О ТТ-рангах приближённых тензоризаций некоторых гладких функций

Высоцкий Л.И.[1,2]

[1] *Институт вычислительной математики им. Г.И. Марчука Российской академии наук, ул. Губкина, 8, Москва, Россия, 119333*

[2] *Факультет вычислительной математики и кибернетики, Московский государственный университет имени М.В. Ломоносова, Ленинские горы, 1, стр.52, ГСП-1, Москва, Россия, 119991*

Аннотация

В данной работе 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 их 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, существенно улучшая известную верхнюю границу.

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

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, то есть 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 [6] как тензор (многомерный массив) размеров $n_1 \times \dots \times n_d$ с элементами

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

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):

$$B(i_1, \dots, i_d) = \sum_{\alpha_1, \dots, \alpha_{d-1}} H_1(i_1, \alpha_1) \dots H_k(\alpha_{k-1}, i_k, \alpha_k) \dots H_d(\alpha_{d-1}, i_d), \tag{1.1}$$

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 (1.1), причём для единообразия определения тензоров H_k считается, что $r_0 = r_d = 1$.

Величины ТТ-рангов играют ключевую роль для применимости ТТ-разложения, ведь требуемая для хранения тензоров H_k память и сложность распространённых операций над тензорами в этом формате растёт 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 её ТТ-ранги ограничены $n + 1$. В [2] было показано, что

E-mail: vysotskylev@yandex.ru (Высоцкий Л.И.)

ТТ-ранги приближений тензоризаций т.н. *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 $-\log_2(\varepsilon)$. 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

Матрицами развёртки тензора $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$ называются матрицы

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

где группы индексов до и после точки с запятой образуют т.н. *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 $\|A\|_\infty = \max |A(i_1, \dots, i_d)|$ и $\|A\|_F = \sqrt{\sum (A(i_1, \dots, i_d))^2}$.

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

Для ТТ-разложения вида (1.1) 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 $\mathcal{R}(B)$ по следующей формуле:

$$\mathcal{R}(B) := \min_{\substack{(H_1, \dots, H_d) - \\ \text{ТТ-разложение } B}} \mathcal{R}(H_1, \dots, H_d).$$

Для оценки ТТ-рангов наилучших приближений нам понадобится функция

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

Эллипс на комплексной плоскости с центром в нуле и осями, параллельными осям координат, чьи полуоси равны $0.5(\rho + \rho^{-1})$ и $0.5(\rho - \rho^{-1})$ для некоторого $\rho > 1$, будем называть эллипсом Бернштейна и обозначать Γ_ρ. Также эллипс Бернштейна можно рассматривать (см. [8]) как образ окружности $\{\rho e^{i\varphi} : \varphi \in [0, 2\pi)\}$ под действием отображения Жуковского $w \to 0.5(w + w^{-1})$. Такой же эллипс, но с центром в точке $z \in \mathbb{C}$, будем обозначать $\Gamma_\rho(z)$.

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 [8], что числа x_j являются корнями *полинома Чебышёва* степени n, то есть $T_n(x) = \cos(\arccos(nx))$.

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

3. Верхние оценки

Лемма 1. *Пусть $f(x)$ — след аналитической внутри эллипса Бернштейна Γ_ρ, $\rho > 1$, функции $f(z)$, причём $|f(z)| \le M$ 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 $x \in [-1, 1]$*

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

Доказательство. Зафиксируем произвольное $x \in [-1,1] \setminus \{x_0, \ldots, x_n\}$. 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 $\Pi_j(x-x_j)$ (то есть рассуждая аналогично [8, Теорема 13.6]), придём к представлению

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

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 [8, Теорема 13.6] для $z \in \Gamma_\rho$ неравенство $|T_{n+1}(z)| \geq 0.5(\rho^{n+1} - \rho^{-n-1})$, а также очевидное для $x \in [-1,1]$ неравенство $|T_{n+1}(x)| \leq 1$, получаем, что при $x \in [-1,1]$

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

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 $w \to 0.5(w + w^{-1})$:

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

Далее вычислим $\min|z-x|^2$ для $z \in \Gamma_\rho$ и $x \in [-1,1]$. Обозначим за r_1 и r_2 большую и меньшую полуоси эллипса Γ_ρ соответственно. Тогда точки этого эллипса параметризуются углом φ: $z = r_1\cos\varphi + ir_2\sin\varphi$. Преобразуем минимизируемое выражение:

$$|z-x|^2 = (r_1\cos\varphi - x)^2 + (r_2\sin\varphi)^2 = (r_1^2 - r_2^2)\cos^2\varphi - 2r_1x\cos\varphi + x^2 + r_2^2. \tag{3.1}$$

