Last Class

1) MF for clustering
2) Non negative MF
3) Building a Rec Sys with multiple data matrices

Today's agenda

1) Netflix Prize Solution
   - Regularization
   - Biases
   - Temporal Dynamics
   - Building a time-sensitive recsys
2) MF: Feature Engineering
   - entities
   - text data
   - co-occurence matrix
   - eigen faces

$i = m =$ movie

$d = 10$

$n_u = 1000$

$n_i = 500$

$\#\text{params} = 10000 + 5000 + 500 + 1000 + 1\ (\mu)$

$b_u = 1000$

$b_i = 500$

$q_i = 500 \times 10 = 5000$

$p_u = 1000 \times 10 = 10{,}000$

$d = 100$

Total q_i params $= 500 \times 100 = \underline{50,000}$

Total p_u params $= 1000 \times 100 = \underline{1,00,000}$

movie vector (common sense perspective)

m_1	is_romance	is_action	movie length	movie director
	0/1	0/1	150	Nolan

	i_1	i_2	i_3
u_1	3.5	5	0.5
u_2	NAN	1	NAN
u_3	NAN	8	9
u_4	7	NAN	2

A_{ij}

user-item interaction matrix

$u_1 = [3.5, 5, 0.5]$

$u_3 = [\text{NAN}, 8, 9]$

cosine similarity

$i_1 = [3.5, \text{NAN}, \text{NAN}, 7]$

cosine

$i_2 = [5, 1, 8, \text{NAN}]$

separate

$q_i = [-1.1,\ 0.7,\ 2.5,\ 8.2,\ \ldots]_d$

$p_u = [-100,\ 2,\ 17.9,\ \ldots\ldots]_d$

$q_1, \quad q_2$ similar

$p_{u_1} \quad p_{u_2}$ similarity

$$\begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}_{A_{m \times n}} = B_{m \times d} \cdot C_{d \times m}$$

$$\underline{B =} \begin{bmatrix} \longleftarrow \text{row}_i \longrightarrow \\ \longleftarrow \text{row}_j \longrightarrow \end{bmatrix}_{m \times d}$$

d-dim representation of word i

$C = \begin{bmatrix} \text{col}_j & \text{col}_i \end{bmatrix}$

$d \times m$

$\text{col}_j \rightarrow$ d-dim representation of word j

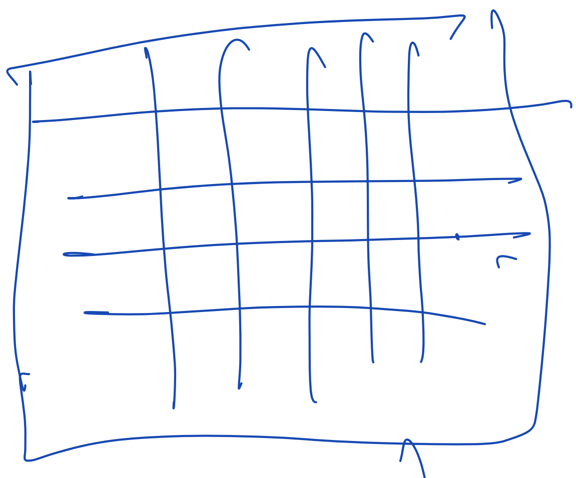
$\text{row}_k \neq \text{col}_i \rightarrow$ d-dim representation of word i

$\text{row}_k \rightarrow B$

$\text{col}_i \rightarrow C$

image :

B/W [Gray Scale]



100 rows

200 columns

BW

R

Red $\rightarrow$ $\begin{bmatrix} \quad \end{bmatrix}$ 100x200

Green $\rightarrow$ $\begin{bmatrix} \quad \end{bmatrix}$ G 100x200

Blue

$\rightarrow$ $\begin{bmatrix} \quad \end{bmatrix}$ 100x200

RGB image (color)

Gray Scale

$\begin{bmatrix} \text{row}_1 \\ \text{row}_2 \\ \text{row}_3 \\ \vdots \end{bmatrix}$ 64x64

$r = 64$

$c = 64$

$r \times c$

$1{,}000{,}000 = n$

$[\ \text{row } 1,\ \text{row } 2,\ \text{row } 3,\ \ldots\ ]$

$64 \times 64 \Longrightarrow 1 \times 64^2$

$\downarrow$

4096

$d = r \times c$

$= 64^2 = 4096$ $\quad 1 \times 4096$ vector

$$X = \begin{bmatrix} \longleftarrow \text{img } 1 \longrightarrow \\ \longleftarrow \text{img } 2 \longrightarrow \\ \vdots \\ \longleftarrow \text{img}_n \longrightarrow \end{bmatrix}$$

$n \times d$

1,000,000 4096

flattened

[row1, row2, row3, ...]

