Устойчивость малорангового приближения тензора к белому шуму

Валиахметов Булат Ильдарович
valiahmetovbulat@mail.ru

Руководитель проекта:
Петров Сергей Владимирович

Секция «Матричные методы»

Обозначения

Обозначим за $K = \{x \in \mathbb{R}^{m_1 \times \cdots \times m_d} : \operatorname{rank} x \leq r\}$ множество тензоров ограниченного ранга. Ранг понимается одним из следующих:

- канонический ранг;
- TT-ранг (тензорный поезд);-
- ранг Таккера.

Обозначения

Пусть $y \in K$ есть тензор малого ранга и

$$z = y + n$$

есть зашумлённый тензор, где $n(i_1, \dots, i_d) \sim \mathcal{N}(0, 1)$ - белый шум. Пусть $P_K(x)$ – оператор поиска приближения для x на множестве K (возможно, не наилучшего), удовлетворяющий следующему условию квази-оптимальности:

$$\|P_K(z) - z\|_F \leq \|y - z\|_F \equiv \|n\|_F.$$

Постановка задачи

Исследовать отношение ошибки аппроксимации к норме шума:

$$\frac{\|P_K(z) - y\|_F}{\|n\|_F}$$

в зависимости от ранга исходного тензора y для различных форматов тензорных разложений.

Ошибка восстановления y

Имеет место следующая оценка:

$$
\begin{aligned}
\|P_K(z) - z\|_F^2 &\leq \|y - z\|_F^2, \\
\|(P_K(z) - y) + (y - z)\|_F^2 &\leq \|y - z\|_F^2, \\
\|P_K(z) - y\|_F^2 &\leq 2\left(P_K(z) - y, z - y\right)_F, \\
\|P_K(z) - y\|_F &\leq 2\left(\frac{P_K(z) - y}{\|P_K(z) - y\|_F}, z - y\right)_F.
\end{aligned}
$$

Тензор $\frac{P_K(\vec{z}) - \vec{y}}{\|P_K(\vec{z}) - \vec{y}\|_2} \in K - K$ имеет единичную норму, откуда

$$
\|P_K(z) - y\|_F \leq 2\|z - y\|_K,
$$

где $\|x\|_K := \sup\limits_{v \in K - K, \|v\|_F = 1} (x, v)$.

Теоретическая оценка для одноранговогo тензора

При $r = 1$ все рассматриваемые тензорные форматы совпадают. В случае $n(i_1, \ldots, i_d) \sim \mathcal{N}(0, 1)$ величина $\mathbb{E}\|n\|_K$ называется гауссовской шириной множества $K - K$.

Утверждение

Для d-мерного тензора $n \in \mathbb{R}^{m \times \cdots \times m}$ ранга $r = 1$ справедлива оценка:

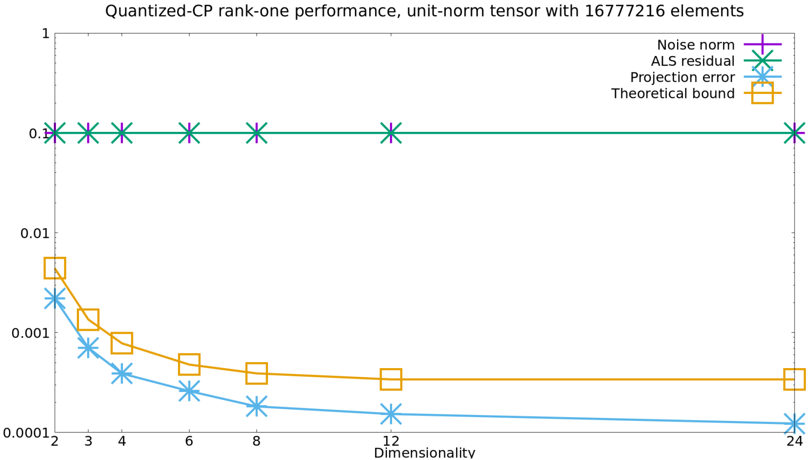
$$\mathbb{E}\|n\|_K \leqslant \sqrt{md \log d}.$$

