Lec - 9 : Recommender System - 3

- Matrix factorization
- PCA
- Singular Value Decomposition

Recap

$$A = \begin{array}{c|cccc} & M_1 & M_2 & M_3 & M_4 \\ \hline \text{Ram} & 1 & 0 & 1 & 0 \\ \text{Shyam} & 0 & 0 & 0 & 1 \\ \text{Niharika} & 1 & 1 & 0 & 1 \\ \text{Tauseef} & 0 & 1 & 0 & 1 \end{array}$$

4×4

→ user item interaction matrix

1 : seen

0 : not seen

$A \rightarrow n \times m$

n = no. of users

m = no. of items

Nature of A : Sparse matrix

$\Rightarrow$ Most of the values are zero

$$A = \begin{array}{c|cccc} & M_1 & M_2 & M_3 & M_4 \\ \hline \text{Ram} & 4 & 0 & 2 & 0 \\ \text{Shyam} & 0 & 0 & 0 & 1 \\ \text{Niharika} & 3 & 5 & 0 & 1 \\ \text{Tauseef} & 0 & 2 & 0 & 4 \end{array}_{4\times 4}$$

user item interaction : Ratings

$$A = \begin{array}{c|cccc} & M_1 & M_2 & M_3 & M_4 \\ \hline \text{Ram} & 4 & 0 & 2 & 0 \\ \text{Shyam} & 0 & 0 & 0 & 1 \\ \text{Niharika} & 3 & 5 & 0 & 1 \\ \text{Tauseef} & 0 & 2 & 0 & 4 \end{array}_{4\times4}$$

Sparse

$\longrightarrow$ **Matrix Completion** $\longrightarrow$

$\longrightarrow$ $\boxed{\text{MF + optimization}}$ $\longrightarrow$

$$A' = \begin{array}{c|cccc} & M_1 & M_2 & M_3 & M_4 \\ \hline \text{Ram} & 4 & 2.5 & 2 & 4.5 \\ \text{Shyam} & 2 & 3 & 4.5 & 1 \\ \text{Niharika} & 3 & 5 & 2.2 & 1 \\ \text{Tauseef} & 1.2 & 2 & 4.6 & 4 \end{array}_{4\times4}$$

Dense

Matrix factorization

Eg. $6 = 2 \times 3$ (2 and 3 are factors of 6)

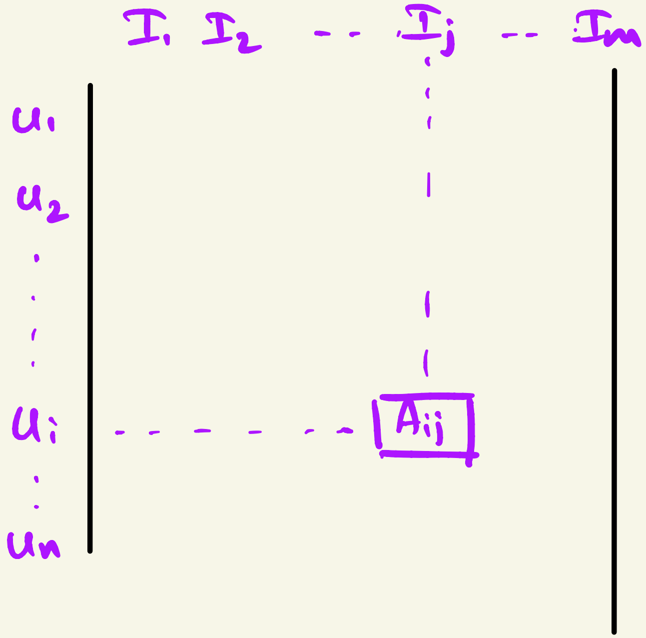
$30 = 2 \times 3 \times 5$

factors

→ extend this to matrices → $\underset{n \times m}{A} = \underset{n \times d}{B} \; \underset{d \times m}{C}$

