \operatorname{Im} \int_0^{2\pi} e^{ip\xi} e^{i\beta\xi^2} (1 + i\beta 2\xi^2) d\xi.$$

Поэтому верно

$$\left| \int_0^{2\pi} \cos(p\xi + \beta\xi^2) d\xi \right| \leq \frac{2\beta}{p^2} \int_0^{2\pi} (1 + 2\beta\xi^2) d\xi = \frac{4\beta\pi}{p^2} + \frac{4\beta^2}{3p^2} 8\pi^3 = \frac{1}{2} \left(4 + \frac{32\pi}{3} \right).$$

Отсюда и из (5) следует, что

$$|d_{s,t} - \delta_{s,t}| \leq \frac{1}{8\pi(s+t)^2} \left(4 + \frac{32\pi}{3} \right) \leq \frac{2}{(s+t)^2}.$$

Лемма 3. Для всех $s, t \in \{0, \dots, b^k - 1\}$ верно неравенство

$$|\hat{d}_{s,t} - d_{s,t}| \leq 2\pi b^{k-d} (2s + 2t + 4).$$

Доказательство. Напомним, что формулой левых прямоугольников для вычисления интеграла $\int_{x_i}^{x_{i+1}} \varphi(x) dx$ называется выражение $(x_{i+1} - x_i)\varphi(x_i)$, а составной формулой для сетки $\{x_0, \dots, x_n\}$ с шагом h называется сумма $\sum_{i=0}^{n-1} \varphi(x_i)h$. Простая формула имеет погрешность:

$$\left| \int_{x_i}^{x_{i+1}} \varphi(x) dx - \int_{x_i}^{x_{i+1}} \varphi(x_i) dx \right| \leq \int_{x_i}^{x_{i+1}} |\varphi'(\xi(x))| (x - x_i) dx \leq \|\varphi'\|_{C[x_i, x_{i+1}]} \frac{1}{2} (x_{i+1} - x_i)^2,$$

где $\xi(x) \in [x_i, x_{i+1}]$. При применении составной формулы на отрезке $[L, R]$ с шагом h погрешность не превосходит $\|\varphi'\|_{C[L, R]} (R - L)h/2$.

Ясно видно, что выражение (4) для $\hat{d}_{s,t}$ есть в точности составная формула левых прямоугольников для вычисления интеграла $d_{s,t}$ на отрезке $[0, 2\pi]$. Погрешность аппроксимации интеграла оценивается (подразумевается норма $\|\cdot\|_{C[0,2\pi]}$) в виде

$$\begin{aligned} |\hat{d}_{s,t} - d_{s,t}| &\leq \left\| \frac{d}{d\xi} \left(\frac{1}{\pi} f_s \cdot f_t \right) \right\| 2\pi \frac{2\pi}{2b^{d-k}} = 2\pi b^{k-d} \|f'_s f_t + f_s f'_t\| \leq \\ &\leq 2\pi b^{k-d} (\|f'_s\| + \|f'_t\|) \leq 2\pi b^{k-d} \left(2s + \frac{2\pi}{\pi} + 2t + \frac{2\pi}{\pi} \right) = 2\pi b^{k-d} (2s + 2t + 4). \end{aligned}$$

Лемма 4. Пусть F – главная подматрица матрицы \hat{D} , находящаяся на пересечении строк и столбцов с индексами $s, t \in [l, u]$, где

$$u = \left\lfloor \frac{b^{\frac{d-k}{2}}}{8} \right\rfloor, \quad \ell = \left\lceil \frac{b^{\frac{d-k}{2}}}{16} \right\rceil,$$

причем $b^{d-k} \geq 2^{10}$. Тогда младшее сингулярное число $\sigma_{\min}(F) > 1/2$.

Доказательство. Обратим внимание, что заявленное в начале раздела условие $k \geq d/3$ гарантирует, что $\ell, u < b^k$, поэтому выбирать подматрицу с указанными индексами в матрице $\hat{D} \in \mathbb{R}^{b^k \times b^k}$ корректно.

