Feb 18, 2020

Last lecture: NMF

$$A \geq 0$$

$$\|A - LR\|_F^2 + \lambda\left(\|L\|_F^2 + \|R\|_F^2\right)$$

s.t. $L, R \geq 0$

separable NMF: $X = WH$, $W = X(:, K)$

$K \subseteq \{1, \ldots, n\}$

scale X to $\bar{X}$: column sums to 1

$$\bar{X} = \bar{X}(:, K)H$$

$$\bar{X}\Pi = \bar{X}(:, K)[I \;\; G] \leftarrow \text{Example of an "interpolative decomposition"}$$

ID: $A\Pi = C[I \;\; T]$

C is a subset of columns of A

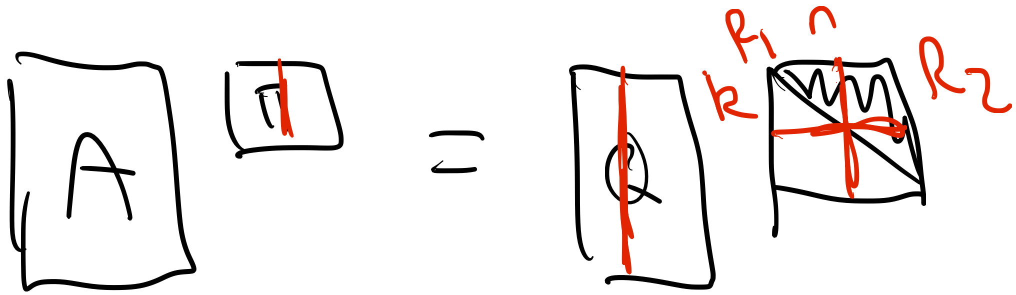
Pivoted QR

- finding representative data points for near-sep NMF (vertices of convex hull)
- LLS $\|Ax-b\|_2$ for feature selection

$$\underset{m\times n}{A} = \underset{m\times n}{Q}\;\underset{n\times n}{R} \qquad Q^T Q = I$$

1. choose col t with largest 2-norm
2. pivot t to front $\quad q_1 = A(:,t)/\|A(:,t)\|_2$
3. Orthog. out $\quad A \leftarrow (I - q_1 q_1^T)A$
4. "Update R" first row: $R(1,:) = q_1^T A$
5. Repeat on $A(:,2:n)$ (next row of R)

At the end $A\Pi = QR$



$$A\Pi_k \approx Q_k [R_1 \ \ R_2]$$

$$= \underbrace{Q_k R_1}_{\text{first } k \text{ cols of } A\Pi} [I \ \ R_1^{-1} R_2]$$

$$A\Pi_{:,1} = q_1 r_{11} = q_1 \|A(:,t)\|_2$$

$$A\Pi_{:,2} = q_1 r_{12} + q_2 r_{22}$$

$$\|A\Pi_{:,2}\|_2^2 = |r_{12}|^2 + |r_{22}|^2 \le \|A\Pi_{:,1}\|^2 = |r_{11}|^2$$

$$|r_{22}|^2 \le |r_{11}|^2 - |r_{12}|^2 \le |r_{11}|^2$$

$|r_{k+1,k+1}|$ small $\Rightarrow$ can truncate at k for an approx

Also, $|r_{12}| \le |r_{11}|$

$A\Pi \approx C[I \;\; T]$ C subset of k columns of A

$A\Pi = C[I \;\; R_1^{-1}R_2]$

More stringent: $|T_{ij}| \leq 2$

PQR: some pathological where
T has large entries

- randomized algs

CUR decomposition

$$\underset{m\times n}{A} = \underset{m\times k}{C}\;\underset{k\times k}{U}\;\underset{k\times n}{R}$$

C: columns of A

R: rows

How to select cols and rows?

Idea 1: ID/PQR on $A \Rightarrow$ cols

$A^T \Rightarrow$ rows

$M^+ = (M^T M)^{-1} \cdot M^T$

$$\min_U \|A - C\underbrace{UR}_{Z}\|_F^2 \qquad \min_Z \|A - CZ\|_F^2 \qquad Z = C^+ A$$

$$\Updownarrow \quad \min_U \|C^+ A - UR\|_F^2 \qquad U = C^+ A R^+$$

(Mahoney+ Drineas 09)

Pre-process: $A \approx U_k \Sigma_k V_k^T$ $O(mnk)$

$\text{ColSelect}(A, k, \varepsilon, c)$ $c = O(k \log k / \varepsilon^2)$

(i) $\pi_j = \frac{1}{k} \| V_k^T(:, j) \|_2^2$ $O(nk)$

(ii) keep j^{th} column w.p. $\min(1, c\pi_j)$ $O(n)$

(iii) return kept cols

CUR: $S = \text{ColSelect}(A, k, \varepsilon, c)$ $O(nk)$ $C = A(:, S)$

$T = \text{ColSelect}(A^T, k, \varepsilon, c)$ $R = A(T, :)$ $O(mk)$

$U = C^+ A R^+$ $O(|T|^2 n)$

$O(|S|^2 m)$

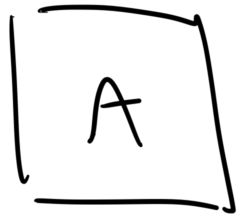
w.h.p. $|T|, |S|$ are $O(c) \geq O(k \log k / \varepsilon^2)$