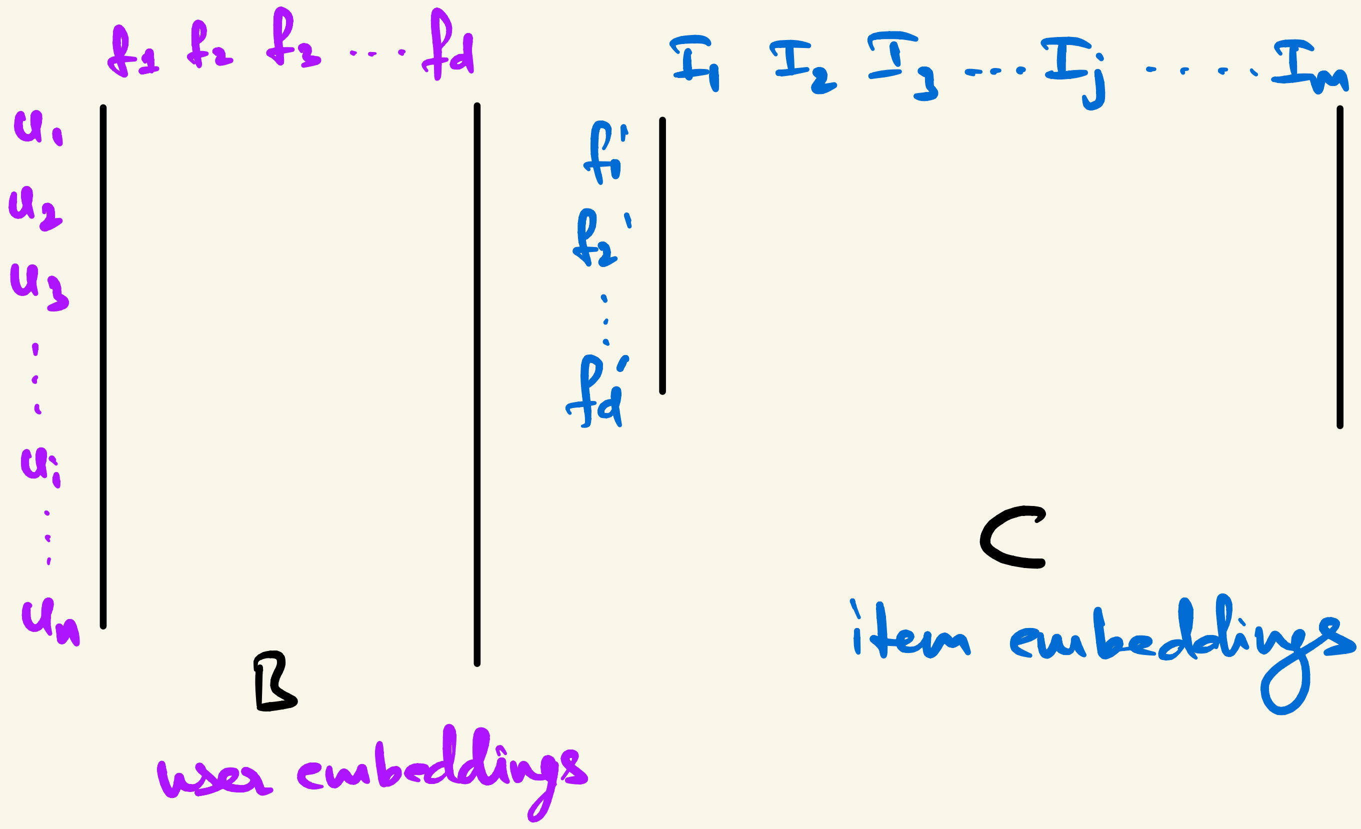
→ $\underset{n \times m}{A} = \underset{n \times n}{B} \cdot \underset{n \times m}{C} \cdot \underset{m \times m}{D}$

$A =$



A_{ij} = interaction b/w U_i and I_j

$$\underset{n \times m}{A} = \underset{n \times d}{B}\ \underset{d \times m}{C} =$$

