Last Class (Dec 4)

1) Leverage, Conviction
2) Summary of market basket analysis
3) Intro to Reco Sys
4) Formulating Problem for Reco Sys
5) Collaborative Filtering
6) Item-item and User-user similarity
7) Cold-start problem
8) Content-based Reco Sys

Today's agenda

1) Matrix Factorization
2) PCA
3) Singular Value Decom
4) Matrix Factorization for Clustering
5) Hyper-param tuning
6) NMF

$12 = 4 \times 3$

$= 2 \times 2 \times 3$

$= 2^2 \times 3$

$\underline{36} = 6 \times 6$

$= 6 \times 3 \times 2$

$= 2 \times 3 \times 3 \times 2$

$= 2 \times 3^2 \times 2$

$= 2^2 \times 3^2$

$$\underbrace{\overset{A}{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}}_{2 \times 2} = \underbrace{\overset{B}{\begin{bmatrix} w & x \\ y & z \end{bmatrix}}}_{2 \times 2} \times \underbrace{\overset{C}{\begin{bmatrix} l & m \\ n & o \end{bmatrix}}}_{2 \times 2}$$

$$= \underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}}_{2 \times 3} \times \underbrace{\begin{bmatrix} c_1 & d_1 \\ c_2 & d_2 \\ c_3 & d_3 \end{bmatrix}}_{3 \times 2}$$

2×2

$$2 \times 2 = (2 \times d) \times (d \times 2)$$

where $d = 1, 2, 3, \ldots, N$

$$A = \begin{array}{cc} & \begin{array}{cccc} I_1 & I_2 & \cdots & I_{10} \end{array} \\ \begin{array}{c} U_1 \\ U_2 \\ \vdots \\ U_{15} \end{array} & \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1,10} \\ r_{21} & r_{22} & \cdots & r_{2,10} \\ \vdots & & & \vdots \\ r_{15,1} & r_{15,2} & \cdots & r_{15,10} \end{bmatrix} \end{array}$$

User 2 for item 10

User 15 for item 2

$A \simeq U \times I$

User-item interaction matrix

$r_{2,3} \rightarrow$ User 2's rating on item 3

$\rightarrow \bar{u}_2 \rightarrow [a_1 \; a_2 \cdots a_d]_{1 \times d}$ d-dim

vector

$\bar{i}_3 \rightarrow [b_1 \; b_2 \cdots b_d]$ d-dim

$$i_3 = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_d \end{bmatrix}_{d \times 1}$$

$$\bar{u}_2 \times \bar{i}_3 = (1 \times d) \times (d \times 1) = \boxed{1 \times 1} \rightarrow \text{single value}$$

$\rightarrow r'_{2,3}$

$$r_{15,1} \simeq \bar{u}_{15} \times \bar{i}_1$$

$r_{1,2} \rightarrow \bar{u}_1 \times \bar{i}_2$

$r_{2,3} \rightarrow \bar{u}_2 \times \bar{i}_3$

$r_{1,1} \rightarrow \bar{u}_1 \times \bar{i}_1$

$A \simeq U \times I$

A: user-item matrix

U: user vector matrix

I: item vectors matrix

$$\underset{(\bar{n} \times \bar{k})}{A} \simeq \underset{(\bar{n} \times d)}{U} \times \underset{(d \times \bar{k})}{I}$$

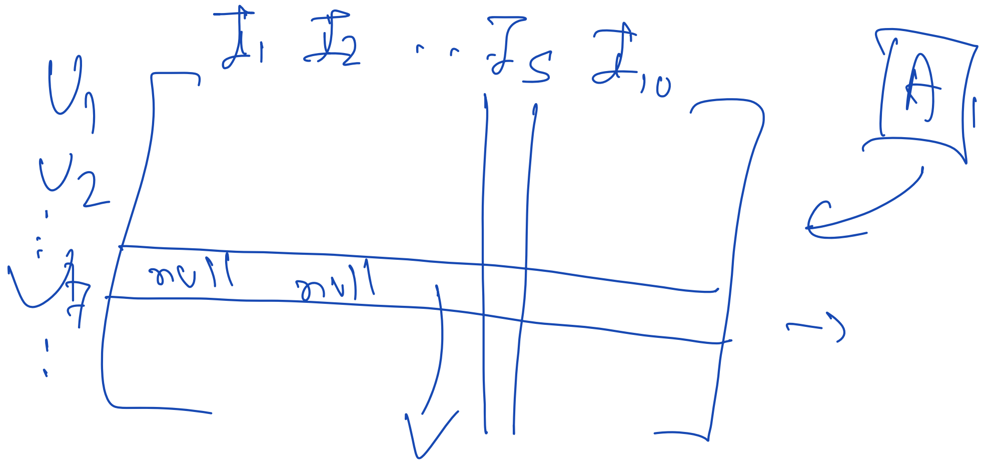
$\bar{n}$: no. of users

$\bar{k}$: no. of items

$\boxed{r_{7,6}}$ = empty in A

$$r'_{7,6} = \bar{u}_7 \times \bar{i}_6$$

$r_{7,6} \rightarrow \bar{u}_7 \quad \bar{i}_6$



user-item

V_7 row is empty

u_7 vector is possible ?

