Лекция 5

Основи матричних вычислений

Рахуба М.В.

15.02.21

Малоранговая аппроксимация многомерных массивов (тензоров)

$X \in \mathbb{R}^{\overbrace{n \times \dots \times n}^{d}}$ n^d $2^{300} \gg$ оценки числа атомов во Вселенной

① Каноническое разложение

в выч. лин. алг. принято называть тензорами

$\boxed{\text{Опр. 1}}$ $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$, $B \in \mathbb{R}^{m_1 \times \dots \times m_D}$

$A \circ B \in \mathbb{R}^{n_1 \times \dots \times n_d \times m_1 \times \dots \times m_D}$ называется внешним (тензорным) произведением:

$$(A \circ B)_{i_1 \dots i_d j_1 \dots j_D} = A_{i_1 \dots i_d} B_{j_1 \dots j_D}$$

$\boxed{\text{Замечание}}$ 1) '$\circ$' также обозначают '$\otimes$'. '$\otimes$' у нас используется для Кронек. произв.

2) Для примера $d = D = 2$:

$A \circ B \in \mathbb{R}^{m_1 \times m_2 \times n_1 \times n_2}$, $A \otimes B \in \mathbb{R}^{m_1 n_1 \times m_2 n_2}$, элементы – одинак., но записаны в разные размеры.

$\boxed{\text{Пример}}$ 1) $a_{ij} = u_i v_j$

$$A = u v^T = u \circ v$$

$$\mathrm{vec}(u \circ v) = v \otimes u$$

2) $A = UV^T \Rightarrow A = \sum_{\alpha=1}^{r} u_\alpha v_\alpha^T = \sum_{\alpha=1}^{r} u_\alpha \circ v_\alpha$

$A = U \Sigma V^T \Rightarrow A = \sum_{\alpha=1}^{r} \sigma_\alpha u_\alpha \circ v_\alpha$

$$a_{ijk} = \sum_{\alpha=1}^{r} u_{i\alpha} v_{j\alpha} w_{k\alpha}$$

$$A = \sum_{\alpha=1}^{r} u_\alpha \circ v_\alpha \circ w_\alpha$$

Опр. 2 1) $A = \sum_{\alpha=1}^{R} \underbrace{u_\alpha^{(1)}}_{\in \mathbb{R}^{n_1}} \circ \underbrace{u_\alpha^{(2)}}_{\in \mathbb{R}^{n_2}} \circ \ldots \circ \underbrace{u_\alpha^{(d)}}_{\in \mathbb{R}^{n_d}}$ – каноническое разлож. $A \in \mathbb{R}^{n_1 \times \ldots \times n_d}$

2) $\min R$ – канонич. ранг

3) $U^{(k)} = [u_1^{(k)}, \ldots, u_R^{(k)}] \in \mathbb{R}^{n_k \times R}$ – канонич. факторы

При $d=2$: $A = \sum_{\alpha=1}^{R} u_\alpha^{(1)} \circ u_\alpha^{(2)} = u_1^{(1)} \circ u_1^{(2)} + \ldots + u_R^{(1)} \circ u_R^{(2)}$

$$U^{(1)} = U, \quad U^{(2)} = V$$

Замечание

1) $U V^T = (U S)(S^{-1} V^T)$ – не единств.

Канонич. разлож. единственно (с точностью до перестановки слагаемых и шкалирования) при некотор. условиях. http://www.kolda.net/publication/TensorReview.pdf

2) Нет устойчивых алгоритмов приближения A тензором ранга $\leq R$. Более того, множество тензоров ранга $\leq R$ не является замкнутым:

$\exists \{A_n\} : \operatorname{rank}(A_n) = 2 \quad : \quad A_n \to A, \quad \operatorname{rank}(A) = 3$

2 Разложение Таккера

Опр 3

$$\underset{\in \mathbb{R}^{m \times n \times l}}{A} = \sum_{\alpha=1}^{R_1} \sum_{\beta=1}^{R_2} \sum_{\gamma=1}^{R_3} g_{\alpha\beta\gamma} \, u_\alpha \circ v_\beta \circ w_\gamma$$

