CP ALS

torek, 19. maj 2020 23:30

Ponovitev tenzorjev:

$$\mathcal{A} \in \mathbb{R}^{M_1 \times M_2 \times M_3} = \sum_{i=1}^{r} \lambda_i \, a_i \circ b_i \circ c_i \qquad \begin{array}{l} a_i \in \mathbb{R}^{M_1} \\ b_i \in \mathbb{R}^{M_2} \\ c_i \in \mathbb{R}^{M_3} \end{array}$$

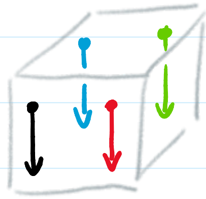
minimalni r je rang tenzorja $\mathcal{A}$

določitev ranga tenzorja je NP težek problem! (za $d>2$)

$\mathrm{vec}(\mathcal{A})$: tenzor $\rightarrow$ vektor $\in \mathbb{R}^{M_1 M_2 M_3}$

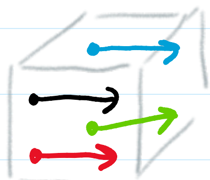
matrifikacija : tenzor $\rightarrow$ matrika

$\mathcal{A}_{(1)} \in \mathbb{R}^{M_1 \times M_2 M_3}$ $\qquad \mathcal{A} =$



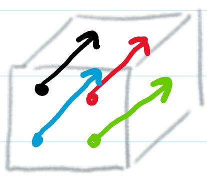
$\mathcal{A}_{(1)} = \begin{bmatrix} \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix}$

$\mathcal{A}_{(2)} \in \mathbb{R}^{M_2 \times M_1 M_3}$ $\qquad \mathcal{A} :$



$\mathcal{A}_{(2)} = \begin{bmatrix} \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix}$

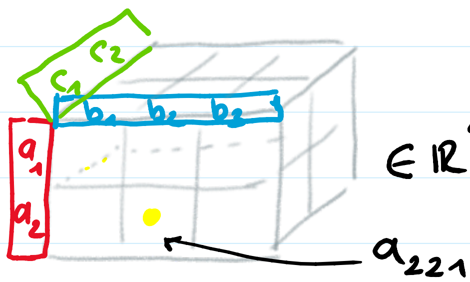
$\mathcal{A}_{(3)} \in \mathbb{R}^{M_3 \times M_1 M_2}$ $\qquad \mathcal{A} :$



$\mathcal{A}_{(3)} = \begin{bmatrix} \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix}$

Poseben primer:

$$\mathcal{A} = a \circ b \circ c = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \circ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \circ \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} =$$



$\in \mathbb{R}^{2 \times 3 \times 2}$

$$\mathcal{A}_{(1)} = \begin{bmatrix} a_1 b_1 c_1 & a_1 b_2 c_1 & a_1 b_3 c_1 & a_1 b_1 c_2 & a_1 b_2 c_2 & a_1 b_3 c_2 \\ a_2 b_1 c_1 & a_2 b_2 c_1 & a_2 b_3 c_1 & a_2 b_1 c_2 & a_2 b_2 c_2 & a_2 b_3 c_2 \end{bmatrix} = a \, (c \otimes b)^T$$

$$c \otimes b = \begin{bmatrix} c_1 b \\ c_2 b \end{bmatrix} = \begin{bmatrix} c_1 b_1 & c_1 b_2 & c_1 b_3 & c_2 b_1 & c_2 b_2 & c_2 b_3 \end{bmatrix}^T$$

$$\mathcal{A}_{(2)} = \begin{bmatrix} a_1 b_1 c_1 & a_2 b_1 c_1 & a_1 b_1 c_2 & a_2 b_1 c_2 \\ a_1 b_2 c_1 & a_2 b_2 c_1 & a_1 b_2 c_2 & a_2 b_2 c_2 \\ a_1 b_3 c_1 & a_2 b_3 c_1 & a_1 b_3 c_2 & a_2 b_3 c_2 \end{bmatrix} = b\,(c \otimes a)^T$$

$$\mathcal{A}_{(3)} = \begin{bmatrix} a_1 b_1 c_1 & a_2 b_1 c_1 & a_1 b_2 c_1 & a_2 b_2 c_1 & a_1 b_3 c_1 & a_2 b_3 c_1 \\ a_1 b_1 c_2 & a_2 b_1 c_2 & a_1 b_2 c_2 & a_2 b_2 c_2 & a_1 b_3 c_2 & a_2 b_3 c_2 \end{bmatrix} = c\,(b \otimes a)^T$$

splošno, če imamo $\mathcal{A} = v_1 \circ v_2 \circ \dots \circ v_d$

$$\mathcal{A}_{(i)} = v_i \,(v_d \otimes \dots \otimes v_{i+1} \otimes v_{i-1} \otimes \dots \otimes v_1)^T$$