$O(mnk)$

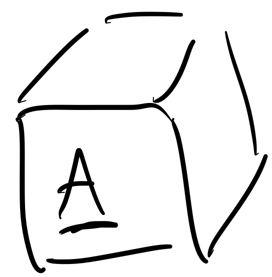
Thm: $\|A - CUR\|_F$

$$\leq (2+\varepsilon) \underbrace{\|A - U_k \Sigma_k V_k^T\|_2}_{\text{OPT}}$$

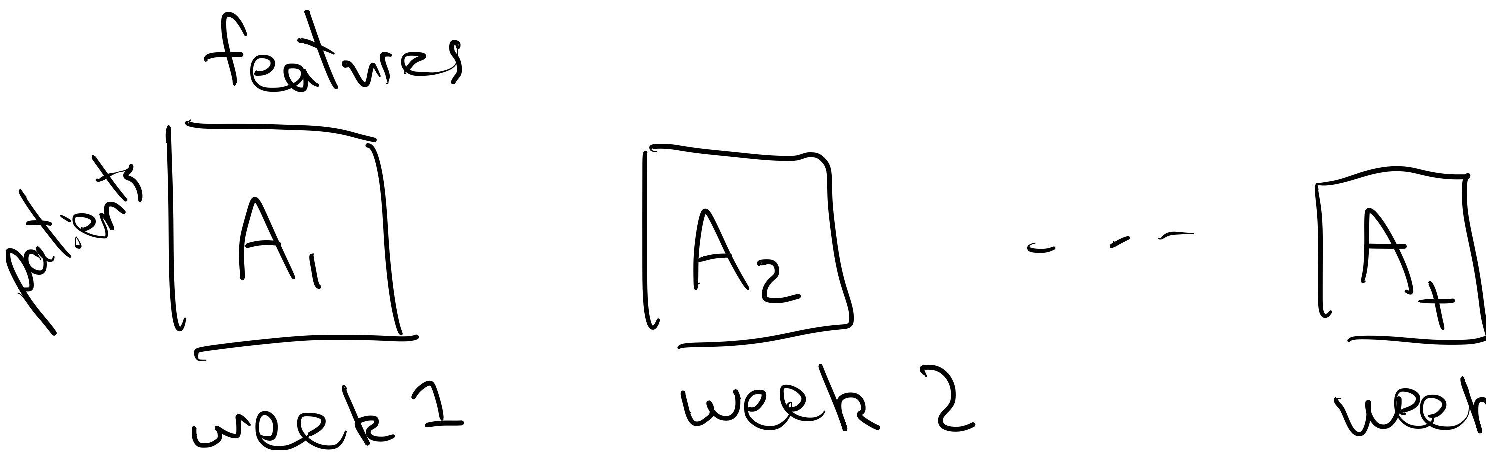
Tensors are a generalization of matrices

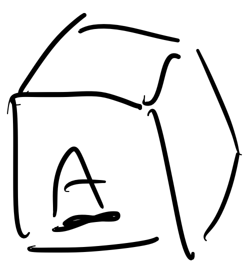


A_{ij}



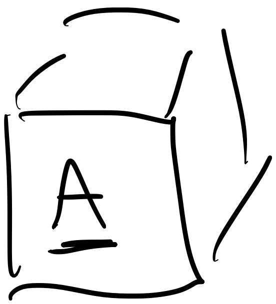
$\underline{A}_{ijk}$





$\underline{A}_{ijk}$ = feature j of patient i in week k

$\boxed{A}$ $\quad A \approx \sum_{s=1}^{r} x_s y_s^T \qquad A_{ij} \approx \sum_{s=1}^{r} (x_s)_i (y_s^T)_j$



$$\underline{A}_{ijk} \approx \sum_{s=1}^{r} (x_s)_i (y_s)_j (z_s)_k$$

$$\underline{A} \approx \sum_{s=1}^{r} x_s \otimes y_s \otimes z_s$$

$z_s = 1$

$$\min_{B} \|A - B\|_F^2 \quad \text{s.t. } B \text{ is rank-}r$$

rank is minimum r s.t. $B = \sum_{s=1}^{r} x_s \otimes y_s$

$$B = U_r \Sigma_r V_r^T$$

Tensor data: "best" low-rank approx

$$\min_{\underline{B}} \|\underline{A} - \underline{B}\|_F^2 \quad \text{s.t. is rank-}r$$

rank is min r s.t. $\underline{B} = \sum_{s=1}^{r} x_s \otimes y_s \otimes z_s$

Problem (de Silva and Lim 08)

Best low-rank approx can be ill-posed!

Might be no best rank-r approx to $\underline{A}$

$$\underline{A}_n = n\left(x_1 + \frac{1}{n}y_1\right) \otimes \left(x_2 + \frac{1}{n}y_2\right) \otimes \left(x_3 + \frac{1}{n}y_3\right) - n\, x_1 \otimes x_2 \otimes x_3$$

rank-2

$$\lim_{n\to\infty} \underline{A}_n = \underline{A}^* = x_1 \otimes x_2 \otimes y_3 + x_1 \otimes y_2 \otimes x_3 + y_1 \otimes x_2 \otimes x_3$$

rank-3

"border rank" = 2

$\underline{A}$ $2 \times 2 \times 2$ is rank 3

Then $\underline{A}$ does not have a best rank-2 approx

$\underline{A}$ $d_1 \times d_2 \times d_3$ $2 \leq s \leq \max(d_1, d_2, d_3)$

$\{ \underline{A} \mid \underline{A}$ do not have best rank-r approx, $r \leq s \}$

has positive volume