Feb 11, 2020

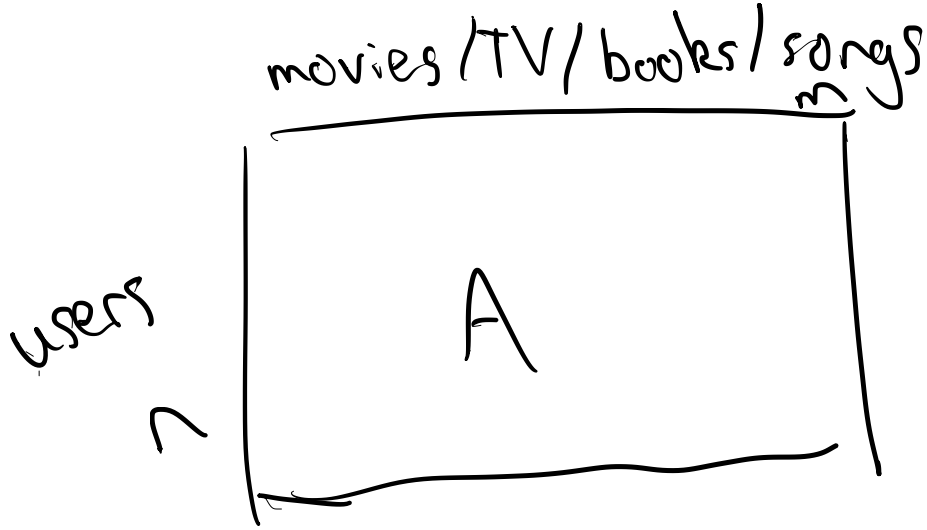
Last time: PCA, robust PCA

truncated SVD

Today: latent factor models thru matrix completion

HW1 due Thurs 11:59pm ET CMS

Recommender systems



$A_{ij} \approx$ how much user i likes song j (# plays)

song $j \in \mathbb{R}^k \Rightarrow y_j$

user $i \in \mathbb{R}^k \Rightarrow x_i$

$(y_j)_1 \approx$ how much "rap"

$(y_j)_2 \approx$ length

$A_{ij} \approx x_i^T y_j$

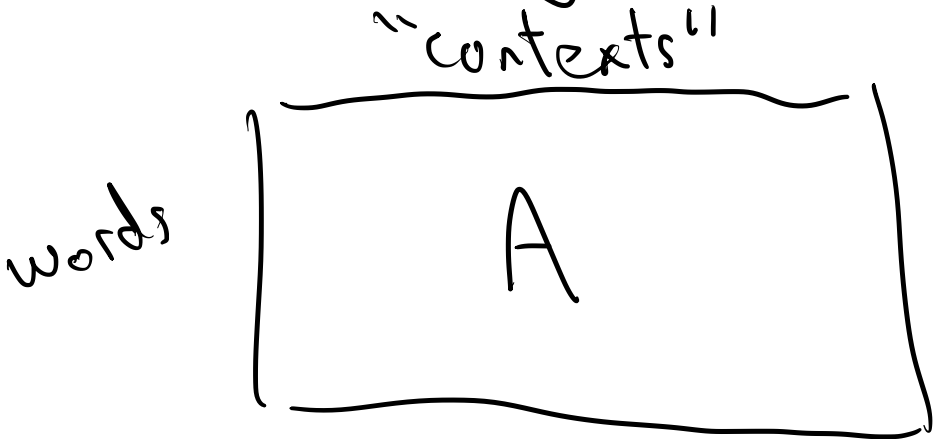
$(x_i)_1 \approx$ proclivity for rap

$(x_i)_2 \approx$ short songs

should user r like song s? $x_r^T y_s$

$A \approx XY^T \quad X = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} \quad Y^T = [y_1 \cdots y_m]$

Word embeddings



context $j \in \mathbb{R}^k \Rightarrow y_j$

word $i \in \mathbb{R}^k \Rightarrow x_i$

want: $\Pr(\text{choose } i \mid \text{context } j) \approx \dfrac{\exp(x_i^T y_j)}{\sum_r \exp(x_r^T y_j)}$

$$A_{ij} = \log(\text{Prob}(\text{choose } i \mid \text{context } j))$$

$$= x_i^T y_j - \log\left(\sum_r \exp(x_r^T y_j)\right)$$

constant per column
force to zero

$$A \approx XY^T \qquad X = \begin{bmatrix} x_1^T \\ \vdots \\ x_N^T \end{bmatrix} \qquad Y^T = [y_1 \cdots y_M]$$