6.3 Iskanje najboljše aproksimacije s tenzorjem ranga r

Imamo $\mathcal{X} \in \mathbb{R}^{M_1 \times M_2 \times M_3}$, iščemo najboljšo aproksimacijo s tenzorjem ranga r. Rang predpišemo naprej.

$$\mathcal{X} \approx \sum_{i=1}^{r} \lambda_i\, a_i \circ b_i \circ c_i \quad , \quad \lambda_i \in \mathbb{R},\ \underbrace{a_i \in \mathbb{R}^{M_1},\ b_i \in \mathbb{R}^{M_2},\ c_i \in \mathbb{R}^{M_3}}_{\text{normirani}}$$

$$= [\![\lambda; A, B, C]\!] \qquad \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{bmatrix}$$

$$A = [a_1 \dots a_r], \quad B = [b_1 \dots b_r], \quad C = [c_1, \dots, c_r]$$

Če tenzor zapišemo kot vsoto tenzorjev ranga 1, je to kanonični poliadični (CP) razcep.

„Direktni" algoritem, kot je npr. SVD za matrike, ne obstaja.

Iščemo (za dani r), $\lambda \in \mathbb{R}^r$, $A \in \mathbb{R}^{M_1 \times r}$, $B \in \mathbb{R}^{M_2 \times r}$, $C \in \mathbb{R}^{M_3 \times r}$, kjer je dosežen „minimum"

$$\| \mathcal{X} - [\![\lambda; A, B, C]\!] \|$$

Za reševanje tega optimizacijskega problema uporabimo metodo alternirajočih najmanjših kvadratov (ALS).

A, B, C ni nujno da imajo normirane stolpce (to lahko naredimo v zadnjem koraku) ⇒ iščemo minimum $\| \mathcal{X} - \underbrace{[\![A, B, C]\!]}_{\sum_{i=1}^{r} a_i \circ b_i \circ c_i} \|$

Glavna ideja metode ALS:

začnemo z A_0, B_0, C_0

fiksiramo B_0, C_0 in poiščemo A_1, ki minimizira $\| \mathcal{X} - [\![A_1, B_0, C_0]\!] \|$

to je navaden predoločen sistem za elemente A_1

fiksiramo A_1, C_0 in poiščemo B_1, ki min $\| \mathcal{X} - [\![A_1, B_1, C_0]\!] \|$

fiksiramo A_1, B_1 in poiščemo C_1, ki min. $\| \mathcal{X} - [\![A_1, B_1, C_1]\!] \|$

iščemo min $f(x,y)$

metoda koordinatnega spusta

(x_0, y_0)

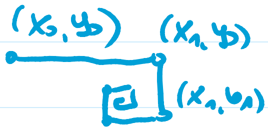
min $f(x, y_0)$ je pri $x = x_1$

min $f(x_1, y)$ je pri $y = y_1$

(x_0, y_0)

(x_1, y_0)

(x_1, y_1)



Algoritem za ALS-CP

izberi $A_0 \in \mathbb{R}^{M_1 \times r}$, $B_0 \in \mathbb{R}^{M_2 \times r}$, $C_0 \in \mathbb{R}^{M_3 \times r}$ z norm. stolpci

$k = 0, 1, \dots$

poišči A_{k+1}, ki minimizira $\| \mathcal{X} - [\![A_{k+1}, B_k, C_k]\!] \|$ in normiraj stolpce A_{k+1}

poišči B_{k+1}, -"- $\| \mathcal{X} - [\![A_{k+1}, B_{k+1}, C_k]\!] \|$ -"- B_{k+1}

poišči S, -"- $\| \mathcal{X} - [\![A_{k+1}, B_{k+1}, S]\!] \|$

zapiši $S = \mathrm{diag}(\lambda)\, C_{k+1}$, kjer so stolpci C_{k+1} normirani

Izvajamo toliko časa, dokler se $\|\mathcal{X} - \tilde{\mathcal{X}}\|$ ne zmanjšuje več.

Konvergenca do globalnega minimuma (ki lahko sploh ne obstaja) ni zagotovljena, niti do lokalnega ne, prav tako je lahko zelo počasna.

V večini primerov pa metoda deluje zadovoljivo.

Zgled: Tabeliramo $u(x,y,z) = \sin(2x^2 + 5y - 3z^2 + 5xy^2z) + x^2 + y^2 - z^2$ na $[0,1]\times[0,1]\times[0,1]$ s tenzorjem $\mathcal{A} \in \mathbb{R}^{200\times 200\times 200}$

Kako rešimo LS problem v algoritmu ALS-CP?

