Matrične i tenzorske metode u analizi podataka

©Zlatko Drmač
PMF Matematički Odsjek, Zagreb

Zagreb 2022/23

Sadržaj

Uvod

$(\mathbb{R}^n, +, \cdot)$ je univerzalni predstavnik n-dimenzionalnog vektorskog prostora nad $\mathbb{R}$. Njegove elemente zovemo vektori, $x \in \mathbb{R}^n$, i u nekoj odabranoj bazi imamo reprezentaciju

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad x_i \in \mathbb{R} \text{ skalari.}$$

Iako sastavljeni od skalara x_i, vektori $x \in \mathbb{R}^n$ imaju bogatu strukturu

- skalarni produkt, npr. $\langle x, y \rangle_2 = \sum_{i=1}^n y_i x_i = y^T x$
- norma, npr. $\|x\|_2 = \sqrt{x^T x}$, $\|x\|_1$, $\|x\|_\infty$
- kut među vektorima, projekcija, ortonormirana baza, potprostori, Gram-Schmidtova ortogonalizacija
- $\mathbb{R}^n$ može npr. reprezentirati prostor polinoma $\mathcal{P}_{n-1}$
- kao struktura podataka u razvoju koda je niz (polje) indeksiran s jednim indeksom, $x = (x_1, \ldots, x_n)^T$

Uvod

Sada pogledajmo matrice, npr. realne dimenzija $m \times n$.

$$A = (a_{ij})_{i,j=1}^{m,n} \in \mathbb{R}^{m\times n}; \quad A = \underbrace{\begin{pmatrix} \star & \circ & \bullet \\ \star & \circ & \bullet \\ \star & \circ & \bullet \\ \star & \circ & \bullet \end{pmatrix}}_{\substack{kolekcija \\ stupaca}} = \underbrace{\begin{pmatrix} \star & \star & \star \\ \bullet & \bullet & \bullet \\ \ast & \ast & \ast \\ \circ & \circ & \circ \end{pmatrix}}_{\substack{kolekcija \\ redaka}}$$

$$A = \begin{pmatrix} \star & \circ & \bullet \\ \star & \circ & \bullet \\ \star & \circ & \bullet \\ \star & \circ & \bullet \end{pmatrix} = (a_1, a_2, \ldots, a_n), \quad a_i \in \mathbb{R}^m. \quad \mathrm{vec}(A) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^{m\cdot n}$$

$\mathrm{vec} : \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{m\cdot n}$ je izomorfizam vektorskih prostora. Lako se provjeri $\mathrm{vec}(\alpha A + \beta B) = \alpha \mathrm{vec}(A) + \beta \mathrm{vec}(B)$; α, β skalari, $A, B \in \mathbb{R}^{m\times n}$.
Frobeniusov skalarni produk i norma
$\langle A, B\rangle_F = \langle \mathrm{vec}(A), \mathrm{vec}(B)\rangle_2 = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij} = \mathrm{Trag}(B^T A)$;
$\|A\|_F = \sqrt{\langle A, A\rangle_F}$.
Naravno, matrice su puno više od kolekcije skalara a_{ij}, ili kolekcije vektora (stupaca, redaka).

Uvod

Prostor matrica ima bogatiju strukturu:

- Matrica je zapis linearnog operatora u paru baza. Ako je $A \in \mathbb{R}^{m \times n}$ i $x \in \mathbb{R}^n$ onda imamo $y = Ax \in \mathbb{R}^m$, $y_i = \sum_k a_{ik} x_k$.
- Produkt matrica odgovarajućih dimenzija, $C = AB$, definiran s $c_{ij} = \sum_k a_{ik} b_{kj}$, daje matricu koja reprezentira kompoziciju linearnih operatora koji odgovaraju A i B
- Za kvadratne matrice imamo pojam regularne matrice i inverza; strukturu algebre itd.
- Imamo svojstvene vrijednosti i vektore, SVD dekompoziciju, QR faktorizaciju, rang, optimalnu aproksimaciju nižim rangom, pseudo-inverz, itd.

Iako kao strukturu podataka matricu reprezentiramo kao kolekciju skalara a_{ij} indeksiranu s dva indeksa, unutar koje imamo podstrukture stupaca (indeks j fiksiran) i redaka (indeks i fiksiran), jasno nam je da je akcija u prostoru matrica na razini višoj od skalara i vektora.

Vektorski model podataka

Puno različitih vrsta podataka možemo matematički reprezentirati kao vektore u n-dimenzionalnom vektorskom prostoru $\mathbb{R}^m$ ili kao $m \times n$ matrice. Na primjer:

- Račun u dućanu

$$= \begin{pmatrix} \vdots \\ 17.99 \\ 23.99 \\ \vdots \end{pmatrix} = a_k \in \mathbb{R}^m$$

Kolekcija računa $A = (a_1, a_2, \ldots, a_n)$ sadrži informacije o navikama kupaca. Koordinate u $\mathbb{R}^m$ su jednostavno imenovane prema artiklima koji su u ponudi i koji imaju svoj jedinstven ID.

Slično, pikseli u $m \times n$ matrici definiraju digitalnu sliku.

SVD dekompozicija (Singular Value Decomposition)

Neka je A $m \times n$ matrica A ranga r. Tada postoje unitarne matrice $U = (u_1, \ldots, u_m)$, $V = (v_1, \ldots, v_n)$ ($U^*U = I_m$, $V^*V = I_n$) i dijagonalna Σ tako da je (npr. $m \geq n$)

$$\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} = \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \begin{pmatrix} \star & 0 & 0 \\ 0 & \star & 0 \\ 0 & 0 & \star \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x & x & x \\ x & x & x \\ x & x & x \end{pmatrix}$$

$$A = U\Sigma V^* = \sum_{k=1}^{r} \sigma_i u_i v_i^*$$

$$\Sigma = \mathrm{diag}(\sigma_i), \quad \sigma_1 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_{\min(m,n)}$$

σ_i= singularna vrijednost;
u_i= lijevi singularni vektor; v_i=desni singularni vektor. $Av_i = \sigma_i u_i$
$H = A^*A = V\Sigma^*\Sigma V^*$, $M = AA^* = U\Sigma\Sigma^* U^*$. $Hv_i = \sigma_i^2 v_i$, $Mu_i = \sigma_i^2 u_i$
$A^\dagger = V\Sigma^\dagger U^* = \sum_{k=1}^{r} (1/\sigma_i) v_i u_i^*$, $\|Ax - b\|_2 \to \min$, $x = A^\dagger b$

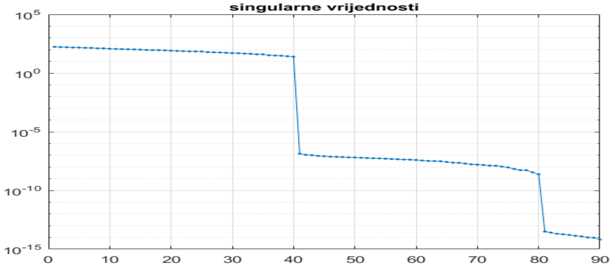
SVD dekompozicija (Singular Value Decomposition)

Theorem (Eckart–Young–Mirsky–Schmidt)

Neka je $\ell < r$ *i* $A_\ell = \sum_{i=1}^{\ell} \sigma_i u_i v_i^*$. *Tada je*

$$\min_{rang(X)\leq \ell} \|A - X\|_F = \|A - A_\ell\|_F = \sqrt{\sum_{i=\ell+1}^{r} \sigma_i^2}$$

$$\min_{rang(X)\leq \ell} \|A - X\|_2 = \|A - A_\ell\|_2 = \sigma_{\ell+1}$$



Hadamardov i Kroneckerov produkt

Hadmardov matrični produkt $C = A \circ B$ $(C = A * B)$

Za $A, B \in \mathbb{R}^{m \times n}$, njihov Hadamardov produkt $C = A \circ B$ je definiran s $c_{ij} = a_{ij} b_{ij}$. Ovakvo množenje je praktično kod obrade digitalnih slika gdje je A slika a B ovim množenjem mijenja točno odgovarajuće piksele.