В силу симметрии достаточно рассмотреть $x \geq 0$. При фиксированном x выражение (3.1) — это квадратный трёхчлен относительно $\cos\varphi$, причём вершина соответствующей параболы имеет абсциссу $r_1x/(r_1^2 - r_2^2) = r_1x$. При $r_1x \leq 1$ минимум по φ достигается при $\cos\varphi = r_1x$ и равен r_2^2. При $r_1x > 1$ минимум достигается при $\cos\varphi = 1$ и равен $(r_1 - x)^2$, причём это выражение достигает минимума по x при $x = 1$. Так как $r_1 - 1 < r_2$, то и $\min|z-x| = r_1 - 1 = 0.5(\rho+\rho^{-1}) - 1$. Итого,

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

□

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

Фиксируем произвольное $\varepsilon > 0$. Тогда для каждой матрицы развёртки $A_k \in \mathbb{R}^{b^k \times b^{d-k}}$ 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

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

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

Доказательство. Обозначим 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

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

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

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 $j = \mu, \mu+1, \ldots, b^k - \mu - 1$ круг $U(\beta(\rho+\rho^{-1})/2, \alpha)$ лежит внутри круга с диаметром $[L, R]$, то есть

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

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

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

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 $\widehat{P}_{s,j}(y)$ степени s на Чебышёвской сетке для $g(y)$ верно неравенство

$$\left|g_j(y) - \widehat{P}_{s,j}(y)\right| \leq \frac{\rho+\rho^{-1}}{\frac{1}{2}(\rho+\rho^{-1}) - 1}\frac{M}{\rho^{s+1} - \rho^{-s-1}} = \frac{4\mu+2}{2\mu}\frac{M}{\rho^{s+1}-\rho^{-s-1}} \leq \frac{3M}{\rho^{s+1}-1} \leq \frac{3M}{3M/\varepsilon + 1 - 1} = \varepsilon. \tag{3.2}$$

Обозначим $P_{s,j}(x) := \widehat{P}_{s,j}\left(\frac{2}{h}(x - L - jh) - 1\right)$. Из неравенства (3.2) получаем, что $|f(x) - P_{s,j}(x)| \leq \varepsilon$ на отрезке $[L+jh, L+(j+1)h]$ для $j = \mu, \mu+1, \ldots, b^k - \mu - 1$. Поэтому строки $a_j^\top$ с этими индексами приблизим строками

$$\vec{P}_{s,j}^\top := \left[P_{s,j}\left(L + h\left(j + \ell b^{k-d}\right)\right)\right]_{\ell=0}^{b^{d-k}-1}.$$

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

Строки $a_j^\top$ матрицы A_k с индексами $j \in \{0, \ldots, \mu-1\} \cup \{b^k - \mu, \ldots, b^k - 1\}$ "приблизим" ими самими, то есть матрицу B_k построим в виде

$$\left[a_0^\top, \ldots, a_{\mu-1}^\top, \vec{P}_{s,\mu}^\top, \ldots, \vec{P}_{s,b^k-\mu-1}^\top, a_{b^k-\mu}^\top, \ldots, a_{b^k-1}^\top\right]^\top.$$

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

Следствие 1. *В условиях Теоремы 1 для любого* $\widehat{\varepsilon} > 0$ *выполнено*

$$\mathcal{R}_{\widehat{\varepsilon}}(T_{b,d,[L,R]}(f)) \leq \widehat{s} + 3, \text{ где}$$

$$\widehat{s} = \left\lfloor \log_\rho\left(3Mb^{\frac{d}{2}}\sqrt{d-1}\,\widehat{\varepsilon}^{-1} + 1\right)\right\rfloor, \quad \rho = 3 + 2\sqrt{2},$$

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

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

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

Теорема 2.2 из [9] гарантирует существование тензора $\widehat{B}$ с ТТ-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 развёртки A_k. Положим $\varepsilon_k := \varepsilon b^{d/2}$, тогда $r_k \le \operatorname{rank} B_k \le \widehat{s}+3$ и при этом

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

□

Теорема 2. *Пусть $\widehat{\Gamma}$ — образ эллипса Γ_ρ под действием линейного отображения $\xi(w) := \frac{R+L}{2} + \frac{R-L}{2}w$, а $f(z)$ — аналитическая внутри $\widehat{\Gamma}$ функция, для которой $|f(z)| \le M$ для $z \in \widehat{\Gamma}$ и имеющая вещественнозначный след на $[L, R]$.*