unflattening

$$\begin{bmatrix} \text{row } 1 \\ \text{row } 2 \\ \text{row } 3 \\ \vdots \end{bmatrix}$$

orig img

cov matrix S

$X_{n \times d}$ ⟶ $d \times d$

$$S = \boxed{V}\, \Sigma\, V^T$$

$d \times d \qquad d \times d \qquad d \times d$

$$\begin{bmatrix} \lambda_1 & 0 & 0 & 0 & \cdots \\ 0 & \lambda_2 & 0 & 0 & \cdots \\ 0 & 0 & \lambda_3 & 0 & 0 \\ & & & \ddots & \end{bmatrix}$$

$$\begin{bmatrix} e_1 & e_2 & \cdots & e_d \end{bmatrix}$$

d-dim

d eigenvectors

$\lambda_1, \lambda_2, \lambda_3 \ldots \lambda_k,$

$\cancel{[\lambda_{k+1}, \ldots \lambda_d]}$

$X_{n\times d} \rightarrow \underline{X'_{n\times k}}$

$X'_{n\times k} = X_{n\times d} * \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ e_1 & e_2 & \cdots & e_k \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}_{d\times k}$

$= \underline{X'_{n\times k}}$

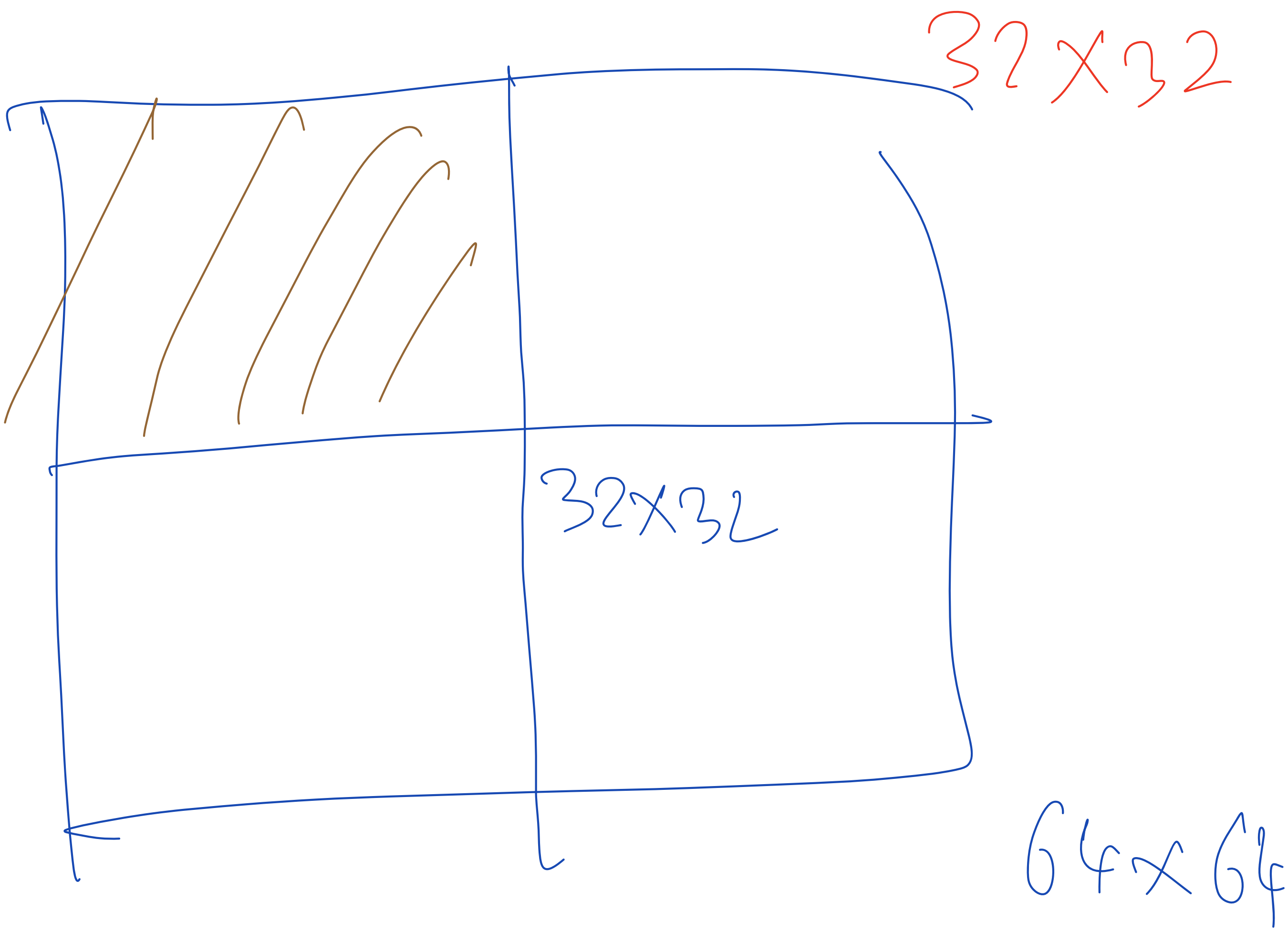
$d = 64^2$
$= 4096$

$k = 1024$

$$\begin{bmatrix} \leftarrow \text{row}_1' \longrightarrow & 1024 \\ \leftarrow \text{row}_2' \longrightarrow & \\ \vdots & \end{bmatrix}_{n\times k}$$

1,000,000
$\times$
1024

for ; → 1024



$$\frac{32 \times 32}{64 \times 64} = \frac{1}{4}$$

$$\boxed{X_{n \times k}} \longrightarrow X_{n \times d}$$

$$\downarrow$$

compressed version