Lecture 3: Matrix multiplication, linear independence, & Rank

Exams:

10/11, 7:15pm, 1800 EH

11/15, 7:15pm, 1800 EH

Last time:

inner product $\langle \underline{x}, \underline{w} \rangle = \sum_{j=1}^{p} w_j x_j = \underline{x}^T \underline{w} = \underline{w}^T \underline{x}$

"weighted sum of elements of x"

matrix-vector mult: $X\underline{w} = \begin{bmatrix} - \underline{x}_1^T - \\ - \underline{x}_2^T - \\ \vdots \\ - \underline{x}_n^T - \end{bmatrix} \underline{w} = \begin{bmatrix} \underline{x}_1^T \underline{w} \\ \underline{x}_2^T \underline{w} \\ \vdots \\ \underline{x}_n^T \underline{w} \end{bmatrix}$

Next: matrix-matrix mult.

Example: Recommender System

	Becca	Noah	Yudon	Tanvi
Star Wars	10	10	8	5
Pride + Prejudice	10	2	1	10
La La Land	8	3	1	10
Taken	1	7	10	2
Walking Dead	1	9	10	4

$= X$

Let's write X as product of 2 matrices T (tastes) and W (weights)

$$\begin{bmatrix} \\ & X & \\ \\ \end{bmatrix} = \begin{bmatrix} \\ T \\ \\ \end{bmatrix} \begin{bmatrix} & W & \end{bmatrix}$$

$X \in \mathbb{R}^{n \times p}$ $\quad$ $T \in \mathbb{R}^{n \times r}$ $\quad$ $W \in \mathbb{R}^{r \times p}$

n movies, p people

Think of

T = taste profiles of r representative customers

W = weights on each representative profile
(1 set of weights for each customer)

e.g.

$$T = \begin{array}{cc} \text{action lover} & \text{romance lover} \\ \end{array}$$

$$T = \begin{array}{c|c} \begin{bmatrix} 10 & 4 \\ 1 & 10 \\ 1 & 10 \\ 8 & 1 \\ 10 & 1 \end{bmatrix} & \begin{array}{l} SW \\ PP \\ LLL \\ T \\ WD \end{array} \end{array}$$

$$\underline{W} = \begin{bmatrix} 3/8 \\ 5/8 \end{bmatrix}$$

so

$$T\underline{W} = \begin{bmatrix} 6.2 \\ 6.6 \\ 6.6 \\ 3.6 \\ 4.3 \end{bmatrix}$$

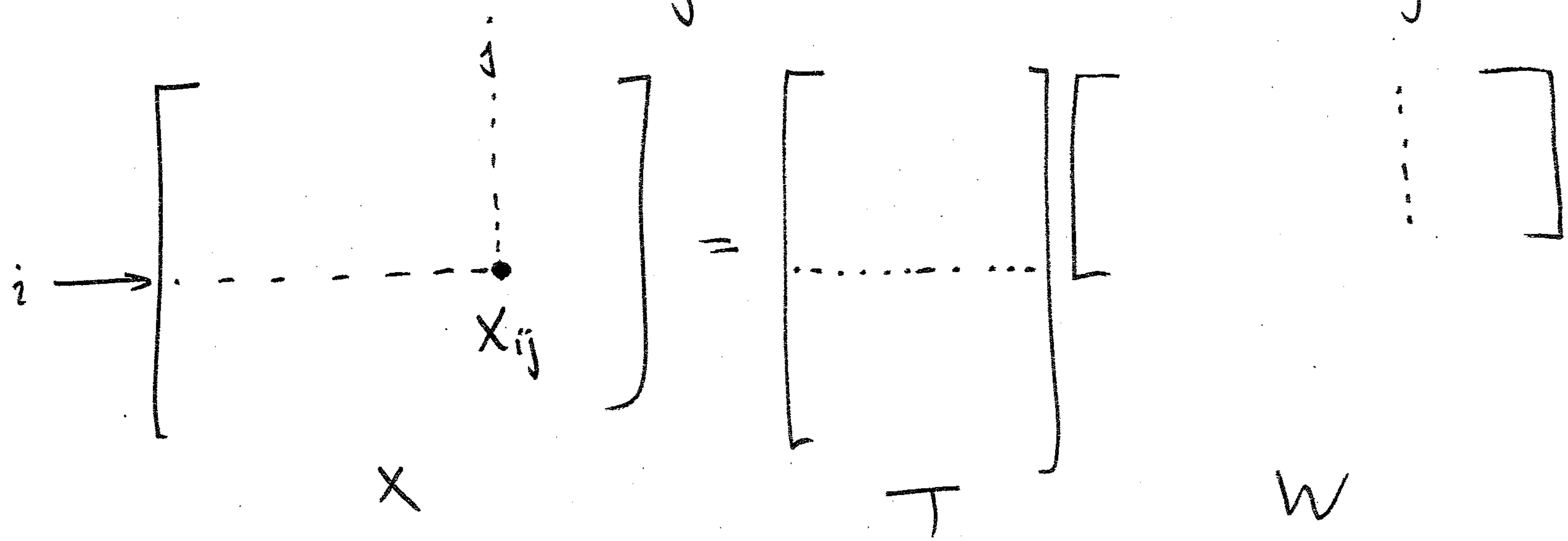
= expect prefs of customer who is 3/8 action lover + 5/8 romance lover