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

$$\mathcal{R}(B) \le \lfloor \log_\rho (3M/\varepsilon + 1) \rfloor, \ \ \|T_{b,d,[L,R]}(f) - B\|_\infty \le \varepsilon.$$

Доказательство. Приблизим $(f\circ\xi)(x)$ полиномом Лагранжа $\widehat{P}_s(x)$ степени s на Чебышёвской сетке на $[-1,1]$. $(f\circ\xi)(x)$ — след аналитической в Γ_ρ функции $(f\circ\xi)(z)$. Аналогично доказательству Теоремы 1 можно показать, что для параметра s из условия выполнено неравенство

$$\begin{aligned} \left|(f\circ\xi)(x) - \widehat{P}_s(x)\right| &\le \varepsilon \ \ \forall x \in [-1,1] \Rightarrow \\ \left|(f\circ\xi)\left(\xi^{-1}(y)\right) - \widehat{P}_s\left(\xi^{-1}(y)\right)\right| &\le \varepsilon \ \ \forall y \in [L,R] \Rightarrow \\ \left|f(y) - \left(\widehat{P}_s \circ \xi^{-1}\right)(y)\right| &\le \varepsilon \ \ \forall y \in [L,R]. \end{aligned}$$

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

4. Нижние оценки

В данном разделе будет рассмотрен пример функции $f(x)$, являющейся следом аналитической в $\mathbb{C}$ функции. Для приближённых ТТ-рангов тензоризации этой функции будет доказана нижняя оценка (Теорема 3).

Зафиксируем натуральные $b \ge 2$, $d \ge 2$ и $k \ge d/3$. Определим $f(z) := \sin(2\pi b^{2k} z^2)$, $z \in \mathbb{C}$, и обозначим $f(x)$ её (вещественнозначный) след на $\mathbb{R}$. Обозначим $A := T_{b,d,[0,1]}(f)$. Введём в рассмотрение функции

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

Строка матрицы развёртки 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 введём величины (здесь $a_{s,i}$ — элементы матрицы A_k)

$$\widehat{d}_{s,t} := \frac{2\pi}{b^{d-k}} \frac{1}{\pi} \sum_{i=0}^{b^{d-k}-1} a_{s,i} a_{t,i}. \tag{4.1}$$

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

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

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

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

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

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

Поэтому

$$\begin{aligned} d_{s,t}&=\frac{1}{\pi}\int_0^{2\pi}f_s(\xi)\cdot f_t(\xi)d\xi=\frac{1}{2\pi}\int_0^{2\pi}\cos\left(2(s-t)\xi\right)-\cos\left(2(s+t)\xi+\frac{1}{\pi}\xi^2\right)d\xi=\\ &=\frac{1}{2\pi}\cdot\begin{cases}2\pi-\int_0^{2\pi}\cos\left(4s\xi+\frac{1}{\pi}\xi^2\right)d\xi, & \text{если } s=t;\\ \quad-\int_0^{2\pi}\cos\left(2(s+t)\xi+\frac{1}{\pi}\xi^2\right)d\xi, & \text{если } s\ne t.\end{cases}\end{aligned}\tag{4.2}$$

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

$$\int_0^{2\pi}e^{ip\xi+i\beta\xi^2}d\xi=\int_0^{2\pi}\frac{1}{ip}e^{i\beta\xi^2}de^{ip\xi}=\frac{1}{ip}\left(e^{i\beta\xi^2}e^{ip\xi}\Big|_0^{2\pi}-\int_0^{2\pi}i\beta2\xi e^{ip\xi}e^{i\beta\xi^2}d\xi\right).$$

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, получим

$$\begin{aligned}\int_0^{2\pi}e^{ip\xi+i\beta\xi^2}d\xi&=\frac{-2i\beta}{ip}\int_0^{2\pi}\xi e^{ip\xi}e^{i\beta\xi^2}d\xi=\frac{-2\beta}{p}\cdot\frac{1}{ip}\int_0^{2\pi}\xi e^{i\beta\xi^2}de^{ip\xi}=\\ &=\frac{-2\beta}{ip^2}\left(\xi e^{i\beta\xi^2}e^{ip\xi}\Big|_0^{2\pi}-\int_0^{2\pi}e^{ip\xi}e^{i\beta\xi^2}\left(1+i\beta2\xi^2\right)d\xi\right)=\\ &=\frac{-2\beta}{ip^2}\left(2\pi-\int_0^{2\pi}e^{ip\xi}e^{i\beta\xi^2}\left(1+i\beta2\xi^2\right)d\xi\right).\end{aligned}$$

