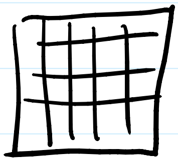
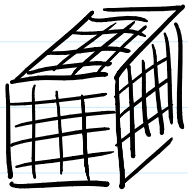
sreda, 13. maj 2020 08:15

VI. Tenzorji

6.1 Uvod

matrika 

tenzor: večdimenzionalna tabela

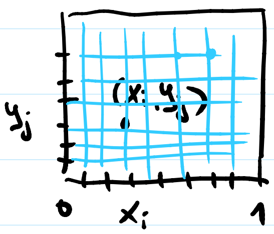


tenzor reda 3

a_{ij}

a_{ijk}

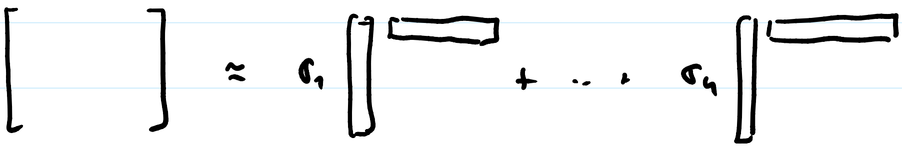
motivacija: rešujemo PDE na $[0,1]\times[0,1]$ $\Rightarrow$ rešitev je $u(x,y)$



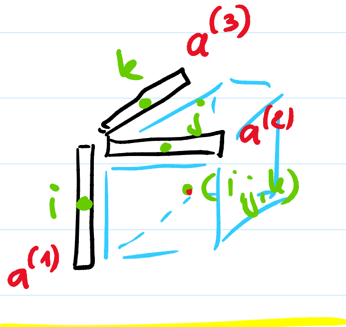
$a_{ij} = u(x_i, y_j)$ $\Rightarrow$ matrika A

Naj bo $u(x,y) = \sin(2x^5 + 5y) + x^2 + y^2$

$$A \approx \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \sigma_3 u_3 v_3^T + \sigma_4 u_4 v_4^T$$



$$xy^T = \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_m \end{bmatrix} \begin{array}{c} \begin{bmatrix} y_1 & \cdot & y_j & \cdot & y_m \end{bmatrix} \\ \begin{bmatrix} x_1y_1 & \cdots & x_1y_m \\ \vdots & x_iy_j & \vdots \\ x_my_1 & \cdots & x_my_n \end{bmatrix} \leftarrow i \\ \uparrow \\ j \end{array}$$



$\mathcal{A} \in \mathbb{R}^{M_1 \times M_2 \times \dots \times M_d}$ je tenzor reda d oziroma d-razsežna matrika (tabela) z elementi

$a_{i_1 i_2 \dots i_d} \in \mathbb{R}$

Za $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{M_1 \times \cdots \times M_d}$ lahko definiramo skalarni produkt

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1=1}^{M_1} \cdots \sum_{i_d=1}^{M_d} a_{i_1 i_2 \cdots i_d} b_{i_1 i_2 \cdots i_d}$$

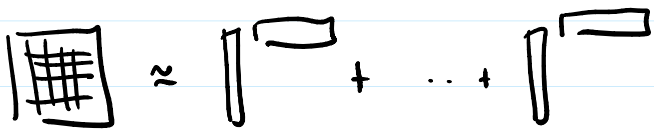
Posplošitev Frobeniusove norme je $\|\mathcal{A}\| = \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$

Tenzor $\mathcal{A}$ ima rang 1, če obstajajo vektorji $a^{(i)} \in \mathbb{R}^{M_i}$ za $i = 1, \ldots, d$, da je

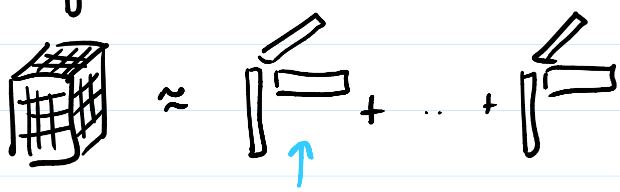
$$a_{i_1 i_2 \cdots i_d} = a^{(1)}_{i_1} a^{(2)}_{i_2} \cdots a^{(d)}_{i_d}$$