Kroneckerov matrični produkt $C = A \otimes B$

Za $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$ Kroneckerovog produkt $A \otimes B$ je matrica

$$C = A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}$$

Primijetimo da je produkt dobro definiran za proizvoljne dimenzije A i B i da je $A \otimes B$ dimenzija $m \cdot p \times n \cdot q$.

$$\begin{pmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{pmatrix} \otimes \begin{pmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{pmatrix} = \begin{pmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \end{pmatrix}$$

Svojstva Kroneckerovog produkta

Ako su $A = (a_1, \ldots, a_m)$, $B = (b_1, \ldots, b_q)$ stupčane particije onda je

$$\begin{aligned} A \otimes B &= (a_1 \otimes B, a_2 \otimes B, \ldots, a_n \otimes B) \\ &= (a_1 \otimes b_1, a_1 \otimes b_2, \ldots, a_1 \otimes b_q, a_2 \otimes b_1, \ldots, a_n \otimes b_1, \ldots, a_n \otimes b_q) \end{aligned}$$

Osnovna svojstva:

- $A \otimes (\alpha B) = \alpha (A \otimes B)$
- $(A + B) \otimes C = A \otimes C + B \otimes C$; $A \otimes (B + C) = A \otimes B + A \otimes C$
- $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
- $(A \otimes B)^T = A^T \otimes B^T$; $(A \otimes B)^* = A^* \otimes B^*$ (za kompleksne matrice)
- $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ (uz odgovarajuće dimenzije)
- $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ (za regularne); $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$
- $\mathrm{vec}(CXD) = (D^T \otimes C)\mathrm{vec}(X)$
- $\mathrm{vec}(xy^T) = y \otimes x$ (x, y vektori stupci)
- $\mathrm{vec}((x, \widehat{x}) \begin{pmatrix} y^T \\ \widehat{y}^T \end{pmatrix}) = y \otimes x + \widehat{y} \otimes \widehat{x}$

Khatri-Raov produkt

Khatri-Raov produkt $C = A \odot B$

Neka su $A = (a_1, \ldots, a_n) \in \mathbb{R}^{m \times n}$, $B = (b_1, \ldots, b_n) \in \mathbb{R}^{p \times n}$ dane stupčanim particijama. Khatri-Raov produkt A i B je

$$A \odot B = (a_1 \otimes b_1, a_2 \otimes b_2, \ldots, a_{n-1} \otimes b_{n-1}, a_n \otimes b_n)$$

Osnovna svojstva

- $(A \odot B) \odot C = A \odot (B \odot C)$
- $(A \odot B)^T (A \odot B) = (A^T A) * (B^T B)$ (ovdje je $*$ Hadamardov produkt)
- $(A \odot B)^\dagger = ((A^T A) * (B^T B))^\dagger (A \odot B)^T$

$C \otimes B$ i $C \odot B$

Primjer:

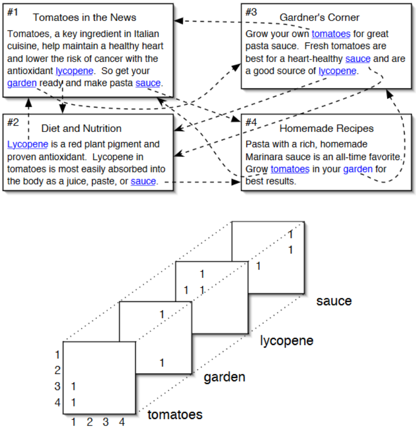
$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = (c_1, c_2), \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = (b_1, b_2)$$

$$C \otimes B = \begin{pmatrix} c_{11}b_{11} & c_{11}b_{12} & c_{12}b_{11} & c_{12}b_{12} \\ c_{11}b_{21} & c_{11}b_{22} & c_{12}b_{21} & c_{12}b_{22} \\ c_{21}b_{11} & c_{21}b_{12} & c_{22}b_{11} & c_{22}b_{12} \\ c_{21}b_{21} & c_{21}b_{22} & c_{22}b_{21} & c_{22}b_{22} \end{pmatrix}$$

$C \odot B$ je podmatrica od $C \otimes B$:

$$C \odot B = \begin{pmatrix} c_{11}b_{11} & c_{12}b_{12} \\ c_{11}b_{21} & c_{12}b_{22} \\ c_{21}b_{11} & c_{22}b_{12} \\ c_{21}b_{21} & c_{22}b_{22} \end{pmatrix} = \begin{pmatrix} c_1 \otimes b_1 & c_2 \otimes b_2 \end{pmatrix}$$

Sljedeći korak: polja s 3 i više indeksa



Stranica i pokazuje na stranicu j u kontekstu pojma k.
Osoba i je povezana s osobom j u kontekstu pojma k.
Osoba i je telefonirala/komunicirala s osobom j i koristila pojam k.

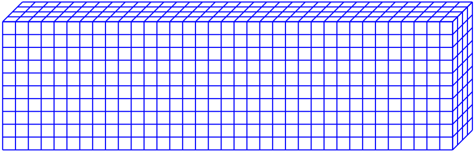
....

Kako poopćiti pojmove izvora, autoriteta? Kako dodijeliti važnost pojedinom kontekstu (pojmu) gledajući cjelinu?

Tamara Kolda, Brett Bader, Joseph Kenny: Higher-order web link analysis using multilinear algebra. Sandia Tech Report.
Tamara Kolda, Brett Bader: The TOPHITS Model for Higher-Order Web Link Analysis . Sandia Tech Report.

Vektorski model podataka

- Digitalna slika. $S = S(i, j) \in \mathbb{R}^{m\times n}$, $S = S(i, j, k) \in \mathbb{R}^{m\times n\times 3}$
- Digitalni video je niz slika (frames)



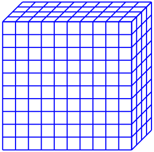
$\longrightarrow$ vrijeme s korakom Δt

Primjene: kompresija, otklanjanje šuma, prepoznavanje (lica, objekata, znamenki, akcije)

Za svaku od navedenih primjena treba razviti matematički model, metode/algoritme, te numeričku/softversku implementaciju. Koriste se teorije i metode linearne i multilinearne algebre, numeričke matematike, statistike, razvoja softvera.

Tenzori reda 3 - osnovni pojmovi

U prethodnom pregledu smo prošli put od skalara, preko vektora do matrica. Vektor je bila struktura u kojoj operiramo s jednim indeksom $x = (x_1, \ldots, x_n)$ i kažemo da je vektor tenzor reda jedan. Matrica je bila dvodimenzionalno polje u kojem strukturu određuju va indeksa $A = (a_{ij})_{i,j=1}^{m,n}$ - kažemo da je matrica tenzor reda dva.
Struktura oblika



je određena s tri indeksa, $\mathcal{A} = (a_{ijk}) \in \mathbb{R}^{I \times J \times K} \equiv \mathbb{R}^{I_1 \times I_2 \times I_3}$. Kažemo da je $\mathcal{A}$ tenzor reda tri.