Взяв вещественную часть от обеих частей полученного равенства, получим

$$\int_0^{2\pi}\cos(p\xi+\beta\xi^2)d\xi=-\frac{2\beta}{p^2}\mathrm{Im}\int_0^{2\pi}e^{ip\xi}e^{i\beta\xi^2}\left(1+i\beta2\xi^2\right)d\xi.$$

Поэтому

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

Отсюда и из (4.2) следует, что

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

□

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

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

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

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

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 (4.1) для $\widehat{d}_{s,t}$ есть в точности составная формула левых прямоугольников для вычисления интеграла $d_{s,t}$ на отрезке $[0,2\pi]$. Погрешность аппроксимации интеграла оценивается так (подразумевается норма $\|\cdot\|_{C[0,2\pi]}$):

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

□

Лемма 4. *Пусть F — главная подматрица матрицы $\widehat{D}$, находящаяся на пересечении строк и столбцов с индексами $s,t\in[\ell,u)$, где $u := \left\lfloor b^{\frac{d-k}{2}}/8\right\rfloor, \ell := \left\lceil b^{\frac{d-k}{2}}/16\right\rceil$, причём $b^{d-k}\ge 2^{10}$.*

Тогда младшее сингулярное число $\sigma_{\min}(F) > 1/2$.

Доказательство. Обратим внимание, что заявленное в начала раздела условие $k\ge d/3$ гарантирует, что $\ell,u<b^k$, поэтому выбирать подматрицу с указанными индексами в матрице $\widehat{D}\in\mathbb{R}^{b^k\times b^k}$ — корректно. Обозначим $G=F-I$ и заметим, что

$$\begin{aligned}\sigma_{\min}(F) &= \sigma_{\min}(I+G) = \left\|(I+G)^{-1}\right\|_2^{-1} = \left\|I-G+G^2-G^3+\dots\right\|_2^{-1} \ge \\ &\ge \left(\|I\|_2+\|G\|_2+\|G\|_2^2+\dots\right)^{-1} = \left(\frac{1}{1-\|G\|}\right)^{-1} = 1-\|G\|_2.\end{aligned}\tag{4.3}$$

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

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

Поэтому

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

По условию $b^{d-k}\ge 2^{10}$, поэтому

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

Отсюда и из неравенства (4.3) следует утверждение Леммы. □

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

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

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

$$\sigma_s(A_k) = \sqrt{\lambda_s(A_k^\top A_k)} = \sqrt{\sigma_s(A_k^\top A_k)}. \tag{4.4}$$

Из определения (4.1) матрицы $\widehat{D}$ следует, что $A_k^\top A_k = 0.5 \cdot b^{d-k}\widehat{D}$. Далее, для главной подматрицы F матрицы $\widehat{D}$ из леммы 4 можно применить теорему о чередовании сингулярных чисел симметричной матрицы [8] и получить, что $\sigma_{u-\ell}(\widehat{D}) \geq 0.5$. Отсюда и из равенства (4.4) получаем

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

Рассмотрим теперь произвольный тензор B со свойством $\|A-B\|_F \leq 0.5b^{(d-k)/2}$. Аналогичное неравенство верно и для матриц развёртки A_k и B_k. По теореме Эккарта–Юнга [10] наилучшее приближение в норме Фробениуса ранга $u-\ell-1$ имеет погрешность как минимум (здесь $N := \min\{b^k, b^{d-k}\}$ — меньший из размеров матрицы A_k)

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

поэтому

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

Рассмотрим произвольное ТТ-разложение тензора B вида (1.1). Для k-ой матрицы развёртки можно написать

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

откуда следует, что $r_k \geq \operatorname{rank} B_k \geq \frac{1}{16} b^{\frac{d-k}{2}} - 2$, 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 берётся 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 $M, d \to \infty$. С другой стороны, для функции $f(z)$ из данного раздела несложно оценить максимум модуля на единичном круге:

$$|f(z)| = \left|\sin(2\pi b^{2k} z^2)\right| = \left|\frac{1}{2i}\left(e^{2\pi i b^{2k} z^2} - e^{-2\pi i b^{2k} z^2}\right)\right| \leq e^{2\pi b^{2k}|z^2|} \leq e^{2\pi b^{2k}}. \tag{4.5}$$

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

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

В последнем равенстве за $M(d)$ обозначено число $e^{2\pi b^{2d/3}}$. Получается, что $\mathcal{R}_{b,\varepsilon,[L,R]}(M(d), d) = \Omega(\sqrt{\ln M(d)})$ при $d \to \infty$.