Matrix W will have a weight vector column for each customer

$$TW = \begin{bmatrix} 10 & 4 \\ 1 & 10 \\ 1 & 10 \\ 8 & 1 \\ 10 & 1 \end{bmatrix} \begin{bmatrix} 7 & 1 & 8 & 3 & \cdots \\ 3 & 9 & 2 & 7 & \end{bmatrix} / 10$$

= product of 2 matrices

if $X = TW \Rightarrow X_{ij} = \langle i^{th} \text{ row of } T, \; j^{th} \text{ col of } W \rangle$



- What is a column of X?

j^{th} column of X = weighted sum of col. of T, where j^{th} col of W tells us the weights

$= T \cdot \underline{w}_j$ = tastes of j^{th} customer

- What is a row of X?

i^{th} row of $X = \underline{x}_i^T = \underline{t}_i^T W$ where $\underline{t}_i^T = i^{th}$ row of T

= how much each customer likes movie i

Inner product representation:

$$TW = \begin{bmatrix} - \underline{t}_1^T - \\ - \underline{t}_2^T - \\ \vdots \\ - \underline{t}_n^T - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \underline{w}_1 & \underline{w}_2 & \cdots & \underline{w}_4 \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \underline{t}_1^T \underline{w}_1 & \underline{t}_1^T \underline{w}_2 & \cdots & \underline{t}_1^T \underline{w}_p \\ \underline{t}_2^T \underline{w}_1 & & & \vdots \\ \vdots & & & \vdots \\ \underline{t}_n^T \underline{w}_1 & & & \underline{t}_n^T \underline{w}_p \end{bmatrix}$$

Outer product representation

$$TW = \begin{bmatrix} | & | & & | \\ T_1 & T_2 & \cdots & T_r \\ | & | & & | \end{bmatrix} \begin{bmatrix} - W_1^T - \\ \vdots \\ - W_r^T - \end{bmatrix}$$

think of each T_k = k^{th} representative taste profile

w_k = k^{th} row of W = affinity of each ~~col of~~ customer with k^{th} representative.

Blocks everything above generalizes to blocks in matrices

$$\begin{bmatrix} 1 & 2 & 0 \\ 3 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 4 & 5 \\ 4 & 4 \end{bmatrix}$$

$$\begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 2 & 2 \end{bmatrix} & \begin{bmatrix} 1 \end{bmatrix} \end{bmatrix}\begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \\ \begin{bmatrix} 0 & 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix}\begin{bmatrix} 0 & 2 \end{bmatrix} \\ \begin{bmatrix} 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 \end{bmatrix}\begin{bmatrix} 0 & 2 \end{bmatrix} \end{bmatrix}$$

Linear Independence

Earlier, we wrote $\underset{n \times p}{\underset{\sim}{X}} = \underset{n \times r}{\underset{\sim}{T}} \; \underset{r \times p}{\underset{\sim}{W}}$

Matrix multiplication holds for any r

~~But~~ given X, what is the smallest r such that we can find T, W so that $X = TW$

How many representatives are needed?