разлож. Таккера

$$G = \{ g_{\alpha\beta\gamma} \}_{\alpha,\beta,\gamma=1}^{R_1, R_2, R_3} \in \mathbb{R}^{R_1 \times R_2 \times R_3} \quad \text{— ядро Таккера}$$

$$\left.\begin{aligned} U &= [u_1, \dots, u_{R_1}] \in \mathbb{R}^{m \times R_1} \\ V &= [v_1, \dots, v_{R_2}] \in \mathbb{R}^{n \times R_2} \\ W &= [w_1, \dots, w_{R_3}] \in \mathbb{R}^{l \times R_3} \end{aligned}\right\} \text{ факторы Таккера}$$

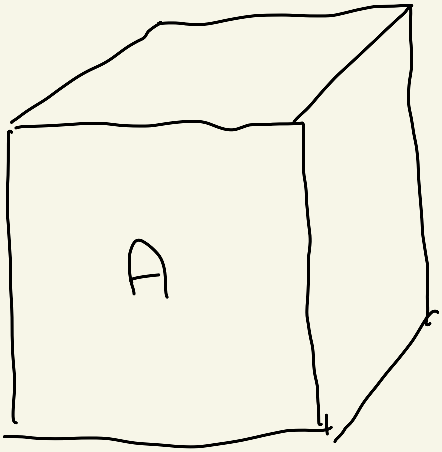
min возможные

$$(R_1, R_2, R_3) \quad \text{— мультилин. ранг.}$$
(Таккеровский ранг)

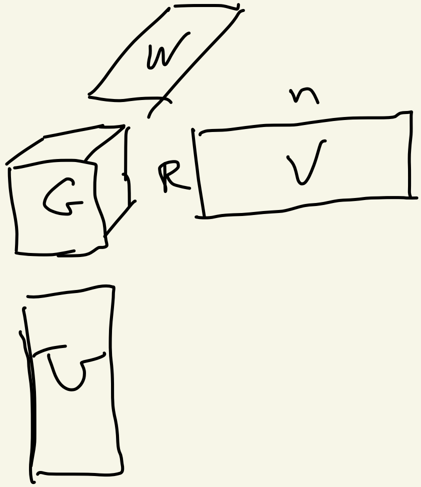
Будем писать $A = [\![G; U, V, W]\!]$

Storage

$(n=m=\ell, \quad R_1=R_2=R_3=R)$



vs.

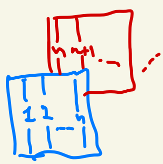


$$n^3 \gg \underbrace{R^3}_{G} + \underbrace{3nR}_{U,\,V,\,W}$$

$(R \ll n)$

JTb 1

$$\begin{bmatrix} \mathrm{vec}(A[:,:,0]) \\ \mathrm{vec}(A[:,:,1]) \\ \vdots \end{bmatrix}$$



$$\mathrm{vec}(A) = \underbrace{(W \otimes V \otimes U)}_{mn\ell \times R_1R_2R_3} \underbrace{\mathrm{vec}(G)}_{R_1R_2R_3}$$

reshape(A, (mnl, 1), order='f')

$\Updownarrow$

$$A = [\![G; U, V, W]\!]$$

$\square$ $\mathrm{vec}(u \circ v \circ w) = w \otimes v \otimes u$

$$vec(A) = \sum_{\alpha}\sum_{\beta}\sum_{\gamma} g_{\alpha\beta\gamma}\, vec(u_\alpha \circ v_\beta \circ w_\gamma) =$$

$$= \sum_{\alpha}\sum_{\beta}\sum_{\gamma} g_{\alpha\beta\gamma}\, w_\gamma \otimes v_\beta \otimes u_\alpha$$

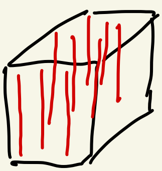
$$W \otimes V \otimes U = [w_1 \otimes v_1 \otimes u_1, \dots, w_{R_3} \otimes v_{R_2} \otimes u_{R_1}]$$

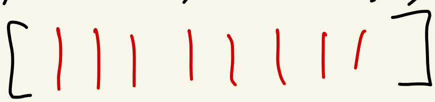
[redacted]