Отсюда с учётом $\mathbb{E}\|n\|_F = m^{\frac{d}{2}}$ и $\|P_K(z) - y\|_F \leq 2\|n\|_K$ имеем

Следствие

$$\mathbb{E}\|P_K(z) - y\|_F \ll \mathbb{E}\|n\|_F.$$

Зависимость $\|P_K(z) - y\|_F$ от размерности тензора



Постановка задачи

В рамках данного проекта было рассмотрено множество K тензоров ранга не выше r для случая $r > 1$.

Исследовалась асимптотическая зависимость величины $\dfrac{\|P_K(z) - y\|_F}{\|n\|_F}$ от ранга тензора y в различных форматах тензорных разложений.

Были реализованы алгоритмы приближения тензоров в малоранговых форматах:

- TT-SVD для тензорного поезда;
- HOSVD для разложения Таккера;
- ALS для канонического разложения.

ТТ-формат

$$y(i_1, \ldots, i_d) = \sum_{\alpha_1, \ldots, \alpha_{d-1}=1}^{r_1, \ldots, r_{d-1}} G_1(i_1, \alpha_1) G_2(\alpha_1, i_2, \alpha_2) \cdots G_d(\alpha_{d-1}, i_d)$$

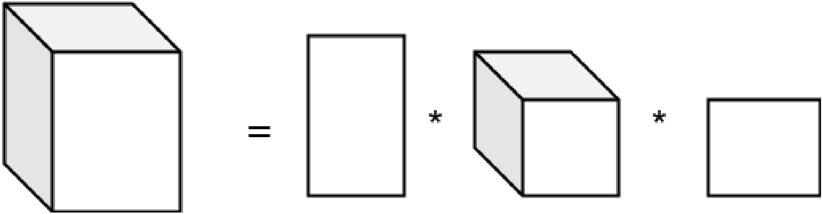


Рис.: Тензор в формате тензорного поезда (TT)

В экспериментах фиксировалось общее количество элементов тензора $N = m^d = 2^{20}$ и рассматривались различные размерности $d \in \{2, 4, 5\}$. В случае тензоров в ТТ-формате наблюдается линейный рост ошибки приближения в зависимости от ранга.

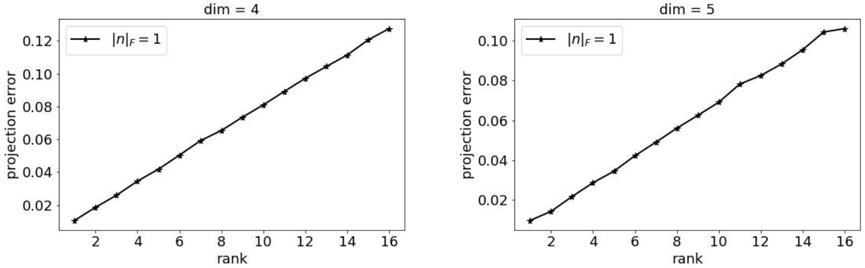


Рис.: Ошибка малорангового приближения в ТТ-формате

$$y(i_1, \ldots, i_d) = \sum_{\alpha_1, \ldots, \alpha_d = 1}^{r_1, \ldots, r_d} G(\alpha_1, \ldots, \alpha_d) U_1(i_1, \alpha_1) \cdots U_d(i_d, \alpha_d)$$

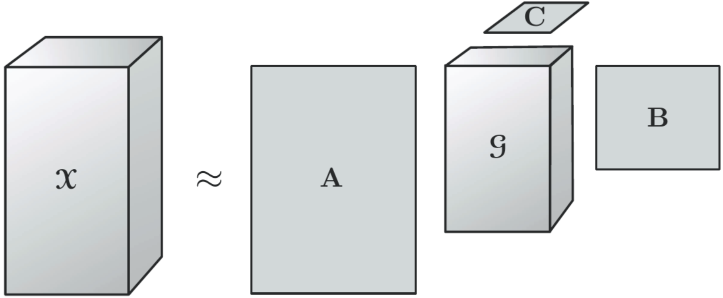


Рис.: Тензор в формате Таккера

В случае тензоров в формате Таккера наблюдается рост ошибки приближения порядка $r^{\frac{d}{2}}$.