Ako je polje $\mathcal{X}$ određeno s N indeksa onda je $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ tenzor reda N nad $\mathbb{R}$, $\mathcal{X} = (x_{i_1 i_2 \ldots i_N})$, $i_1 = 1, \ldots, I_1; \ldots; i_N = 1, \ldots, I_N$.

Primijetimo kako s redom tenzora eksponencijalno raste volumen podataka koje on sadrži. Problem prokletstva dimenzije (curse of dimensionality)!

Anatomija tenzora

Prisjetimo se anatomije matrice $A = (a_{ij})$. Indeks retka i zovemo 1. indeks, a indeks stupca j je 2. indeks.

j-ti stupac matrice A, $A(:, j)$, je

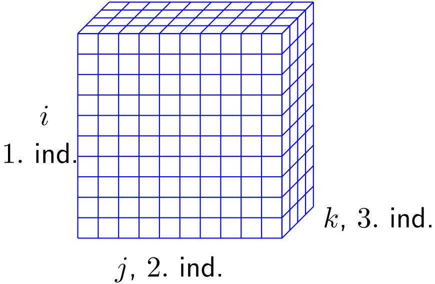
$$A(:, j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad \leftarrow \text{fiksirani svi indeksi osim prvog}$$

i redak matrice A, $A(i, :)$, je

$$A(i, :) = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \quad \leftarrow \text{fiksirani svi indeksi osim drugog}$$

Cilj je pojmove redaka i stupaca poopćiti na tenzore bilo kojeg reda. Mi ćemo zbog jednostavnosti prvo raditi na tenzorima reda tri, a nakon toga općenito.

Anatomija tenzora



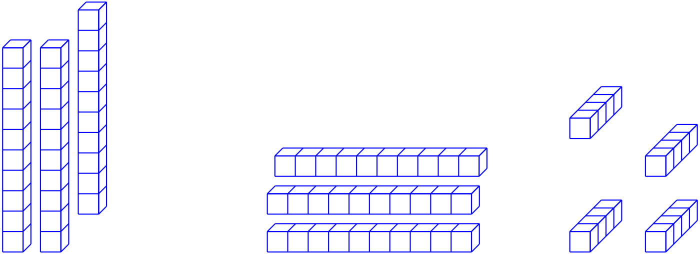
Ako fiksiramo sve indekse osim ℓ-tog onda taj dio tenzora zovemo nit u modu ℓ, $\ell \in \{1, 2, 3\}$. (engl. fiber)

Kod matrica (tenzori reda 2) su

- stupac = nit u modu 1.
- redak = nit u modu 2.

Pogledajmo kako izgledaju niti u svim modovima tenzora reda tri.

Anatomija tenzora



$\mathcal{A}(:, j, k)$
niti u modu 1

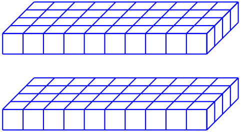
$\mathcal{A}(i, :, k)$
niti u modu 2

$\mathcal{A}(i, j, :)$
niti u modu 3

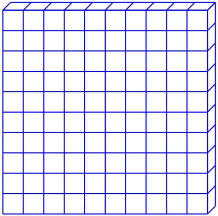
Niti (koje su tenzori reda 1) po dogovoru reprezentiramo kao vektore stupce i sav račun (formule) rade s tom pretpostavkom.

Odsječak (slice)

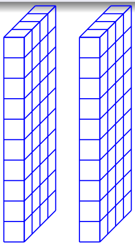
Odsječak (slice) tenzora je 2d-sekcija tenzora definirana fiksiranjem svih osim dva indeksa.



$\mathcal{A}(i,:,:)$
horizontalni slice



$\mathcal{A}(:,:,k)$
frontalni slice



$\mathcal{A}(:,j,:)$
bočni (lateralni) slice

Osnovna svojstva i operacije

Ako je tenzor $\mathcal{X}$ nad $\mathbb{R}$ i ima N modova, pišemo $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, a indekse označavamo s $i_1, i_2, \ldots, i_N$.

Euklidski skalarni produkt i norma

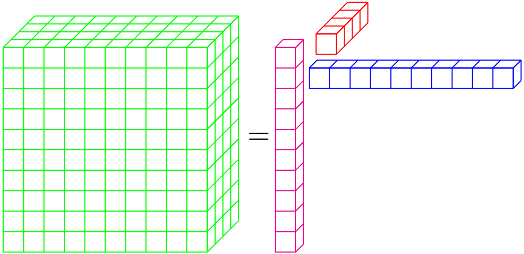
- $\langle \mathcal{X}, \mathcal{Y} \rangle_F = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} x_{i_1 i_2 \ldots i_N} y_{i_1 i_2 \ldots i_N}$
- $\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle_F} = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} x_{i_1 i_2 \ldots i_N}^2}$

Neke specijalne klase tenzora

- Kažemo da je tenzor kockast (cubical) ko su mu svi modovi jednake duljine, $\mathcal{X} \in \mathbb{R}^{I \times I \times \cdots \times I}$.
- Kockasti tenzor je super-simetričan ako se vrijednost elementa ne mijenja permutacijom njegovih indeksa, npr. kod tenzora reda 3 $x_{ijk} = x_{ikj} = x_{jik} = x_{jki} = x_{kij} = x_{kji}$
- Tenzor može biti parcijalno simetričan u nekim modovima, npr. $\mathcal{X}(:,:,k) = \mathcal{X}(:,:,k)^T$, $k = 1, \ldots, K$.
- $\mathcal{X}$ je dijagonalan ako $x_{i_1 i_2 \ldots i_N} \neq 0 \Longrightarrow i_1 = i_2 = \ldots = i_N$.

Tenzor ranga jedan

$$\mathcal{A} = a^{(1)} \circ a^{(2)} \circ \cdots \circ a^{(N)}, \quad a_{i_1 i_2 \ldots i_N} = a^{(1)}_{i_1} a^{(2)}_{i_2} \cdots a^{(N)}_{i_N}$$

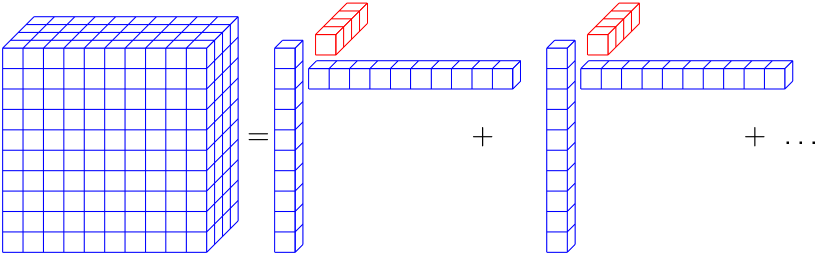


Uoči odnos potrebne memorije za $\mathcal{A}$ i $a^{(1)}, \ldots, a^{(N)}$.

Aproksimacija pomoću tenzora ranga jedan

CP/PARAFAC dekompozicija (esencijalno jedinstvena)

$$\mathcal{A} \approx \sum_{\ell=1}^{L} \sigma_\ell u_\ell \circ v_\ell \circ w_\ell$$



Možemo uzeti $\|u_\ell\| = \|v_\ell\| = \|c_\ell\| = 1$, i sva skaliranja akumulirati u σ_ℓ.
O tome više kasnije; uglavnom kroz seminare i u drugom semestru.
Sada nastavljamo s uvodnim elementarnim pojmovima.