$\boxed{\text{Опр. 4}}$ $A_{(k)}$ – развертка по k-й моде

$$A_{(1)} = [A[:,:,0], A[:,:,1], \dots, A[:,:,-1]] \in \mathbb{R}^{m \times ne}$$

reshape(A, (m, n·e), order='f')

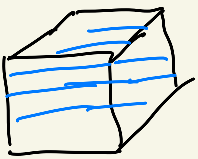


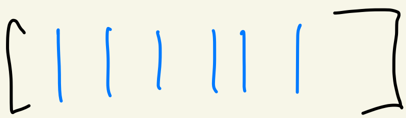


$$A_{(2)} = \text{permute}(A, [1,0,2])_{(1)} \in \mathbb{R}^{n \times me}$$

np.transpose

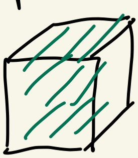
$a_{ijk} \to \tilde{a}_{jik}$





$$A_{(3)} = \text{permute}(A, [2,0,1])_{(1)} \in \mathbb{R}^{e \times mn}$$

$a_{ijk} \to \tilde{a}_{kij}$



$\boxed{\text{Лемма 1}}$

$$A_{(1)} = U\, G_{(1)}\, (W \otimes V)^T$$

$$A_{(2)} = V\, G_{(2)}\, (W \otimes U)^T$$

$$A_{(3)} = W\, G_{(3)}\, (V \otimes U)^T$$

$\square$ для $A_{(1)}$:

$$\mathrm{vec}(A) = \mathrm{vec}(A_{(1)})$$

$$\overset{1.}{=} (W \otimes V) \otimes U\, \mathrm{vec}(G)$$

$$\overset{2.}{=} (W \otimes V) \otimes U\, \mathrm{vec}(G_{(1)}) \overset{3.}{=} \mathrm{vec}\left(U\, G_{(1)}\, (W \otimes V)^T\right) \overset{4.}{=} \mathrm{vec}(A_{(1)})$$

$$B^T \otimes A\, \mathrm{vec}(X) = \mathrm{vec}(AXB)$$

$$\Downarrow$$

$$A_{(1)} = U\, G_{(1)}\, (W \otimes V)^T \qquad \blacksquare$$

$\boxed{\text{Теор}}$ Такк. ранг (R_1, R_2, R_3) равен

$$\left(\mathrm{rank}(A_{(1)}),\ \mathrm{rank}(A_{(2)}),\ \mathrm{rank}(A_{(3)})\right)$$

$\square$ 1) $$A_{(1)} = \underbrace{U}_{m \times R_1} \underbrace{G_{(1)} (W \otimes V)^T}_{R_1 \times nl}$$

$$\text{rank}(A_{(1)}) \le R_1$$
$$\phantom{\text{rank}}(A_{(2)}) R_2$$
$$\phantom{\text{rank}}(A_{(3)}) R_3$$

2) Пусть $A_{(1)} = U_1 \Sigma_1 V_1^T$ - compact SVD

$$\tilde{R}_1 = \text{rank}(A_{(1)})$$

$$U_1 U_1^T A_{(1)} = A_{(1)}$$

$$\text{vec}(A_{(1)}) = \text{vec}(U_1 U_1^T A_{(1)}) =$$

$$= \text{vec}\Big(U_1 \underbrace{U_1^T U G_{(1)}}_{\tilde{G}_{(1)}} (W \otimes V)^T\Big) =$$

$$= \text{vec}\left(U_1 \tilde{G}_{(1)} (W \otimes V)^T\right)$$

$$A_{(1)} = U_1 \tilde{G}_{(1)} (W \otimes V)^T \quad (\text{берём vec})$$

$$\text{vec}(A) = W \otimes V \otimes U_1 \, \text{vec}(\tilde{G})$$

$\tilde{G} \in \mathbb{R}^{\tilde{R}_1 \times R_2 \times R_3}$, аналогично $R_2 \to \tilde{R}_2$, $R_3 \to \tilde{R}_3$

То есть явно построили разлож. с рангами развёрток. $\blacksquare$

③ HOSVD алгоритм
(подробнее на семинаре)