Пусть $G = F - I$. Заметим, что

$$\begin{aligned} \sigma_{\min}(F) = \sigma_{\min}(I + G) &= \|(I + G)^{-1}\|_2^{-1} = \|I - G + G^2 - G^3 + \dots\|_2^{-1} \geq \\ &\geq (\|I\|_2 + \|G\|_2 + \|G\|_2^2 + \dots)^{-1} = \left(\frac{1}{1 - \|G\|} \right)^{-1} = 1 - \|G\|_2. \end{aligned} \tag{6}$$

Из лемм 2 и 3 следует, что все элементы матрицы $\hat{D} - I$, а значит, и $F - I$, по модулю не превосходят

$$\frac{2}{(s+t)^2} + 2\pi b^{k-d} (2s + 2t + 4).$$

Поэтому имеем

$$\begin{aligned} \|G\|_2 = \|F - I\|_2 &\leq \|F - I\|_F \leq (u - \ell) \|F - I\|_{\infty} \leq \\ &\leq (u - \ell) \left(\frac{2}{(s+t)^2} + 2\pi b^{k-d} (2s + 2t + 4) \right) \leq \frac{b^{\frac{d-k}{2}}}{16} \left(\frac{2}{4\ell^2} + \frac{2\pi \cdot 4u}{b^{d-k}} \right) \leq \frac{256}{32b^{\frac{d-k}{2}}} + \frac{8\pi}{128}. \end{aligned}$$

По условию $b^{d-k} \geq 2^{10}$, тогда получаем

$$\|G\|_2 \leq \frac{1}{4} + \frac{\pi}{16} < \frac{1}{2}.$$

Отсюда и из неравенства (6) следует утверждение леммы.

Теорема 3. Пусть $b \geq 2$, $d \geq 2$ и $k \geq d/3$ – натуральные числа, удовлетворяющие условию $b^{d-k} \geq 2^{10}$, $a f(x) = \sin(2\pi b^{2k} x^2)$, $x \in [0, 1]$. Тогда верно неравенство

$$\mathfrak{R}_{\varepsilon}(T_{b,d,[0,1]}(f)) \geq \frac{1}{16} b^{\frac{d-k}{2}} - 2, \quad \text{где} \quad \varepsilon = \frac{1}{2} b^{\frac{d-k}{2}}.$$

Доказательство. Исследуем сингулярные числа $\sigma_s(A_k)$ матрицы A_k . Для этого воспользуемся известным фактом, что

$$\sigma_s(A_k) = \sqrt{\lambda_s(A_k^T A_k)} = \sqrt{\sigma_s(A_k^T A_k)}. \tag{7}$$

Из определения (4) матрицы \hat{D} следует, что $A_k^T A_k = b^{d-k} \hat{D}/2$. Далее, для главной подматрицы F матрицы \hat{D} из леммы можно применить теорему о чередовании сингулярных чисел симметричной матрицы [9] и получить, что $\sigma_{u-\ell}(\hat{D}) \geq 1/2$. Отсюда и из равенства (7) вытекает

$$\sigma_{u-\ell}(A_k) \geq \frac{1}{\sqrt{2}} b^{\frac{d-k}{2}} \sqrt{\sigma_{u-\ell}(\hat{D})} > \frac{1}{2} b^{\frac{d-k}{2}}.$$

Рассмотрим теперь произвольный тензор B со свойством $\|A - B\|_F \leq \frac{1}{2} b^{(d-k)/2}$. Аналогичное неравенство верно и для матриц развертки A_k и B_k . По теореме Эккарта–Юнга [10] наилучшее приближение в норме Фробениуса ранга $u - \ell - 1$ имеет погрешность как минимум

$$\sqrt{\sigma_{u-\ell}^2(A_k) + \dots + \sigma_N^2(A_k)} \geq \sigma_{u-\ell}(A_k) > \frac{1}{2} b^{\frac{d-k}{2}},$$