Matricizacija tenzora

Sjetimo se kako smo $m \times n$ matricu operatorom vec razvukli u vektor.

$$A = (a_1, \dots, a_n) = \begin{pmatrix} \widetilde{a}_1 \\ \vdots \\ \widetilde{a}_m \end{pmatrix}. \quad \mathrm{vec}_1(A) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}; \quad \mathrm{vec}_2(A) = \begin{pmatrix} \widetilde{a}_1^T \\ \vdots \\ \widetilde{a}_m^T \end{pmatrix}$$

To možemo napraviti po stupcima i po retcima, tj. u modovima 1 i 2.

Matricizacija tenzora $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$

Matricizacija tenzora u n-tom modu, $\mathcal{X}_{(n)}$ se dobije tako da se niti u n-tom modu poslože kao stupci matrice u nekom unaprijed zadanom fiksnom poretku.

Matlab: proučiti funkcije `reshape`, `squeeze`.

Matricizacija tenzora

Primjer

$\mathcal{X} \in \mathbb{R}^{3\times4\times2}$ je zadan frontalnim odsječcima

$$X_1 = \begin{pmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 13 & 16 & 19 & 22 \\ 14 & 17 & 20 & 23 \\ 15 & 18 & 21 & 24 \end{pmatrix}$$

Tada su matricizacije u sva tri moda dane s

$$\begin{aligned} X_{(1)} &= \begin{pmatrix} 1 & 4 & 7 & 10 & 13 & 16 & 19 & 22 \\ 2 & 5 & 8 & 11 & 14 & 17 & 20 & 23 \\ 3 & 6 & 9 & 12 & 15 & 18 & 21 & 24 \end{pmatrix} \\ X_{(2)} &= \begin{pmatrix} 1 & 2 & 3 & 13 & 14 & 15 \\ 4 & 5 & 6 & 16 & 17 & 18 \\ 7 & 8 & 9 & 19 & 20 & 21 \\ 10 & 11 & 12 & 22 & 23 & 24 \end{pmatrix} \\ X_{(3)} &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \end{pmatrix} \end{aligned}$$

Primjer (vježba) – operatori $\texttt{unfold}$, $\texttt{fold} = \texttt{unfold}^{-1}$

Primjer ($\texttt{unfold}$) za $\mathcal{A} \in \mathbb{R}^{\ell\times m\times n}$

$$\mathcal{A}_{(1)} = \underbrace{\texttt{unfold}_1(\mathcal{A})}_{\ell\times(m\cdot n)} = \underbrace{(\mathcal{A}(:,1,:), \mathcal{A}(:,2,:), \ldots, \mathcal{A}(:,m,:))}_{\text{niti u modu 1 uzete redom po lateralnim sliceovima}}$$

$$\mathcal{A}_{(2)} = \underbrace{\texttt{unfold}_2(\mathcal{A})}_{m\times(\ell\cdot n)} = \underbrace{(\mathcal{A}(:,:,1)^T, \mathcal{A}(:,:,2)^T, \ldots, \mathcal{A}(:,:,n)^T)}_{\text{transponirani frontalni sliceovi, redom}}$$

$$\mathcal{A}_{(3)} = \underbrace{\texttt{unfold}_3(\mathcal{A})}_{n\times(m\cdot\ell)} = \underbrace{(\mathcal{A}(1,:,:)^T, \mathcal{A}(2,:,:)^T, \ldots, \mathcal{A}(\ell,:,:)^T)}_{\text{transponirani horizontalni sliceovi}}$$

Primjer: kod matrica s $\ell \times m \equiv \ell \times m \times 1$

- $\texttt{unfold}_1(A) = A$
- $\texttt{unfold}_2(A) = A^T$
- $\texttt{unfold}_3(A) = \texttt{vec}_1(A^T)^T$

$\texttt{fold}$ je inv. operacija od $\texttt{unfold}$ i treba deskriptor (dimenzije tenzora).

Množenje u modu matrice tenzorom

Neka je dan tenzor $\mathcal{T} \in \mathbb{R}^{J_1 \times \cdots \times J_{n-1} \times J_n \times J_{n+1} \times \cdots \times J_N}$, $\mathcal{T} = (t_{j_1 \ldots j_N})$ i neka je zadana matrica $A \in \mathbb{R}^{I \times J_n}$.

Množenje u modu n: $\mathcal{T} \times_n A$

Produkt $\mathcal{T} \times_n A \in \mathbb{R}^{J_1 \times \cdots \times J_{n-1} \times I \times J_{n+1} \times \cdots \times J_N}$ je zadan formulama

$$\underbrace{(\mathcal{T} \times_n A)_{j_1 \ldots j_{n-1} i j_{n+1} \ldots j_N}}_{\text{element niti u modu } n} = \sum_{k=1}^{J_n} a_{ik} \underbrace{t_{j_1 \ldots j_{n-1} k j_{n+1} \ldots j_N}}_{\star}, \quad i = 1, \ldots, I$$

$\star = v_k =$ fiksirani svi ind. osim n-tog=nit u modu n; $\sum_{k=1}^{J_n} a_{ik} v_k$

Množenje u modu n: $\mathcal{T} \times_n A$

$$(\mathcal{T} \times_n A)_{(n)} = A\mathcal{T}_{(n)}$$

Koristeći $\mathrm{vec}(AX) = (I \otimes A)\mathrm{vec}(\mathrm{X})$, $\mathrm{vec}(XB) = (B^T \otimes I)\mathrm{vec}(X)$ imamo

$$\mathrm{vec}(A\mathcal{T}_{(n)}) = (I \otimes A)\mathrm{vec}(\mathcal{T}_{(n)})$$

Primjer

Kod tenzora reda 3, za $\mathcal{T} \in \mathbb{R}^{I \times J \times K}$ i $A \in \mathbb{R}^{L \times J}$ imamo npr. da je $\mathcal{X} = \mathcal{T} \times_2 A$ dan s

$$\mathcal{X}(i,:,k) = A\mathcal{T}(i,:,k), \quad i = 1,\dots,I; k = 1,\dots,K.$$

$\mathcal{T} \in \mathbb{R}^{3 \times 4 \times 2}$, $A \in \mathbb{R}^{2 \times 3}$

$$\mathcal{T}(:,:,1) = \begin{pmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{pmatrix}, \quad \mathcal{T}(:,:,2) = \begin{pmatrix} 13 & 16 & 19 & 22 \\ 14 & 17 & 20 & 23 \\ 15 & 18 & 21 & 24 \end{pmatrix}; \quad A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

$$(\mathcal{T} \times_1 A)(:,:,1) = \begin{pmatrix} 22 & 49 & 76 & 103 \\ 28 & 64 & 100 & 136 \end{pmatrix}, \quad (\mathcal{T} \times_1 A)(:,:,2) = \begin{pmatrix} 130 & 157 & 184 & 211 \\ 172 & 208 & 244 & 280 \end{pmatrix}$$

$\mathcal{T} \times_i M = \mathtt{fold}_i(M \cdot \mathtt{unfold}_i(\mathcal{T}), \mathtt{deskriptor})$, tj.
$\mathcal{T} \times_i M = \mathtt{fold}_i(M \cdot \mathcal{T}_{(i)}, \mathtt{deskriptor})$

Osnovna svojstva

Teorem

Neka je $\mathcal{T} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$. Tada vrijedi:

- Za $A \in \mathbb{R}^{I_m \times J_m}$, $B \in \mathbb{R}^{I_n \times J_n}$ i $m \neq n$ je

$$\mathcal{T} \times_m A \times_n B = (\mathcal{T} \times_m A) \times_n B = (\mathcal{T} \times_n B) \times_m A$$

- Za $A \in \mathbb{R}^{I \times J_n}$, $B \in \mathbb{R}^{K \times I}$ je

$$(\mathcal{T} \times_n A) \times_n B = \mathcal{T} \times_n (BA)$$

- Neka je $A \in \mathbb{R}^{I \times J_n}$ punog ranga. Tada

$$\mathcal{X} = \mathcal{T} \times_n A \Longrightarrow \mathcal{T} = \mathcal{X} \times_n A^{\dagger}.$$

  Ako je A ortonormalna $(A^T A = I)$ onda gornje vrijedi s $A^{\dagger} = A^T$.