Iščemo $B \in \mathbb{R}^{n_2\times r}$, ki minimizira $\|\mathcal{X} - \underbrace{[\![A, B, C]\!]}_{\mathcal{Y}}\| = \|\mathcal{X} - \mathcal{Y}\|$

$$\|\mathcal{X} - \mathcal{Y}\| = \|X_{(2)} - Y_{(2)}\|_F$$

$$\mathcal{Y} = \sum_{i=1}^{r} a_i \circ b_i \circ c_i \quad \Rightarrow \quad Y_{(2)} = \sum_{i=1}^{r} b_i (c_i \otimes a_i)^T$$

$$Y_{(2)} = \underbrace{[b_1\ b_2\ \cdots\ b_r]}_{B} \cdot \begin{bmatrix} (c_1\otimes a_1)^T \\ \vdots \\ (c_r \otimes a_r)^T \end{bmatrix} = B(C \odot A)^T$$

$$C\odot A = [c_1\otimes a_1\ \ c_2\otimes a_2\ \cdots\ c_r\otimes a_r]$$ Khatri-Raojev produkt

$C\odot A \in \mathbb{R}^{n_1 n_3 \times r}$

$$\Rightarrow \text{iščemo} \min_{B\in\mathbb{R}^{n_2\times r}} \|X_{(2)} - B(C\odot A)^T\|_F$$

$$= \min_{B\in\mathbb{R}^{n_2\times r}} \|(C\odot A)B^T - X_{(2)}^T\|_F$$

to je predoločen sistem za B^T, matrika je velikosti $n_1 n_3 \times r$

$B^T \in \mathbb{R}^{r\times n_2}$

To lahko rešimo s pomočjo normalnega sistema. Dobimo

$$\boxed{(C\odot A)^T (C\odot A)\, B^T = (C\odot A)^T \mathcal{X}_{(2)}^T}$$

$(C\odot A)^T(C\odot A)$ je matrika velikosti $r\times r$

$$(C\odot A)^T(C\odot A) = (C^TC) * (A^TA)$$

Hadamardov produkt (množenje istoležnih elementov)

$A, B\in\mathbb{R}^{m\times n}$ $\Rightarrow$ $A * B = \begin{bmatrix} a_{11}b_{11} & \cdots & a_{1n}b_{1n} \\ \vdots & & \vdots \\ a_{m1}b_{m1} & \cdots & a_{mn}b_{mn} \end{bmatrix}$

6.4 Tuckerjev razcep

$$(\mathcal{X} \times_1 A)_{(1)} = A\,\mathcal{X}_{(1)}$$

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

$d=3$:

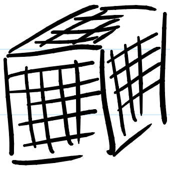
$$\operatorname{vec}(\mathcal{X} \times_1 A) = (I_{M_3} \otimes I_{M_2} \otimes A) \operatorname{vec}(\mathcal{X})$$
$$\operatorname{vec}(\mathcal{X} \times_2 B) = (I_{M_3} \otimes B \otimes I_{M_1}) \operatorname{vec}(\mathcal{X})$$
$$\operatorname{vec}(\mathcal{X} \times_3 C) = (C \otimes I_{M_2} \otimes I_{M_1}) \operatorname{vec}(\mathcal{X})$$

$$\operatorname{vec}(\mathcal{X} \times_1 A \times_2 B \times_3 C) = (C\otimes B\otimes A) \operatorname{vec}(\mathcal{X})$$

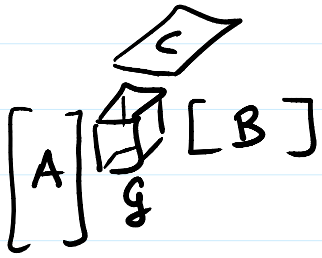
CP razcep: iščemo dobro aproksimacijo oblike

$= [\![\lambda ; A, B, C]\!]$

Tuckeyjev razcep



$$\approx \mathcal{G} \times_1 A \times_2 B \times_3 C =$$



$\mathcal{G} \in \mathbb{R}^{r_1 \times r_2 \times r_3}$: jedro , $A \in \mathbb{R}^{n_1 \times r_1}$, $B \in \mathbb{R}^{n_2 \times r_2}$, $C \in \mathbb{R}^{n_3 \times r_3}$

$= \underbrace{[\![\mathcal{G} ; A, B, C]\!]}_{\mathcal{Y}}$: kjer

$$y_{ijk} = \sum_{p=1}^{r_1} \sum_{q=1}^{r_2} \sum_{s=1}^{r_3} g_{pqs} \, a_{ip} \, b_{jq} \, c_{ks}$$

CP zapis je poseben primer Tuckeyje, če vzamemo $\mathcal{G}$ superdiagonalni tenzor, kjer je
$$g_{ijk} = \begin{cases} \lambda_i & \text{za } i=j=k \\ 0 & \text{sicer} \end{cases}$$