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 $p(x) = p_0 + p_1 x + \dots + p_n x^n$ — полином с вещественными коэффициентами. Обозначим $A := T_{b,d,[0,1]}(f)$ и $M := \sum_{i=0}^{n} |p_i|$. Зафиксируем произвольное $\varepsilon > 0$.*

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

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

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 развёртки при стремлении n к бесконечности, то доказанное утверждение даст оценку вида $O(\ln n)$ в отличии от оценки $O(\sqrt[3]{n})$, полученной в [5].*

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

Тогда для любого $n \ge \lfloor \log_2(1 + 14 b^{k/2} e^{10 b^{2k}}) \rfloor$ существует полином $P_n(x) = p_0 + \dots + p_n x^n$ такой, что:

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

Доказательство. Возьмём функцию $f(x)$ из Теоремы 3. Обозначим за $P_n(x)$ 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 [8]:

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

где x_j суть корни полинома Чебышёва степени $n+1$, то есть

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

Оценим снизу модуль $\omega'(x_j)$:

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

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

$$\sum_{\substack{0 \le j_1 < \dots < j_{n-m} \le n \\ j_\ell \ne j}} (-1)^{n-m} x_{j_1} \dots x_{j_{n-m}},$$

то есть по модулю не превосходит C_n^{n-m}. Итого получаем

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

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

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

Положим $\rho := 2$ и оценим величину M, исходя из рассуждения, аналогичного (4.5):

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

Таким образом, с учётом определения неравенства (из условия) для n для $x \in [-1,1]$ имеем

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

Это означает, что

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

Теперь возьмём произвольный тензор B, приближающий $T_{b,d,[0,1]}(P_n)$ с ошибкой не более ε. Заметим, что по неравенству треугольника для нормы он даёт приближение $T_{b,d,[0,1]}(f)$ с ошибкой не более 2ε. Из Теоремы 3 сразу получаем неравенство $\mathcal{R}(B) \ge \frac{1}{16} b^{\frac{d-k}{2}} - 2$, а в силу произвольности B и свойство 2 из условия. □

Благодарности

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

Список литературы

[1] I. Oseledets, Constructive Representation of Functions in Low-Rank Tensor Formats, Constructive Approximation 37 (2013) 1–18. URL: http://link.springer.com/10.1007/s00365-012-9175-x. doi:10.1007/s00365-012-9175-x.

[2] L. Grasedyck, Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization, Preprint (2010). URL: http://www.dfg-spp1324.de/download/preprints/preprint043.pdf.

[3] B. Khoromskij, $O(d \log N)$-Quantics Approximation of N-d Tensors in High-Dimensional Numerical Modeling, Constructive Approximation 34 (2011) 257–280. URL: https://link.springer.com/article/10.1007/s00365-011-9131-1. doi:10.1007/s00365-011-9131-1.

[4] E. E. Tyrtyshnikov, Tensor approximations of matrices generated by asymptotically smooth functions, Sbornik: Mathematics 194 (2003) 941–954. URL: http://stacks.iop.org/1064-5616/194/i=6/a=A09?key=crossref.759fa240378703afa76f11293a94de57. doi:10.1070/SM2003v194n06ABEH000747.

[5] L. Vysotsky, On Tensor-Train Ranks of Tensorized Polynomials, in: Lecture Notes in Computer Science, volume 11958 LNCS, Springer, 2020, pp. 189–196. URL: https://link.springer.com/chapter/10.1007/978-3-030-41032-2_21. doi:10.1007/978-3-030-41032-2_21.

[6] I. Oseledets, Approximation of $2^d \times 2^d$ matrices using tensor decomposition, SIAM Journal on Matrix Analysis and Applications 31 (2009) 2130–2145. doi:10.1137/090757861.

[7] I. Oseledets, Tensor-Train Decomposition, SIAM Journal on Scientific Computing 33 (2011) 2295–2317. URL: http://epubs.siam.org/doi/10.1137/090752286. doi:10.1137/090752286.

[8] . . Тыртышников, Методы численного анализа, Академия, Москва, 2007.

[9] I. Oseledets, E. Tyrtyshnikov, TT-cross approximation for multidimensional arrays, Linear Algebra and Its Applications 432 (2010) 70–88. doi:10.1016/j.laa.2009.07.024.

[10] C. Eckart, G. Young, The approximation of one matrix by another of lower rank, Psychometrika 1 (1936) 211–218. URL: http://link.springer.com/10.1007/BF02288367. doi:10.1007/BF02288367.