Skica dokaza za drugu tvrdnju:

$$((\mathcal{T} \times_n A) \times_n B)_{(n)} = B(\mathcal{T} \times_n A)_{(n)} = B(A\mathcal{T}_{(n)}) = (BA)\mathcal{T}_{(n)} = (\mathcal{T} \times_n (BA))_{(n)}.$$

Tenzorski SVD i kompresija tenzora

Prisjetimo se SVD dekompozicije matrica i njene uloge u kompresiji (količine podataka u matrici).

$B \in \mathbb{R}^{m\times n}$, $m \geq n$; $B = U\Sigma V^T$, $U^T U = I_m$, $V^T V = I_n$,
$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$; $\sigma_1 \geq \cdots \geq \sigma_r \gg \sigma_{r+1} = \ldots = \sigma_n = 0$

$$\begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} = \begin{pmatrix} + & + & + & + & + & + & + \\ + & + & + & + & + & + & + \\ + & + & + & + & + & + & + \\ + & + & + & + & + & + & + \\ + & + & + & + & + & + & + \\ + & + & + & + & + & + & + \\ + & + & + & + & + & + & + \end{pmatrix} \begin{pmatrix} \bullet & & & & \\ & \bullet & & & \\ & & \bullet & & \\ & & & \bullet & \\ & & & & \bullet \end{pmatrix} \begin{pmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{pmatrix}$$

$B_k = U(:, 1:k)\Sigma(1:k, 1:k)V(:, 1:k)^T$; $\|B - B_k\|_2 = \sigma_{k+1}$

$$\begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}_{m\times n} \approx \begin{pmatrix} + & + \\ + & + \\ + & + \\ + & + \\ + & + \\ + & + \\ + & + \end{pmatrix}_{m\times k} \begin{pmatrix} \bullet & \\ & \bullet \end{pmatrix}_{k\times k} \begin{pmatrix} x & x & x & x & x \\ x & x & x & x & x \end{pmatrix}_{k\times n}$$

Tenzorski SVD i kompresija tenzora

$B = U\Sigma V^T$, $U = (u_1, \ldots, u_m)$, $V = (v_1, \ldots, v_n)$; $B = \sum_i \sigma_i u_i v_i^T$

U Frobenijusovom skalarnom produktu $\langle X, Y\rangle_F = \mathrm{Trag}(Y^T X)$ vrijedi

$$\langle u_i v_i^T, u_j v_j^T\rangle_F = \mathrm{Trag}(v_j u_j^T u_i v_i^T) = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

Matricizacije $\ell \times m \equiv \ell \times m \times 1$

$\mathtt{unfold}_1(M) = M$; $\mathtt{unfold}_2(M) = M^T$; $\mathtt{unfold}_3(M) = \mathtt{vec}_1(M^T)^T$
$\mathcal{T} \times_2 M = \mathtt{fold}_2(M\mathcal{T}_{(2)})$

$\Sigma \times_2 V = (V\Sigma^T)^T = \Sigma V^T$; $\Sigma \times_1 U = U\Sigma$

$$B = U\Sigma V^T = \Sigma \times_1 U \times_2 V; \;\; \Sigma = U^T B V = B \times_1 U^T \times_2 V^T$$

$\Sigma = \mathrm{diag}(\sigma_i)$, $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$ :: stupci međusobno okomiti; retci međusobno okomiti; norme stupaca padaju $\rightarrow$; norme redaka padaju $\downarrow$

HOSVD - tenzorski SVD

Teorem

Tenzor $\mathcal{A} \in \mathbb{R}^{\ell \times m \times n}$ možemo zapisati kao

$$\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$$

gdje su $U^{(1)} \in \mathbb{R}^{\ell \times \ell}$, $U^{(2)} \in \mathbb{R}^{m \times m}$, $U^{(3)} \in \mathbb{R}^{n \times n}$ ortogonalne matrice, $\mathcal{S}$ je tenzor istih dimenzija kao i $\mathcal{A}$ i ima svojstvo potpune ortogonalnosti: svaka dva različita odsječka istog tipa su okomiti: za sve $i \neq j$ vrijedi

$$\langle \mathcal{S}(i,:,:), \mathcal{S}(j,:,:) \rangle_F = \langle \mathcal{S}(:,i,:), \mathcal{S}(:,j,:) \rangle_F = \langle \mathcal{S}(:,:,i), \mathcal{S}(:,:,j) \rangle_F = 0.$$

Singularne vrijednosti u modu 1 su

$$\sigma_j^{(1)} = \|\mathcal{S}(j,:,:)\|_F, \quad j = 1, \ldots, \ell$$

i uređene su tako da je $\sigma_1^{(1)} \geq \sigma_2^{(1)} \geq \cdots \geq \sigma_\ell^{(1)} \geq 0$. Ako je $\ell > mn$, onda je $\mathcal{S}(i,:,:) = \mathbf{0}$ za $i > mn$.
Analogno se definiraju singularne vrijednosti u ostalim modovima.

HOSVD

Vrijedi i da je

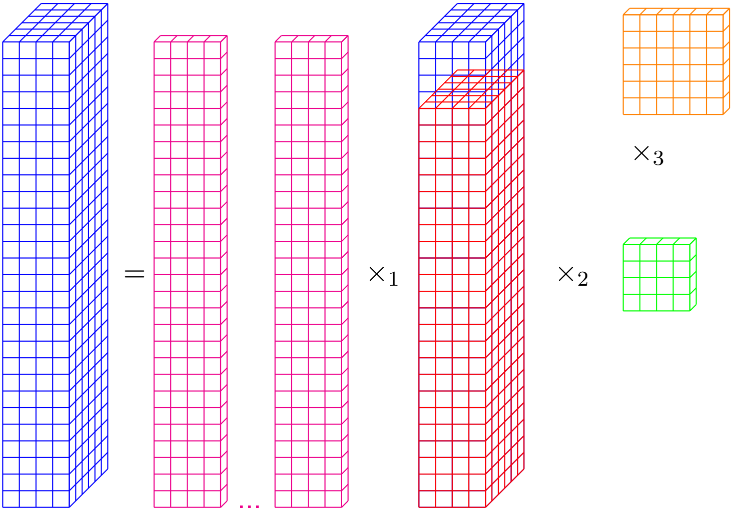
$$\mathcal{S} = \mathcal{A} \times_1 (U^{(1)})^T \times_2 (U^{(2)})^T \times_3 (U^{(3)})^T.$$

Program za računanje HOSVD:

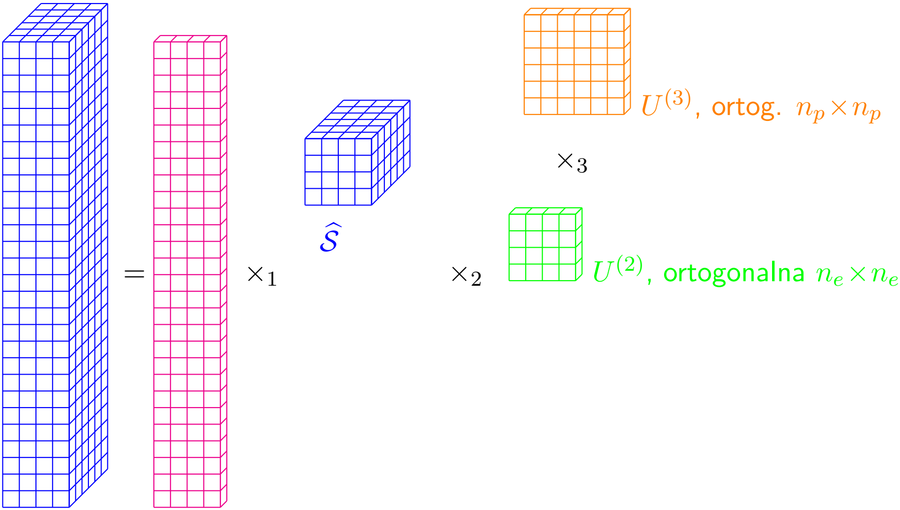
1. `[U1,S1,V1] = svd(unfold(A,1))`
   `([U1,~,~ ] = svd(unfold(A,1)))`
2. `[U2,S2,V2] = svd(unfold(A,2))`
3. `[U3,S3,V3] = svd(unfold(A,1))`
4. `S=ten_mat_m(ten_mat_m(ten_mat_m(A,U1',1),U2',2),U3',3)`

- `svd` računa matrični SVD
- `ten_mat_m` je funkcija za množenje tenzora i matrice u zadanom modu

HOSVD



$$\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \ (\ell > mn,\ \mathcal{S}(i,:,:) = \mathbf{0} \text{ za } i > mn)$$



$$\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} = \widehat{\mathcal{S}} \times_1 \widehat{U^{(1)}} \times_2 U^{(2)} \times_3 U^{(3)}$$

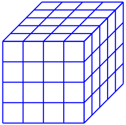
HOSVD – $\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$ – anatomija

Opet motivaciju tražimo u matričnom SVD-u i pokušavamo poopćiti.

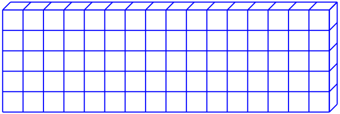
$$M = U\Sigma V^T = \Sigma \times_1 U \times_2 V = \sum_{i=1}^{n} \sigma_i \underbrace{\begin{pmatrix} + \\ + \\ + \end{pmatrix}}_{u_i} \underbrace{\begin{pmatrix} x & x & x \end{pmatrix}}_{v_i^T}; \ \langle u_i v_i^T, u_j v_j^T \rangle_F = 0.$$

$$\begin{aligned}
\mathcal{A} &= \overbrace{(\mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)})}^{\mathcal{B}} \times_3 U^{(3)} \\
&= \mathtt{fold}_3(\underbrace{U^{(3)}}_{n\times n} \underbrace{\mathtt{unfold}_3(\mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)})}_{n\times \ell\cdot m}) = \mathtt{fold}_3(\underbrace{U^{(3)}}_{n\times n} \underbrace{\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}}_{n\times \ell\cdot m}) \\
&= \mathtt{fold}_3(\underbrace{\begin{pmatrix} u_1^{(3)} & \dots & u_n^{(3)} \end{pmatrix}}_{n\times n} \underbrace{\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}}_{n\times \ell\cdot m}) = \mathtt{fold}_3(\sum_{i=1}^{n} u_i^{(3)} r_i) = \sum_{i=1}^{n} \mathtt{fold}_3(u_i^{(3)} r_i)
\end{aligned}$$

Ovdje je $\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$ retčana particija matrice čiji stupci su niti u modu 3 tenzora $\mathcal{B}$. Pogledajmo koji su to elementi u retku r_i.



$\mathcal{B}$, $\ell \times m \times n$

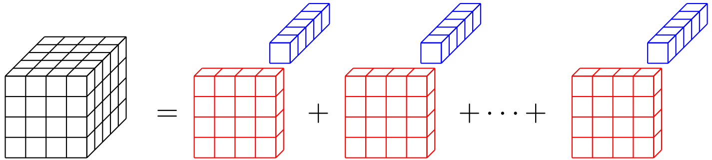


$\mathtt{unfold}_3(\mathcal{B})$, $n \times (\ell \cdot m)$

r_i (i-ti redak u $\mathtt{unfold}_3(\mathcal{B})$) je $1 \times (\ell \cdot m)$ i njemu su elementi od $\mathcal{B}(:,:,i)$, tj. $r_i = \mathtt{unfold}_3(\mathcal{B}(:,:,i))$. Sada nastavljamo s formulama s prethodne stranice:

$$\begin{aligned}
\mathcal{A} &= \sum_{i=1}^{n} \mathtt{fold}_3(u_i^{(3)} r_i) = \sum_{i=1}^{n} \mathtt{fold}_3(u_i^{(3)} \mathtt{unfold}_3(\mathcal{B}(:,:,i))) \\
&= \sum_{i=1}^{n} \mathcal{B}(:,:,i) \times_3 u_i^{(3)} = \sum_{i=1}^{n} \underbrace{(\mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)})}_{(\mathcal{S} \times_2 U^{(2)}) \times_1 U^{(1)}}(:,:,i) \times_3 u^{(i)} \\
&= \sum_{i=1}^{n} {\color{red}(\mathcal{S}(:,:,i) \times_1 U^{(1)} \times_2 U^{(2)})} \times_3 u_i^{(3)} = \sum_{i=1}^{n} {\color{red}\mathbb{A}_i} \times_3 u_i^{(3)}
\end{aligned}$$

HOSVD – $\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$ – anatomija

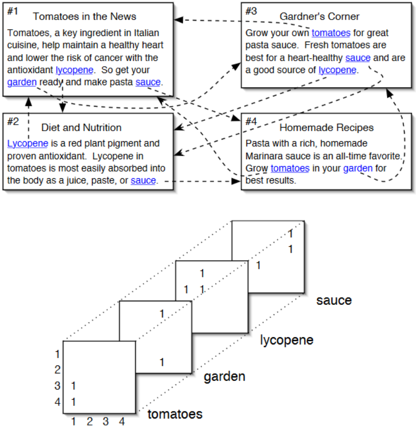


$$\mathcal{A} = \sum_{i=1}^{n} \mathbb{A}_i \times_3 u_i^{(3)}; \;\; \mathcal{A}(:,:,j) = \sum_{i=1}^{n} \mathbb{A}_i (u_i^{(3)})_j, \; j = 1, \ldots, n.$$

$\mathbb{A}_i = (\mathcal{S}(:,:,i) \times_1 U^{(1)} \times_2 U^{(2)}) = U^{(1)} \mathcal{S}(:,:,i) (U^{(2)})^T$ pa je

$$\begin{aligned} \langle \mathbb{A}_i, \mathbb{A}_j \rangle_F = \mathrm{Trag}(\mathbb{A}_j^T \mathbb{A}_i) &= \mathrm{Trag}(U^{(2)} \mathcal{S}(:,:,j)^T (U^{(1)})^T U^{(1)} \mathcal{S}(:,:,i) (U^{(2)})^T) \\ &= \mathrm{Trag}(\mathcal{S}(:,:,j)^T \mathcal{S}(:,:,i)) = 0 \quad \text{za } i \neq j \text{ (iz def. HOSVD)} \end{aligned}$$

Prisjetimo se prvog primjera ...