где $N = \min\{b^k, b^{d-k}\}$ – меньший из размеров матрицы A_k . Поэтому верно

$$\text{rank } B_k \geq u - \ell \geq \left(\frac{1}{8} b^{\frac{d-k}{2}} - 1\right) - \left(\frac{1}{16} b^{\frac{d-k}{2}} - 1\right) = \frac{1}{16} b^{\frac{d-k}{2}} - 2.$$

Рассмотрим произвольное ТТ-разложение тензора B вида (1). Для k -й матрицы развертки можно написать

$$B_k(i_1, \dots, i_k; i_{k+1}, \dots, i_d) = \sum_{\alpha_k=1}^{r_k} H'(i_1, \dots, i_k, \alpha_k) H''(\alpha_k, i_{k+1}, \dots, i_d),$$

откуда следует, что $r_k \geq \text{rank } B_k \geq \frac{1}{16} b^{\frac{d-k}{2}} - 2$, а в силу произвольности B и

$$\mathcal{R}_\varepsilon(A) \geq \frac{1}{16} b^{\frac{d-k}{2}} - 2.$$

Чтобы наглядно связать верхнюю оценку из данного раздела и нижние оценки из предыдущего, имеет смысл зафиксировать некоторые величины и перейти к асимптотической нотации. Именно, зафиксируем натуральное $b \geq 2$, произвольно малое $\varepsilon > 0$ и отрезок $[L, R]$ на вещественной прямой. Определим функцию

$$\mathcal{R}_{b,\varepsilon,[L,R]}(M, d) = \max_{|f(z)| \leq M \text{ на } U} \mathcal{R}_\varepsilon(T_{b,d,[L,R]}(f)).$$

Здесь максимум берется по всем аналитическим в круге U с диаметром $[L, R]$ функциям $f(z)$, имеющим вещественный след на $[L, R]$.

Следствие 1 гарантирует оценку $\mathcal{R}_{b,\varepsilon,[L,R]}(M, d) = O(\ln M + d)$ при $M, d \rightarrow \infty$. С другой стороны, для функции $f(z)$ из данного раздела несложно оценить максимум модуля на единичном круге:

$$|f(z)| = \left| \sin(2\pi b^{2k} z^2) \right| = \left| \frac{1}{2i} \left(e^{2\pi i b^{2k} z^2} - e^{-2\pi i b^{2k} z^2} \right) \right| \leq e^{2\pi b^{2k} |z|^2} \leq e^{2\pi b^{2k}}. \tag{8}$$

Пусть $k = \lceil d/3 \rceil$, тогда по теореме 3 максимальный ТТ-ранг есть как минимум:

$$\frac{1}{16} b^{\frac{d-k}{2}} - 2 = \Omega(b^k) = \Omega(\sqrt{\ln M(d)}) \quad \text{при } d \rightarrow \infty.$$

В последнем равенстве через $M(d)$ обозначено число $e^{2\pi b^{2d/3}}$. Получается, что

$$\mathcal{R}_{b,\varepsilon,[L,R]}(M(d), d) = \Omega(\sqrt{\ln M(d)}) \quad \text{при } d \rightarrow \infty.$$

5. ПРИМЕНЕНИЕ К ПОЛИНОМАМ

В работе [5] рассматривались ТТ-ранги приближений к тензоризациям полиномов. В данном разделе мы применим результаты предыдущих разделов к полиномам, существенно улучшив

верхние оценки из указанной работы. Также мы получим нижние оценки ТТ-рангов ε -приближений для этого класса функций.

Утверждение 1. Пусть $p(x) = p_0 + p_1x + \dots + p_nx^n$ — полином с вещественными коэффициентами. Пусть $A = T_{b,d,[0,1]}(f)$ и $M = \sum_{i=0}^n |p_i|$. Зафиксируем произвольное $\varepsilon > 0$.