i_5 vector is not possible

$$\underset{n\times 1}{Y} = \underset{(n\times d)}{X}\ \underset{(d\times 1)}{\boxed{W}}\ \boxed{+\,C} \rightarrow \text{ignore}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1,d} \\ x_{21} & x_{22} & \cdots & x_{2,d} \\ \vdots & & & \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix}$$

$\times$

SGD

$Y \simeq XW$

Reco Sys

$A \simeq U \times I$

$$\sum_{i=1}^{n} (y_i - w \cdot x_i)^2 \quad \text{Loss fn}$$

$$\sum_{j=1}^{k} \sum_{i=1}^{n} \left(A_{ij} - \bar{u}_i \cdot \bar{i}_j\right)^2$$

MSE loss fn

$\bar{u}_i$ for all $i = 1, 2, 3, \ldots, n$

$\bar{i}_j$ for all $j = 1, 2, 3, \ldots, k$

$A \simeq U \times I$

$$U = \begin{bmatrix} \bar{u}_{1,1} & \bar{u}_{1,2} & \cdots & \bar{u}_{1,d} \\ \vdots & & & \\ \bar{u}_{n,1} & \bar{u}_{n,2} & \cdots & \bar{u}_{n,d} \end{bmatrix}$$

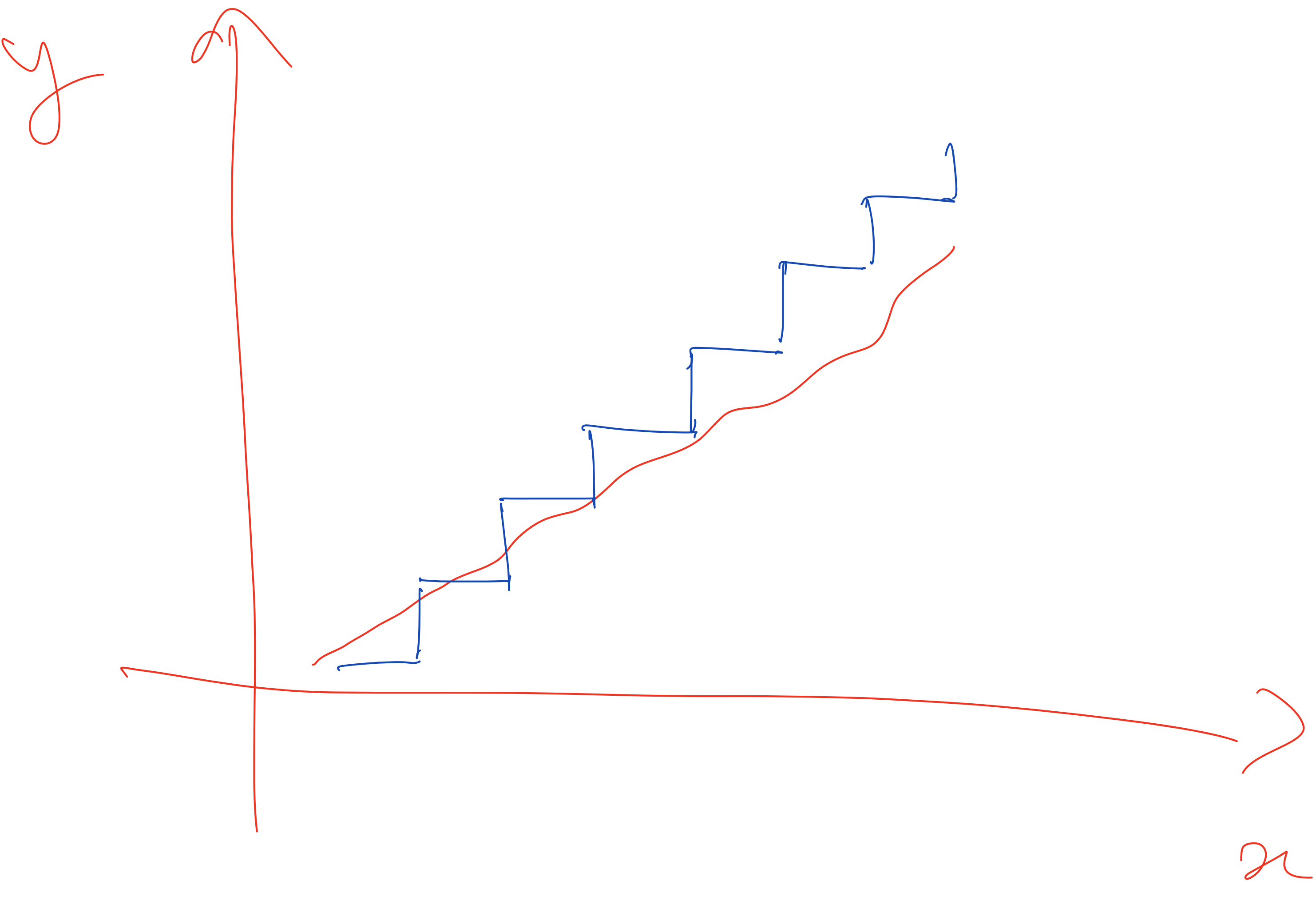
user vectors are along rows

$n \times d$

$$I = \begin{bmatrix} \bar{i}_{1,1} & \cdots & \bar{i}_{1,k} \\ \bar{i}_{2,1} & & \vdots \\ \vdots & & \\ \bar{i}_{d,1} & & i_{d,k} \end{bmatrix}$$