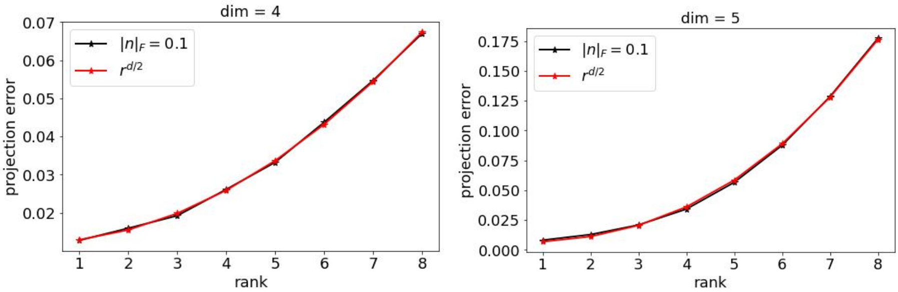


Рис.: Ошибка малорангового приближения в формате Таккера

Канонический формат

$$y(i_1, \dots, i_d) = \sum_{\alpha=1}^{r} u_1(i_1, \alpha) \cdots u_d(i_d, \alpha)$$

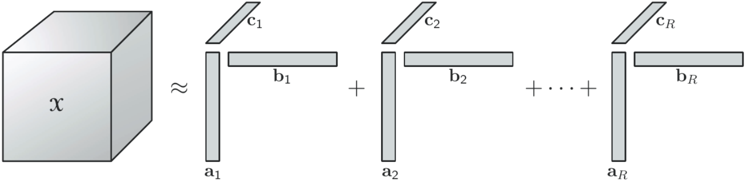


Рис.: Тензор в каноническом формате

Канонический формат

В случае тензоров в каноническом формате наблюдается рост ошибки приближения порядка $\sqrt{r}$.

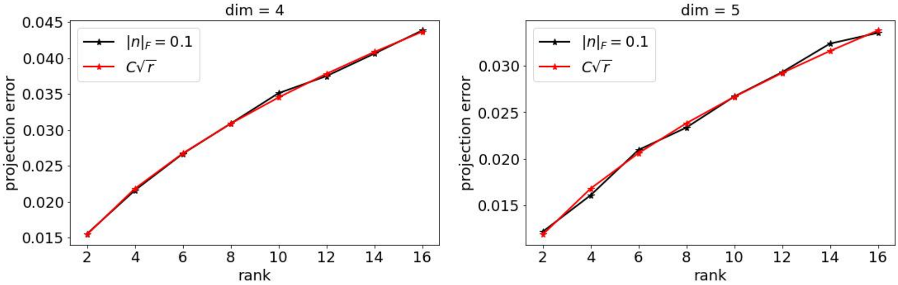


Рис.: Ошибка малорангового приближения в каноническом формате

Результаты

Для трёх исследованных форматов представления тензоров **рост ошибки приближения** зашумлённого тензора **с рангом пропорционален корню из числа параметров** в аппроксимации:

- ТТ: $\|P_K(z) - y\|_F \sim r$ при $N_{\text{appr}} = (d-1)mr^2$;
- Таккер: $\|P_K(z) - y\|_F \sim r^{\frac{d}{2}}$ при $N_{\text{appr}} = r^d + dmr$;
- Канонический: $\|P_K(z) - y\|_F \sim \sqrt{r}$ при $N_{\text{appr}} = dmr$.