Тогда для каждой матрицы развертки A_k тензора A , $k = 1, \dots, d - 1$, и натурального $\mu \in [1, b^k - 1]$ существует матрица B_k такая, что

$$\text{rank } B_k \leq 2\mu + s + 1 \quad \text{и} \quad \|A_k - B_k\|_\infty \leq \varepsilon,$$

где

$$s = \lfloor \log_\rho(3M/\varepsilon + 1) \rfloor, \quad \rho = 2\mu + 1 + 2\sqrt{\mu^2 + \mu}.$$

Доказательство. Достаточно заметить, что на единичном круге на комплексной плоскости (а тем более, в круге с диаметром $[0, 1]$) аналитическое продолжение $p(z)$ полинома $p(x)$ ограничено по модулю суммой $\sum_i |p_i|$, и применить теорему 1.

Замечание. Если в духе работы [5] ограничить коэффициенты p_i по модулю фиксированной константой и поинтересоваться асимптотическим поведением ε -рангов матриц развертки при стремлении n к бесконечности, то доказанное утверждение даст оценку вида $O(\ln n)$ в отличие от оценки $O(\sqrt[3]{n})$, полученной в [5].

Утверждение 2. Пусть $b \geq 2$, $d \geq 2$ и $k \geq d/3$ — натуральные числа, удовлетворяющие условию $b^{d-k} \geq 2^{10}$.

Тогда для любого $n \geq \lceil \log_2(1 + 14b^{k/2}e^{10b^{2k}}) \rceil$ существует полином $P_n(x) = p_0 + \dots + p_nx^n$ такой, что верно следующее:

1. $\sum_{i=0}^n |p_i| \leq 2^{2n}(n + 1)^2$;
2. $\mathcal{R}_\varepsilon(T_{b,d,[0,1]}(P_n)) \geq \frac{1}{16}b^{\frac{d-k}{2}} - 2$, где $\varepsilon = \frac{1}{4}b^{\frac{d-k}{2}}$.

Доказательство. Возьмем функцию $f(x)$ из теоремы 3. Обозначим через $P_n(x)$ полином Лагранжа степени n , интерполирующий $f(x)$ на чебышёвской сетке с $n + 1$ узлом на $[-1, 1]$.

Сначала оценим коэффициенты $p(x)$. Полином Лагранжа можно записать в следующем виде [9]:

$$P_n(x) = \sum_{j=0}^n \frac{f(x_j)\omega(x)}{(x - x_j)\omega'(x_j)}, \quad \omega(x) = (x - x_0)\cdots(x - x_n),$$

где x_j суть корни полинома Чебышёва степени $n + 1$, т.е.

$$x_j = \cos\left(\frac{\pi}{2(n+1)} + \frac{\pi}{n+1}j\right), \quad \omega(x) = 2^{-n} \cos((n+1)\arccos x).$$

Оценим снизу модуль $\omega'(x_j)$:

$$\omega'(x_j) = 2^{-n}(n+1) \frac{1}{\sqrt{1-x_j^2}} \sin((n+1)\arccos x_j) = \frac{2^{-n}(n+1)}{\sin\left(\frac{\pi}{2(n+1)} + \frac{\pi}{n+1}j\right)} \sin\left(\frac{\pi}{2} + \pi j\right).$$

Поэтому $|\omega'(x_j)| \geq 2^{-n}(n+1)$. Коэффициент при x^m многочлена $\omega(x)/(x - x_j)$ есть, очевидно,

$$\sum_{\substack{0 \leq j_1 < \dots < j_{n-m} \leq n \\ j_i \neq j}} (-1)^{n-m} x_{j_1} \cdots x_{j_{n-m}},$$