Defn: linear independent vectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n \in \mathbb{R}^p$

$$\sum_{j=1}^{n} \alpha_j \underline{v}_j = 0 \iff \alpha_j = 0 \quad \text{for } j = 1, \ldots, n$$

ex. $p = 3$, $V_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $V_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ yes, Lin. indep

$$\alpha_1 V_1 + \alpha_2 V_2 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_2 \end{bmatrix} = 0 \text{ only if } \alpha_1 = \alpha_2 = 0$$

ex. $p = 3$. $V_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $V_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $V_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

$$\alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 V_2 = \begin{bmatrix} \alpha_1 \\ \alpha_2 + \alpha_3 \\ \alpha_2 \end{bmatrix} \Rightarrow \text{yes, Lin. indep.}$$

ex. $p = 3$, $V_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $V_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $V_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $V_4 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

$$V_4 = V_1 + V_2 - V_3$$

$$\alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 V_3 + \alpha_4 V_4 = \begin{bmatrix} \alpha_1 + \alpha_4 \\ \alpha_2 + \alpha_3 \\ \alpha_2 + \alpha_4 \end{bmatrix}$$

if $\alpha_1 = -\alpha_4 = \alpha_2$
$= -\alpha_3$

Linear independence $\Longrightarrow$ $p \geq n$

$n > p$ $\Longrightarrow$ linearly dependent

$\boxed{\text{Matrix rank}}$ = max # of linearly independent columns.

if $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{p \times n}$

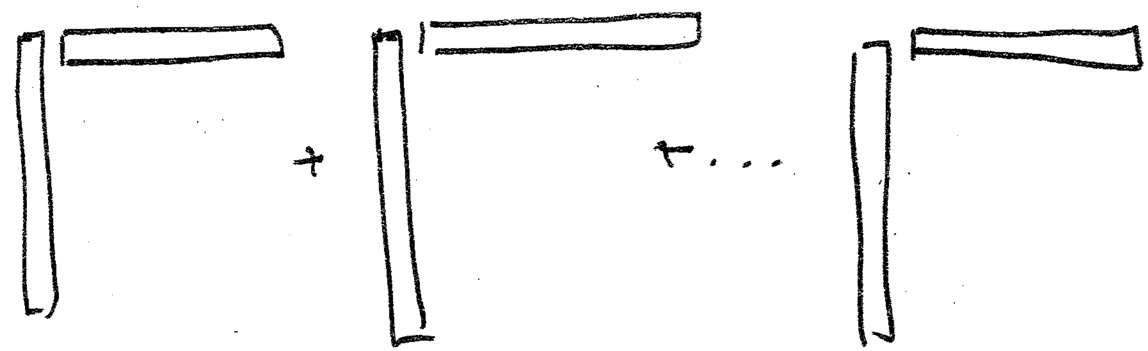
$\text{rank}(X) \leq \min(p, n)$

ex: $X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \Longrightarrow \text{rank}(X) = 2$

ex. $X = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \Longrightarrow \text{rank}(X) = 2$

ex. Recall outer product representation

$$TW = \sum_{k=1}^{r} \underline{t}_k \, \underline{w}_k^T =$$



rank = 1

sum of r rank-1 matrices

$\Longrightarrow \text{rank}(TW) = r$

Distance matrix

We have n items in p-dim space

$$\underline{x}_i \in \mathbb{R}^p \quad , \quad i = 1, \dots, n$$

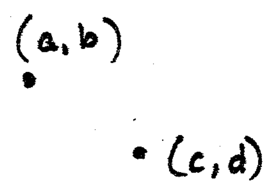
(e.g. $\underline{x}_i$ = location in space of i^{th} satellite)

look at $(\text{distances})^2$ between pairs of satellites

how measure distance?

$$\text{dist}(\underline{x}_i, \underline{x}_j) = \left(\sum_{k=1}^{p} (x_{ik} - x_{jk})^2 \right)^{1/2}$$

e.g. $p = 2$



$$\text{dist} = \sqrt{(a-c)^2 + (b-d)^2}$$

form (distance)^2 matrix $D \in \mathbb{R}^{n \times n}$

$$D_{ij} = (\text{dist}(\underline{x}_i, \underline{x}_j))^2$$

algebra

$$= \langle \underline{x}_i - \underline{x}_j \,,\, \underline{x}_i - \underline{x}_j \rangle$$

$$= \langle \underline{x}_i, \underline{x}_i \rangle - 2\langle \underline{x}_i, \underline{x}_j \rangle + \langle \underline{x}_j, \underline{x}_j \rangle$$

Q1: if $\langle \underline{x}_i, \underline{x}_i \rangle = 1$ for all i

what is rank(D)? (assume $p < n$, items are not aligned $\Rightarrow X = [x_1, x_2, \ldots, x_n]$ full rank (rank = p))

Q2: What if $\langle \underline{x}_i, \underline{x}_i \rangle$ is not assumed to be 1? what is rank(D)?