Stranica i pokazuje na stranicu j u kontekstu pojma k.
Osoba i je povezana s osobom j u kontekstu pojma k.
Osoba i je telefonirala/komunicirala s osobom j i koristila pojam k.
....
Kako poopćiti pojmove izvora, autoriteta? Kako dodijeliti važnost pojedinom kontekstu (pojmu) gledajući cjelinu?

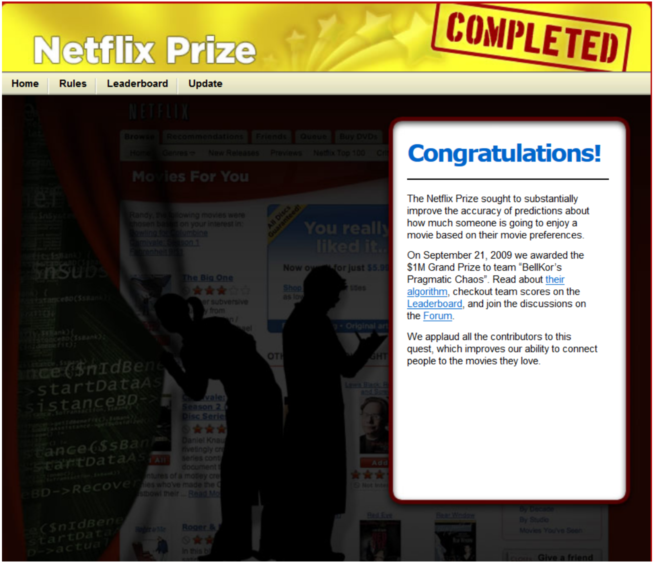
Tamara Kolda, Brett Bader, Joseph Kenny: Higher-order web link analysis using multilinear algebra. Sandia Tech Report.
Tamara Kolda, Brett Bader: The TOPHITS Model for Higher-Order Web Link Analysis . Sandia Tech Report.

Primjer: Analiza i predviđanje linkova

Imamo neusmjeren graf G koji reprezentira povezanosti unutar grupe od n objekata/osoba, koji je zadan matricom susjedstva A: $A_{ij} = 1$ ako postoji brid između vrhova i, j; inače $A_{ij} = 0$. G može biti i bipartitni graf koji sadrži npr. preferencije m korisnika pri izboru n artikala/filmova.
Sada pretpostavimoi da se G dinamički mijenja, $G = G_t$, $A = A_t$, $t = 1, 2, \ldots, T$. Da li možemo u trenutku T prognozirati stanje povezanosti za $T+1, T+2, \ldots, T+L$, tj. predvidjeti najvjerojatnije veze u budućnosti? Da li možemo u nekom momentu t predvidjeti/sugerirati postojanje linka koji nije registriran u grafu?
Primjene

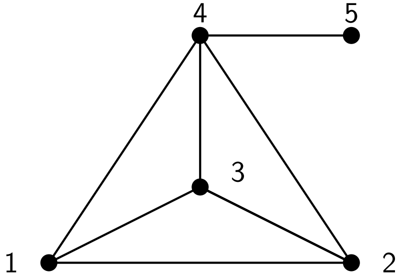
- Sugeriranje prijateljstva na društvenim mrežama
- Širenje zaraznih bolesti . Širenje (dez)informacija, online reklama.
- Data mining u znanstvenim istraživanjima (npr. biologija, genetika)
- Kriminalističke i obavještajne analize i operacije
- Sugeriranje filma/Netflix (recommender sys., collaborative filtering). Primjer: https://www.kaggle.com/tags/recommender-systems

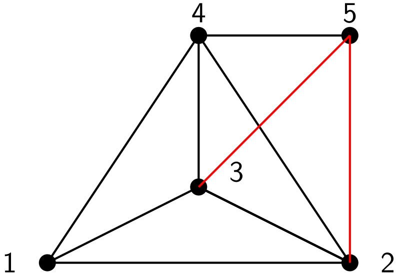
Primjer: Analiza i predviđanje linkova. Cijena? Sitnica.



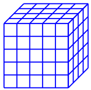
https://www.netflixprize.com/

Primjer: Analiza i predviđanje linkova





$$A_1 = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix},$$



$A_1, A_2, \ldots, A_T$ kao frontalni odsječci tenzora $\mathbb{T}$, $\mathbb{T}(:,:,t) = A_t$.
Matricizacija: $A = \sum_{t=1}^{T} A_t$ ili $A = \sum_{t=1}^{T} (1-\theta)^{T-t} A_t$. Gubi se t-dimenzija! Tenzor reda 3 prirodna struktura.

Primjer: pretraga i prepoznavanje znakova/znamenki

Zadana je kolekcija rukom pisanih znamenki, skenirane i spremljene kao $\ell \times m$ slike (bitmape). Za svaku konkretnu znamenku imamo n uzoraka, tako da svakoj znamenci pripada $\ell \times m \times n$ tenzor. (Zbog jednostavnosti oznaka je n isti za sve znamenke.) Neka je $\mathcal{A}_k$ tenzor k-te znamenke. Kako smo upravo naučili, možemo pisati

$$\mathcal{A}_k = \sum_{i=1}^{n} \mathbb{A}_i^{(k)} \times_3 u_i^{(3,k)}, \quad k = 0, \ldots, 9.$$

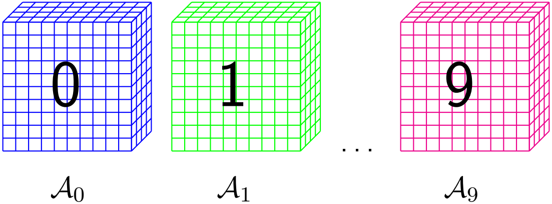
Ako trebamo klasificirati (prepoznati) znamenku danu kao $Z \in \mathbb{R}^{\ell \times m}$, pokušavamo je aproksimirati u smislu najmanjih kvadrata

$$\underbrace{\min_{z_i^{(k)}} \|Z - \sum_i z_i^{(k)} \mathbb{A}_i^{(k)}\|_F}_{\epsilon_k}, \quad k = 0, \ldots, 9$$

i onda klasificiramo kao $\arg\min_k \epsilon_k$

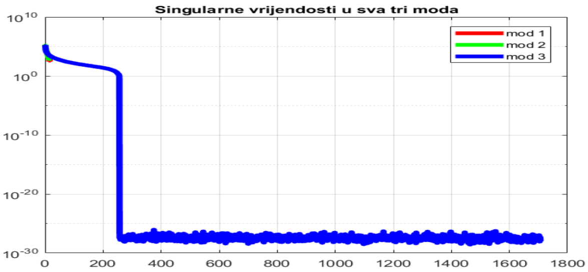
Primjer: pretraga i prepoznavanje znakova/znamenki

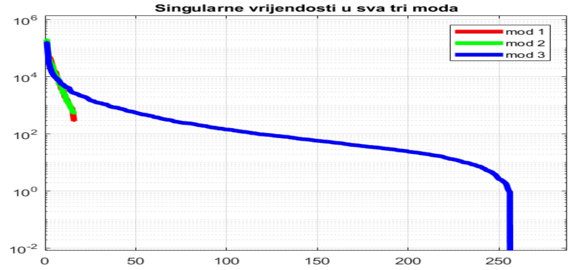
Primjer iz [1] Lars Eldén: *Matrix Methods in Data Mining and Pattern Recognition*, SIAM 2007; [2] B. Savas, L. Eldén: *Handwritten digit classification using higher order singular value decomposition*, Pattern Recognition Volume 40, Issue 3, March 2007, 993-1003.