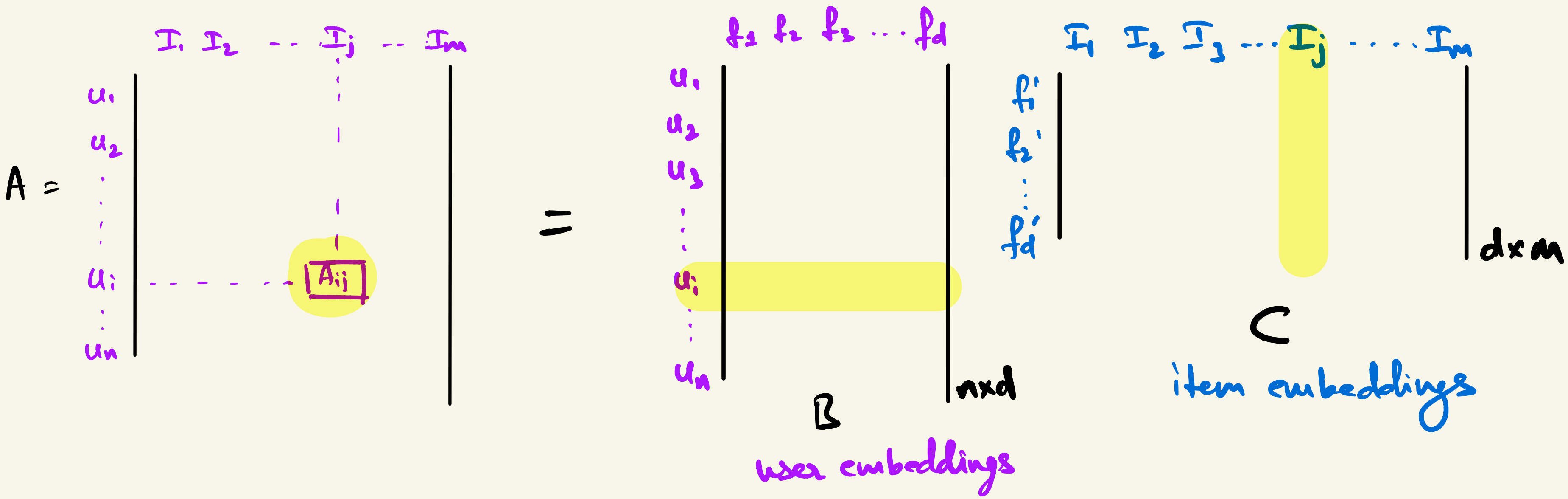


Eg. Recap of matrix multiplication

$$B = \begin{vmatrix} 2 & 3 \\ 3 & 1 \\ -1 & 0 \end{vmatrix}_{3\times 2} \cdot \begin{vmatrix} 1 & 0 & 2 \\ 5 & -1 & 3 \end{vmatrix}_{2\times 3} = C$$

$$3\times 2 \,,\; 2\times 3$$
$$\shortparallel$$
$$3\times 3$$

$$BC = \begin{vmatrix} 2\times1+3\times5 & 2\times0+3\times-1 & 2\times2+3\times3 \\ 3\times1+1\times5 & 3\times0+3\times-1 & 3\times2+1\times3 \\ & A_{32} & \end{vmatrix}_{3\times 3}$$



$$A_{ij} = u_i^T I_j = B_i^T C_j$$

$$B \longleftrightarrow U$$

$$C \longleftrightarrow I$$

Recap of vectors

$$\bar{B}_i = \begin{vmatrix} 2 \\ 0 \\ 3 \end{vmatrix} \rightarrow \bar{B}_i^T = \begin{vmatrix} 2 & 0 & 3 \end{vmatrix}$$

$$C_j = \begin{vmatrix} 3 \\ 1 \\ 2 \end{vmatrix}$$

Dot product of $\bar{B}_i$ and $\bar{C}_j = \bar{B}_i^T \bar{C}_j$

$$= 2 \times 3 + 0 \times 1 + 3 \times 2$$

Continued

⇒ B_i = user embedding for the i^{th} user

C_j = item embedding for the j^{th} item

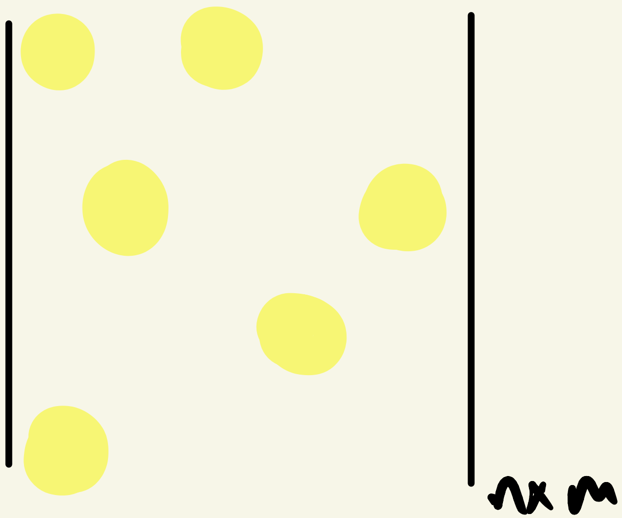
$B_i^T C_j$ = predicted interaction b/w the i^{th} user and j^{th} item