U_k – r_k левых синг. вект. $A_{(k)}$:

$$\| A_{(k)} - U_k U_k^* A_{(k)} \|_F \le \varepsilon \|A\|_F$$

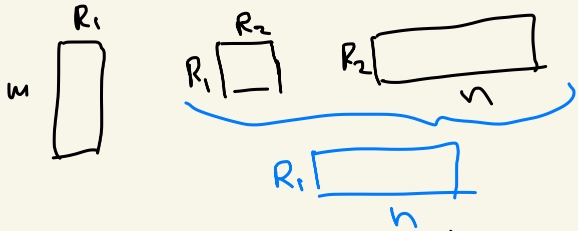
$$G = [\![A ; U_1^T, U_2^T, U_3^T]\!]$$

ф-ла для ядра.

$$\| A - [\![G ; U_1, U_2, U_3]\!] \|_F \le \sqrt{3}\, \varepsilon \|A\|_F$$

Для матриц разложение Таккера (и канонич.) переходит в скелетное:

$$A = U G V^T$$

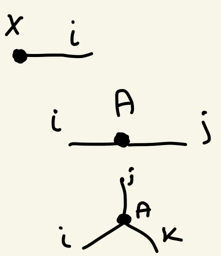


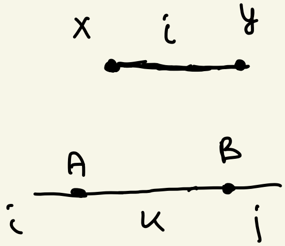
④ Тензорные сети (не входит в курс)

x_i

a_{ij}

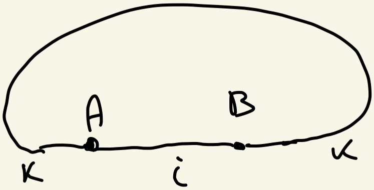
a_{ijk}





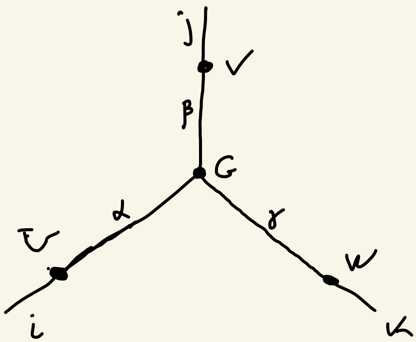
$$\sum x_i y_i$$

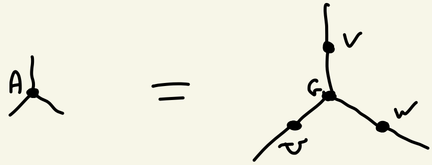
$$\sum_k A_{ik} B_{kj}$$



$$\sum_{i,k} A_{ki} B_{ik} = \mathrm{Tr}(AB)$$

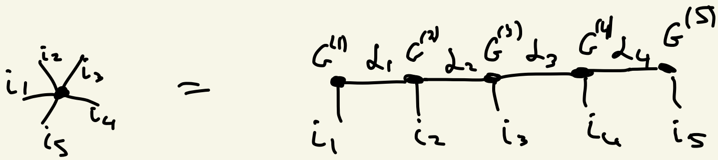
Tucker:





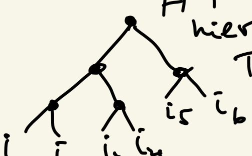
$$\sum_{\alpha,\beta,\gamma=1}^{R_1,R_2,R_3} g_{\alpha\beta\gamma}\, u_{i\alpha}\, v_{j\beta}\, w_{k\gamma}.$$

Other decomp.:

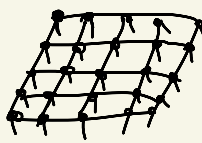


— Tensor Train (TT) a.k.a. MPS разлож.

$$\sum_{\alpha_4=1}^{R_4} \sum_{\alpha_3=1}^{R_3} \sum_{\alpha_2=1}^{R_2} \sum_{\alpha_1=1}^{R_1} G^{(1)}_{i_1\alpha_1} G^{(2)}_{\alpha_1 i_2 \alpha_2} G^{(3)}_{\alpha_2 i_3 \alpha_3} G^{(4)}_{\alpha_3 i_4 \alpha_4} G^{(5)}_{\alpha_4 i_5}$$



HT hierarchical Tucker



PEPS

MERA
(погуглите картинку)