т.е. по модулю не превосходит C_n^{n-m} . Итого получаем

$$\sum_{m=0}^n |p_m| \leq \sum_{j=0}^n \left| \frac{f(x_j)}{\omega'(x_j)} \right| \sum_{m=0}^n C_n^{n-m} \leq (n+1)2^n (n+1)2^n = 2^{2n}(n+1)^2.$$

Применим лемму 1 для оценки отклонения $f(x)$ от $P_n(x)$ на $[-1, 1]$:

$$|f(x) - P_n(x)| \leq \frac{M}{\rho^{n+1} - \rho^{-n-1}} \frac{\rho + \rho^{-1}}{\frac{1}{2}(\rho + \rho^{-1} - 1)}, \quad M = \max_{z \in \Gamma_\rho} |f(z)|.$$

Положим $\rho = 2$ и оценим величину M , исходя из рассуждения, аналогичного (8):

$$M \leq e^{2\pi \left(\frac{\rho + \rho^{-1}}{2}\right)^2 b^{2k}} \leq e^{10b^{2k}}.$$

Таким образом, с учетом неравенства для n (из условия утверждения) имеем для всех $x \in [-1, 1]$:

$$|f(x) - P_n(x)| \leq \frac{e^{10b^{2k}}}{14b^{k/2} e^{10b^{2k}}} \frac{2.5}{0.75} \leq \frac{1}{4} b^{-k/2}.$$

Это означает, что

$$\|T_{b,d,[0,1]}(f) - T_{b,d,[0,1]}(P_n)\|_F \leq \frac{1}{4} b^{\frac{d-k}{2}} = \varepsilon.$$

Теперь возьмем произвольный тензор B , приближающий $T_{b,d,[0,1]}(P_n)$ с ошибкой не более ε . Заметим, что по неравенству треугольника для нормы он дает приближение $T_{b,d,[0,1]}(f)$ с ошибкой не более 2ε . Из теоремы 3 сразу получаем неравенство

$$\mathcal{R}(B) \geq \frac{1}{16} b^{\frac{d-k}{2}} - 2,$$

а в силу произвольности B и свойство из условия.

Автор выражает благодарность Виктории Владимировне Высоцкой за помощь в оформлении статьи.

СПИСОК ЛИТЕРАТУРЫ

1. *Oseledets I.* Constructive representation of functions in low-rank tensor formats // *Constructive Approximat.* 2013. V. 37. № 1. P. 1–18.
2. *Grasedyck L.* Polynomial approximation in hierarchical Tucker format by vector–tensorization. Marburg: Philipps-Universität, 2010.
3. *Khoromskij B.* $O(d \log N)$ -Quantics approximation of N -d tensors in high-dimensional numerical modeling // *Constructive Approximat.* 2011. V. 34. № 2. P. 257–280.
4. *Tyrtyshnikov E.E.* Tensor approximations of matrices generated by asymptotically smooth functions // *Mat. Sb.* 2003. V. 194. I. 6. P. 941–954.
5. *Vysotsky L.* On tensor-train ranks of tensorized polynomials // *Large-Scale Scientific Computing. LSSC 2019. Lect. Notes in Comp. Sc.* 2020. V. 11958. P. 189–196.
6. *Oseledets I.* Approximation of $2^d \times 2^d$ matrices using tensor decomposition // *SIAM Journal on Matrix Anal. and Appl.* 2009. V. 31. № 4. P. 2130–2145.
7. *Oseledets I.* Tensor-train decomposition // *SIAM Journal on Scientific Comp.* 2011. V. 33. № 5. P. 2295–2317.
8. *Oseledets I., Tyrtyshnikov E.* TT-cross approximation for multidimensional arrays // *Linear Algebra and Its Appl.* 2010. V. 432. № 1. P. 70–88.
9. *Тыртышников Е.Е.* Методы численного анализа. М.: Академия, 2007.
10. *Eckart C., Young G.* The approximation of one matrix by another of lower rank // *Psychometrika.* 1936. V. 1. № 3. P. 211–218.