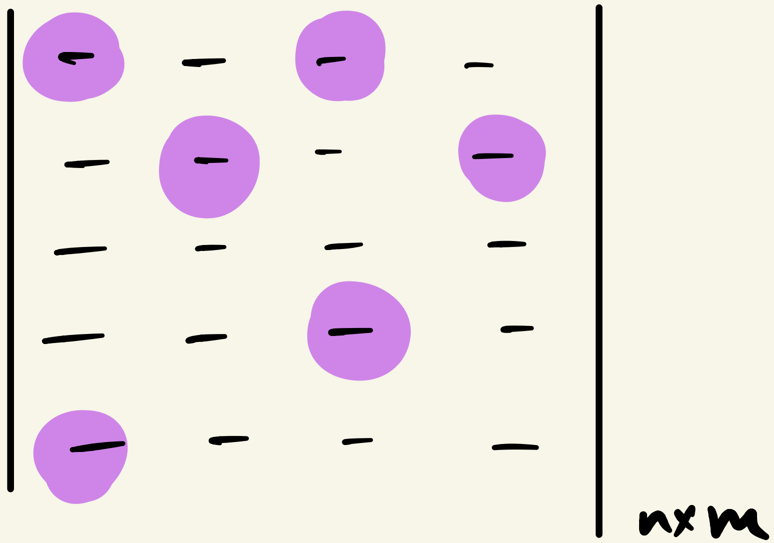
A_{ij} = Actual interaction b/w the i^{th} user and j^{th} item

Process

S1) Randomly initialize all $\bar{B}_i$'s and $\bar{C}_j$'s

$\Rightarrow BC = A'$ = predicted interaction matrix

$A =$ 

$A' =$ 

 → observed in real

$$\text{Loss} = \text{error} = \sum_{i,j \in obs} \left(A_{ij} - \bar{B}_i^T \bar{C}_j \right)^2$$

obs : observed

Next :

→ <u>optimization</u> ↗ Stochastic G.D.
↘ Coordinate Descent (weighted alternating least squares)

M1) Stochastic G.D.

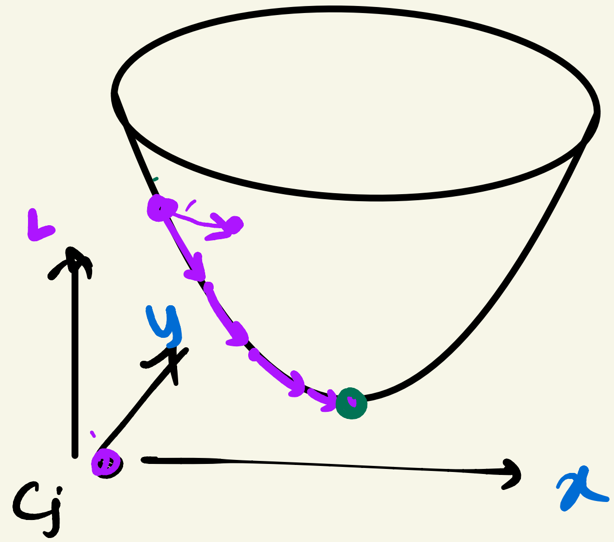
update rule : $\bar{B}_i^{t+1} = \bar{B}_i^{t} - \eta \bar{\nabla}_{\bar{B}_i} L \quad \forall B_i s$

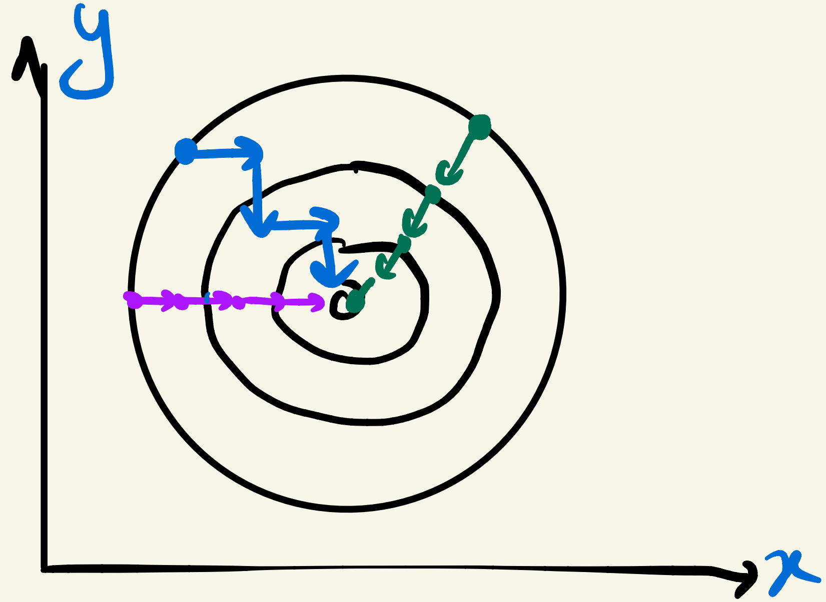
$$\bar{C}_j^{t+1} = C_j^{t} - \eta \bar{\nabla}_{\bar{C}_j} L$$