U našem primjeru su svi tenzori dimenzije $16 \times 16 \times 1707$, što znači da je netrivijalni dio HOSVDa u modu 3 dimenzije najviše 256.
Sada ćemo formirati te tenzore, izračunati HOSVD i pogledati izračunate singularne vrijednosti u svim modovima.

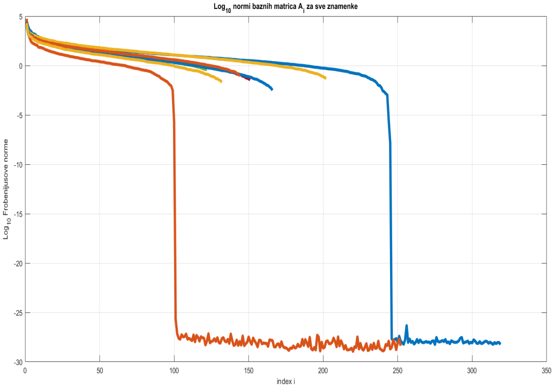
Singularne vrijednosti (HOSVD) kolekcije svih znamenki





Koliko su norme od $\mathbb{A}_i$?

Kako je i bilo za očekivati, za svaku znamenku k, Frobenijusove norme od $\mathbb{A}_i^{(k)} = (\mathcal{S}^{(k)}(:,:,i) \times_1 U^{(k,1)} \times_2 U^{(k,2)}) = U_k^{(1)} \mathcal{S}^{(k)}(:,:,i)(U^{(k,2)})^T$ padaju s rastućim indeksom i.
U našem primjeru, $\log_{10} \|\mathbb{A}_i^{(k)}\|_F$, $k = 0, \ldots, 9$, izgledaju ovako



Primjer: pretraga i prepoznavanje znakova/znamenki

Kako minimiziramo? Klasika: Stavimo

$$\begin{aligned}
F(z_1^{(k)}, ..., z_n^{(k)}) &= \frac{1}{2}\|Z - \sum_i z_i^{(k)} \mathbb{A}_i^{(k)}\|_F^2 \\
&= \frac{1}{2}\langle Z - \sum_i z_i^{(k)} \mathbb{A}_i^{(k)}, Z - \sum_j z_j^{(k)} \mathbb{A}_j^{(k)} \rangle_F \\
&= \frac{1}{2}\|Z\|_F^2 + \frac{1}{2}\sum_i (z_i^{(k)})^2 \|\mathbb{A}_i\|_F^2 - \sum_i z_i^{(k)} \langle Z, \mathbb{A}_i \rangle_F \\
\frac{\partial F(z_1^{(k)}, ..., z_n^{(k)})}{\partial z_j^{(k)}} &= -\langle Z, \mathbb{A}_j \rangle_F + z_j^{(k)} \|\mathbb{A}_j\|_F^2 = 0, \ \text{ tj. } \ z_j^{(k)} = \frac{\langle Z, \mathbb{A}_j \rangle_F}{\|\mathbb{A}_j\|_F^2}.
\end{aligned}$$

To smo zapravo odmah znali iz teorema o projekciji ($\mathbb{A}_1, \ldots, \mathbb{A}_n$ su međusobno okomiti pa su $z_i^{(k)}$ Fourierovi koeficijenti). Dakle, možemo izračunati

$$\epsilon_k = \min_{z_i^{(k)}} \|Z - \sum_i z_i^{(k)} \mathbb{A}_i^{(k)}\|_F, \quad k = 0, \ldots, 9$$

i odabrati k za kojeg je ϵ_k najmanji.

Komentar o kompresiji tenzora

U HOSVD $\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$ znamo da su u jezgrenom tenzoru $\mathcal{S}$ odsječci u svakom modu padajući u Frobenijusovoj normi. Možemo postaviti neki prag tolerancije i sve odsječke ispod tog praga proglasiti nulama. Tako umjesto $\mathcal{S}$ dimenzije $\ell \times m \times n$ dobijemo

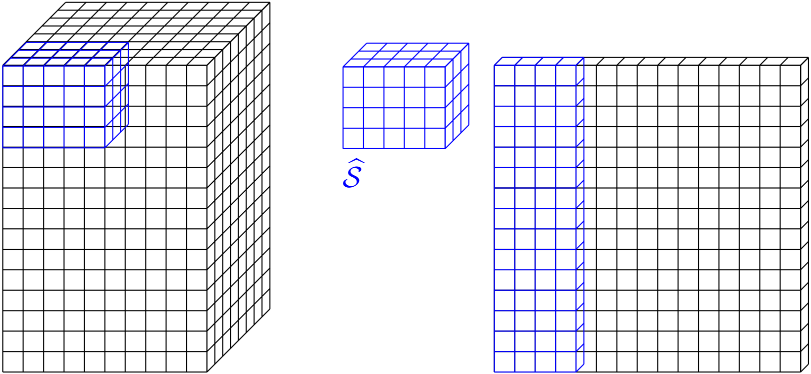
$$\widehat{\mathcal{S}} = \mathcal{S}(1:k_1, 1:k_2, 1:k_3)$$

Ako stavimo $\widehat{U}^{(i)} = U^{(i)}(:, 1:k_i)$, $i = 1, 2, 3$, onda dobijemo aproksimaciju

$$\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \approx \widehat{\mathcal{S}} \times_1 \widehat{U^{(1)}} \times_2 \widehat{U^{(2)}} \times_3 \widehat{U^{(3)}}$$

Jasno je da $\widehat{\mathcal{S}}$ i $\widehat{U}^{(i)} = U^{(i)}(:, 1:k_i)$ zauzimaju puno manje memorije.

Komentar o kompresiji tenzora



$\mathcal{S}, \mathcal{A}$, $\ell \times m \times n$

$\widehat{\mathcal{S}}$

$\widehat{U^{(1)}}$ $\quad U^{(1)} = (\widehat{U^{(1)}}, \square \ldots \square)$

$$\mathcal{A} \approx \widehat{\mathcal{S}} \times_1 \widehat{U^{(1)}} \times_2 \widehat{U^{(2)}} \times_3 \widehat{U^{(3)}}$$

Literatura.

Pročitati (dostupno na internetu; samo "zagooglajte" naslov)

- Tamara G. Kolda, Brett W. Bader: Tensor Decompositions and Applications. SIAM Review 2009.
- Tamara G. Kolda: Multilinear operators for higher-order decompositions.

Baciti pogled na primjene

- Tamara Kolda, Brett Bader, Joseph Kenny: Higher-order web link analysis using multilinear algebra. Sandia Tech Report.
- https://www.kolda.net/

Knjiga i rad o prepoznavanju znamenki

- Lars Eldén: Matrix Methods in Data Mining and Pattern Recognition, SIAM 2007.
- B. Savas, L. Eldén: *Handwritten digit classification using higher order singular value decomposition*, Pattern Recognition Volume 40, Issue 3, March 2007, 993-1003

Vježbe

Vježbe

- Dokazati sve tvrdnje navedene na predavanju. Familijarizirati se s različitim matričnim množenjima i sa strukturom tenzora reda tri.
- Napisati funkcije (Matlab ili neki drugi software) za `unfold`, `fold` i množenje tenzora i matrica u svim modovima.
- U folderu VJEZBE/Klasteriranje_Prepoznavanje_ZNAMENKE su skenirane znamenke (sjetite se ranijih vježbi). Učitajte ih u strukturu tenzora reda tri tako da svaki frontalni odsječak sadrži jednu znamenku. Implementiraje metodu prepoznavanja znamenki, kako smo opisali.