$d \times k$

item vectors are along columns



I) $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & -1 & 3 \\ 1 & 4 & 2 & 3 \end{bmatrix} \rightarrow$ Can we get all $\bar{u}_i$'s & $\bar{i}_j$'s ?

II) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ \cancel{0} & 1 & -1 & \cancel{0} \\ 1 & \cancel{4} & \cancel{0} & 3 \end{bmatrix} \rightarrow$ all user vectors and item vectors possible

III) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0} \\ 1 & 4 & 2 & 3 \end{bmatrix}$ ✗

null = $\cancel{0}$ ✓

null $\neq 0$

PCA

$X \rightarrow n \times d$

$\bar{X} \rightarrow$ zero-mean

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix}$$

$\overline{0 \quad 0 \quad 0}$

zero-mean

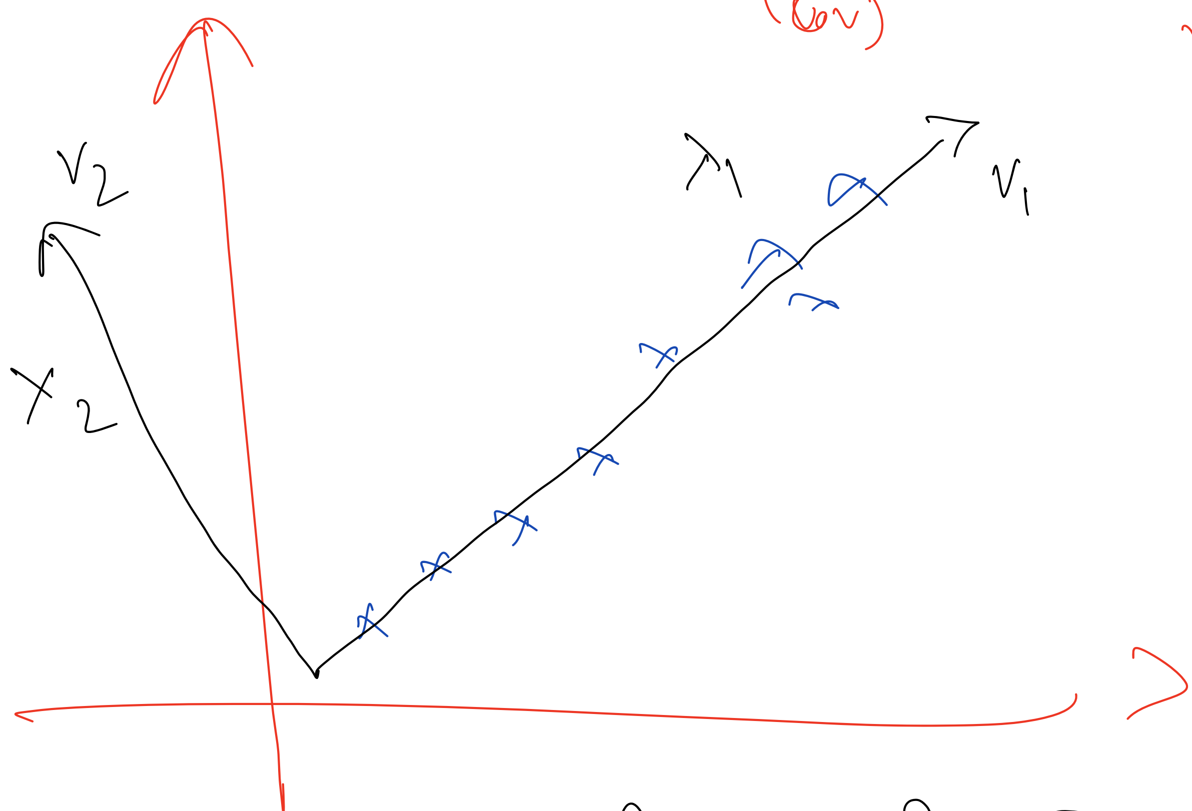
$\overline{X} = (n \times d)$

$Cov = d \times d$

$$\frac{1}{n}\Big(\underbrace{\overline{X}^T}_{d \times n} \; \underbrace{\overline{X}}_{n \times d}\Big)$$

$S \rightarrow$ covariance matrix

(Cov)



$A \approx B \cdot C$

$$A \simeq U \times I$$

$$V \times I \simeq A$$

$$= A \times$$

$$\begin{bmatrix} \sigma_x^2 & \sigma_{x,y} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}^T$$

$Cov(x, y)$

$$\begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$