M2)

First update all the B_i's keeping all the C_j's fixed

Then update all C_j's keeping all the B_i's fixed





$L = x^2 + y^2$

Summarize:

$$A = \left| \quad\quad\quad \right| \xrightarrow{\text{matrix completion}} A' : \left| \quad\quad\quad \right|$$

sparse $\qquad\qquad\qquad\qquad$ dense

① Randomly initiate B and C

$A = BC \rightarrow$ ② $L = \sum_{i,j \in obs} \left(A_{ij} - \bar{B}_i^T \bar{C}_j \right)^2$

③ optimization → SGD
→ CD/ALS

↓

$$B^* \text{ and } C^* \Rightarrow \underbrace{A'}_{n \times m} = \underbrace{B^*}_{n \times d} \underbrace{C^*}_{d \times m}$$

How to generate recommendations.

$$A' = \begin{array}{c} u_1 \\ u_2 \\ \\ u_i \\ \vdots \\ u_n \end{array} \left| \begin{array}{c} \\ \\ \\ \hphantom{xxxxxxxxxxxxxxxx} \\ \\ \\ \end{array} \right|$$

Actual

$$u_i = [4 \quad 0 \quad 0 \quad \ldots\ldots]$$

$$u_i' = [4.3 \quad 3.7 \quad 4.2 \quad \ldots\ldots]$$

Predicted

recommend this to the users

Quiz

$$A = \begin{array}{c|cccccccc|} u_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & & & \end{array}$$

$$L = \sum_{i,j \in obs} \left(A_{ij} - B_i^T C_j \right)^2$$

$$\nabla_{\bar{B}_i} L = 0 \text{ for all updates}$$

$$B = \begin{array}{c|c|} B_1 & \\ B_2 & \\ \vdots & \\ B_i & \end{array}_{n \times d}$$

$$C = \begin{array}{|cccc|} C_1 & C_2 & \cdots & C_j \cdots \end{array}_{d \times m}$$

$\Rightarrow \bar{B}_i$ final

$= \bar{B}_i$ initial

= randomly initiated one

MF + optimization $\longrightarrow$ matrix completion

$\downarrow$

Recommendation Sys

Principal Component Analysis

$X\bar{v} = \lambda\bar{v}$ $\Rightarrow$ $\bar{v}$ is an eigen vector of X and its eigenvalue is λ

$\Rightarrow S_{d\times d} = X^T_{d\times n} X_{n\times d} = \mathrm{cov}(X)$

$$\boxed{S_{d\times d} = W_{d\times d} \Lambda_{d\times d} W^T_{d\times d}} \longrightarrow \text{Matrix factorization (eigen value decomposition)}$$

$$= \begin{bmatrix} | & | & | & & | \\ \bar{w}_1 & \bar{w}_2 & \bar{w}_3 & \cdots & \bar{w}_d \\ | & | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots \\ 0 & \lambda_2 & & \\ 0 & \cdots & \lambda_3 & \\ 0 & \cdots & & \ddots & \lambda_d \end{bmatrix} \begin{bmatrix} - & \bar{w}_1^T & - \\ - & \bar{w}_2^T & - \\ & \vdots & \\ - & w_d^T & - \end{bmatrix}$$

Constraints : ① $\lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_d$

② $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_d$ are orthogonal to each other

③ S should be a square matrix

Singular value Decomposition (SVD)

$$X_{n \times d} = \underset{n \times n}{U} \; \underset{n \times d}{\Sigma} \; \underset{d \times d}{V^T}$$

$$\Sigma = \begin{vmatrix} s_1 & & & & \\ & s_2 & & & \\ & & s_3 & & \\ & & & \ddots & \\ & & & & s_d \end{vmatrix}$$

U : left singular matrix

V : Right singular matrix

Σ : singular value matrix

Q) Can we do SVD of the user item interaction matrix?

Ans) Yes.