Krajši zapis je $\mathcal{A} = a^{(1)} \circ a^{(2)} \circ \cdots \circ a^{(d)}$ zunanji produkt vektorjev $a^{(1)}, \ldots, a^{(d)}$

Če se matriko da dobro aproksimirati z matriko nizkega ranga:



Posplošitev za tenzor reda 3:



tenzor ranga 1

$$\mathcal{A} \approx \sum_{i=1}^{k} \lambda_i \, u_i \circ v_i \circ w_i$$

$\Rightarrow$ namesto $O(M^3)$ prostora potrebujemo le $O(kM)$ $\quad M_1 = M_2 = M_3 = M$

$A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$. Kroneckerjev produkt $A \otimes B$ je matrika $mp \times nq$ oblike

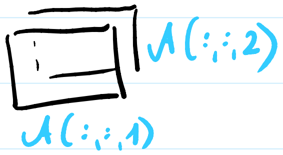
$$A \otimes B = \begin{bmatrix} a_{11} B & a_{12} B & \cdots & a_{1n} B \\ \vdots & & & \vdots \\ a_{m1} B & a_{m2} B & \cdots & a_{mn} B \end{bmatrix}$$

Za $x \in \mathbb{R}^m$ in $y \in \mathbb{R}^n$ je $x \otimes y = \begin{bmatrix} x_1 y \\ x_2 y \\ \vdots \\ x_m y \end{bmatrix} \in \mathbb{R}^{mn}$

Vektorizacija je operacija, kjer vse elemente tenzorja zložimo v vektor.
Pri tem uporabimo obratno leksikografsko urejenost:

$$i_1 i_2 \cdots i_d < j_1 j_2 \cdots j_d \iff i_k < j_k \text{ in } i_\ell = j_\ell \text{ za } \ell = k+1,\dots,d$$
za neki k

$\mathcal{A} \in \mathbb{R}^{2\times 3\times 2}$ $\qquad \mathcal{A}(:,:,1) = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}, \mathcal{A}(:,:,2) = \begin{bmatrix} 7 & 9 & 11 \\ 8 & 10 & 12 \end{bmatrix}$



$$\mathrm{vec}(\mathcal{A}) = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ 11 \\ 12 \end{bmatrix} = \begin{bmatrix} a_{111} \\ a_{211} \\ \vdots \\ a_{132} \\ a_{232} \end{bmatrix}$$

111
211
121
⋮
132
232

Pri matrifikaciji elemente tenzorja zložimo v matriko. Poseben primer so razpetja v izbrani smeri (indeksu)

$\mathcal{A}(:,:,1) = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}, \mathcal{A}(:,:,2) = \begin{bmatrix} 7 & \boxed{9} & 11 \\ 8 & 10 & 12 \end{bmatrix}$

$$\mathcal{A}_{(1)} = \begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 11 \\ 2 & 4 & 6 & 8 & 10 & 12 \end{bmatrix}$$

(1,1) (2,1) (3,1) (1,2) (2,2) (3,2) → (2,3)

1, 2 ↓ (1)

a_{231}

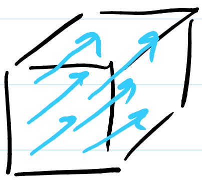
$$\mathcal{A}_{(2)} = \begin{bmatrix} 1 & 2 & 7 & 8 \\ 3 & 4 & \boxed{9} & 10 \\ 5 & 6 & 11 & 12 \end{bmatrix}$$

(1,1) (2,1) (1,2) (2,2) → (1,3)

1, 2, 3 ↓ (2)

a_{122}

$$\mathcal{A}_{(3)} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 7 & 8 & 9 & 10 & 11 & 12 \end{bmatrix}$$



Tenzor $\mathcal{X} \in \mathbb{R}^{m_1\times\cdots\times m_d}$ množimo v smeri $j \in \{1,\dots,d\}$ z matriko $A \in \mathbb{R}^{m_j \times m_j}$ in dobimo tenzor

$$\mathcal{X} \times_j A \in \mathbb{R}^{m_1\times\cdots\times m_{j-1}\times m_j \times m_{j+1}\times\cdots\times m_d},$$ kjer je

$$(\mathcal{X} \times_j A)_{i_1\cdots i_{j-1} k i_{j+1}\cdots i_d} = \sum_{i_j=1}^{m_j} x_{i_1\cdots i_d}\, a_{k i_j}$$

To je ekvivalentno: $\underline{(\mathcal{X} \times_j A)_{(j)} = A\, \mathcal{X}_{(j)}}$

$$\mathrm{vec}(\mathcal{X} \times_j A) = \underline{(I_{M_d} \otimes \cdots \otimes I_{M_{j+1}} \otimes A \otimes I_{M_{j-1}} \otimes \cdots \otimes I_{M_1})\, \mathrm{vec}(\mathcal{X})}$$

Velja
$$(\mathcal{X} \times_j A) \times_k B = (\mathcal{X} \times_k B) \times_j A \quad \text{za } j \neq k$$

$$(A \otimes B)(C \otimes D) = AC \otimes BD \qquad A, B, C, D \text{ ustreznih dimenzij}$$
$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

$$(\mathcal{X} \times_j A) \times_j B = \mathcal{X} \times_j (BA)$$

6.2 Rang tenzorja

Rang tenzorja $\mathcal{X} \in \mathbb{R}^{M_1 \times M_2 \times M_3}$ je najmanjši r, za katerega je možno zapisati $\mathcal{X}$ kot vsoto r tenzorjev ranga 1.

$$\mathcal{X} = \sum_{i=1}^{r} \lambda_i\, a_i \circ b_i \circ c_i \qquad a_i \in \mathbb{R}^{M_1},\ b_i \in \mathbb{R}^{M_2},\ c_i \in \mathbb{R}^{M_3}$$

$$\mathcal{X} = [\![\lambda ; A, B, C]\!] \qquad \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{bmatrix} \qquad \begin{matrix} A = [a_1 \cdots a_r] \\ B = [b_1 \cdots b_r] \\ C = [c_1 \cdots c_r] \end{matrix}$$

kompakten zapis

za vektorje a_i, b_i, c_i lahko zahtevamo, da so normirani.

Med rangom tenzorja $(d > 2)$ in matrike $(d = 2)$ so pomembne razlike:

① pri tenzorjih lahko dobiš drugačen rang, če gledamo nad $\mathbb{R}$ ali nad $\mathbb{C}$.

$2\times 2\times 2$ tenzor $\mathcal{X}(:,:,1)=\begin{bmatrix}1&0\\0&1\end{bmatrix}$, $\mathcal{X}(:,:,2)=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$

ima rang 3 nad $\mathbb{R}$ in rang 2 nad $\mathbb{C}$:

$\mathcal{X}=[A,B,C]$ za a) $A=\begin{bmatrix}1&0&1\\0&1&-1\end{bmatrix}$, $B=\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}$

$C=\begin{bmatrix}1&1&0\\-1&1&0\end{bmatrix}$

b) $A=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\-i&i\end{bmatrix}$, $B=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\i&-i\end{bmatrix}$, $C=\begin{bmatrix}1&1\\i&-i\end{bmatrix}$

(2) $\mathcal{X}=a_1\circ b_1\circ c_2+a_1\circ b_2\circ c_1+a_2\circ b_1\circ c_1$: rang 3

a_1, a_2
b_1, b_2
c_1, c_2

$$\mathcal{Z}_\alpha=\alpha\left(a_1+\tfrac{1}{\alpha}a_2\right)\circ\left(b_1+\tfrac{1}{\alpha}b_2\right)\circ\left(c_1+\tfrac{1}{\alpha}c_2\right)-\alpha\, a_1\circ b_1\circ c_1 \quad : \text{rang } 2$$

$$\mathcal{X}-\mathcal{Z}_\alpha=\frac{1}{\alpha}\left(a_2\circ b_2\circ c_1+a_2\circ b_1\circ c_2+a_1\circ b_2\circ c_2+\frac{1}{\alpha}a_2\circ b_2\circ c_2\right)$$

$$\lim_{\alpha\to\infty}\|\mathcal{X}-\mathcal{Z}_\alpha\|=0$$

(3) če izberemo naključen $\mathbb{R}$ tenzor $2\times 2\times 2$

$\Rightarrow$ 79% verjetnost. da bo imel rang 2

21% -||- rang 3

(4) maks rang $2\times 2\times 2$ tenzorja je 3

$3\times 3\times 3$ je 5

maksimalno $m\times n$

$\min(m,n)$

splošna formula ni znana