Lecture 1: Intro

Data can be rep as a mtx

e.g.

$$X = \begin{array}{c} \begin{bmatrix} 6.1 & 100 & 10 \\ 5.5 & 150 & 20 \\ 7.3 & 200 & 30 \end{bmatrix} \\ \begin{array}{ccc} \text{height} & \text{weight} & \text{age} \end{array} \end{array}$$

\# observations

$N \times P$

\# variables

$N = 3$
$P = 3$

Can view a mtx as a collection of rows

$$X = \begin{bmatrix} \text{---} \; x_1 \; \text{---} \\ \text{---} \; x_2 \; \text{---} \\ \text{---} \; x_3 \; \text{---} \end{bmatrix}$$

$x_n \in \mathbb{R}^P$
(lower-case = observations)

e.g. $x_1 = (6.1, 100, 10)$

Can also view X as collection of cols

Can also view X as collection of cols

$$X = \begin{bmatrix} | & | & | & \\ X_1 & X_2 & X_3 & \cdots \\ | & | & | & \end{bmatrix} \quad X_p \in R^n$$

(upper = vars)

e.g. $X_1 = (6.1, 5.5, 7.3)$.

Inner Products

$a, b \in \mathbb{R}^p$ then the inner product is

$$a^T b = \sum_{k=1}^{p} a_k b_k$$

a number

outer product

$$\underbrace{a}_{p\times 1}\underbrace{b^T}_{1\times p} = \begin{bmatrix} b_1 a & b_2 a & b_3 a & \cdots \end{bmatrix}$$

$\{ \qquad \}$

$\uparrow$ a $p \times p$ mtx

Norm: $\|a\| = \sqrt{\sum_k a_k^2} = \sqrt{a^T a}$

What about matrices? needs to match

$A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$

then $AB \in \mathbb{R}^{m \times p}$

4 ways to define AB

① Sum of Inner Products

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

= row i of A $\cdot$ col j of B

(2) Lin. Combn of Cols of A

$$B = \begin{bmatrix} | & | & | & & | \\ B_1 & B_2 & B_3 & \cdots & B_p \\ | & | & | & & | \end{bmatrix}$$

then

$$AB = \begin{bmatrix} | & | & | & & | \\ AB_1 & AB_2 & AB_3 & \cdots & AB_p \\ | & | & | & & | \end{bmatrix}$$

lin. combn of Cols of A

(3) LC of Rows of B

$$A = \begin{bmatrix} — & a_1 & — \\ — & a_2 & — \\ & \vdots & \\ — & a_m & — \end{bmatrix}$$

LC of rows of B

then $$AB = \begin{bmatrix} — & a_1 B & — \\ — & a_2 B & — \\ & \vdots & \\ — & a_m B & — \end{bmatrix}$$

④ Sum of Outer Products

$$AB = \sum_{k=1}^{n} \overbrace{A_k}^{m\times 1} \overbrace{b_k^T}^{1\times p}$$

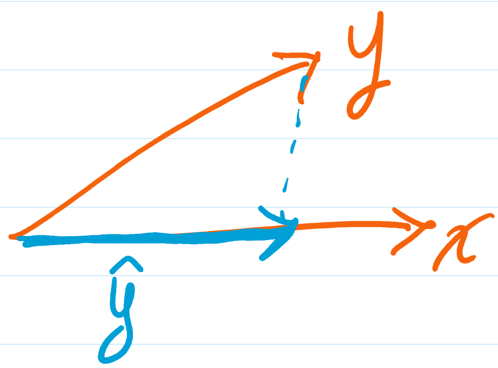
k^{th} row of B

k^{th} col of A

Projection

$x, y \in \mathbb{R}^p$ then the proj. of y onto x:

$$\hat{y} = u_x u_x^T y$$



u_x = unit vec. in dir of x

$$u_x = \frac{x}{\|x\|}$$

$$= \frac{x}{\|x\|} \frac{x^T}{\|x\|} y$$

$$\|x\| = \sqrt{x^T x}$$

$$= \frac{x}{\|x\|}\frac{x^T}{\|x\|}\, y$$

$$\|x\| = \sqrt{x^T x}$$

$$\|x\|^2 = x^T x$$

$$= \frac{x x^T y}{\|x\|^2}$$

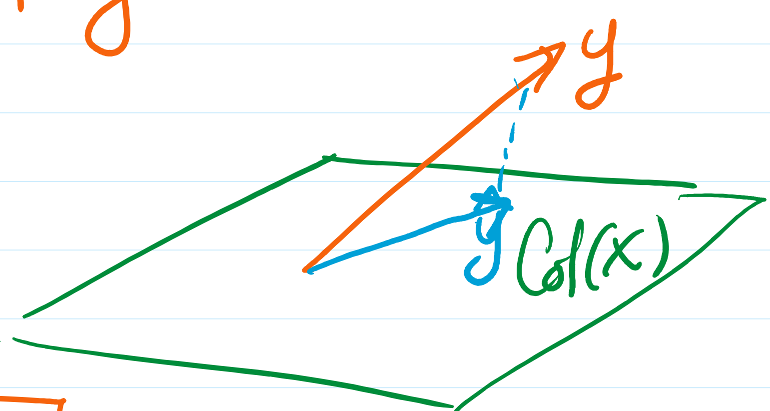
$$= x x^T y / x^T x$$

$$= \boxed{x(x^T x)^{-1} x^T y = \hat{y}}$$

---

what abut matrices?

$\hat{y}$ = proj. of y onto Col(X)



$$\boxed{\hat{y} = \underbrace{X\overbrace{(X^T X)^{-1}}^{p\times p} X^T}\, y}$$

Col(X)

$0 \qquad 0$

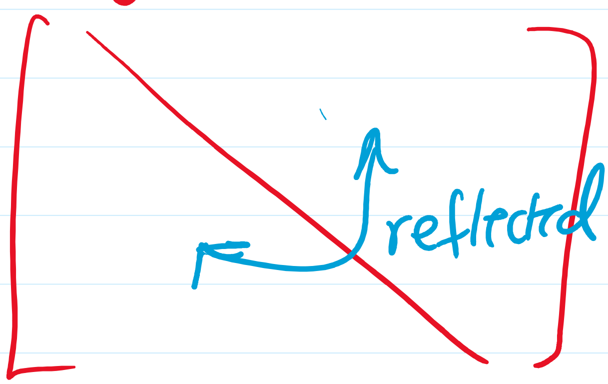
P_X = proj. onto Col(X)

Orthogonality: $u^T v = 0$

Special Mtxes

Symmetric: X is symmetric if $X = X^T$

$X =$



e.g. $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is symm.

$\begin{bmatrix} 1 & 0 & 5 \\ 0 & 2 & 4 \\ 5 & 4 & 3 \end{bmatrix}$ is symm.

$[5\ 4\ 3]$

Diagonal: zeros off main diag

$$D = \begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & d_3 & \\ & 0 & & \ddots & \\ & & & & d_n \end{pmatrix}$$

$$= \operatorname{diag}(d_1, d_2, \ldots, d_3)$$

properties:

① Symmetric : $D = D^T$

② $D^{-1} = \operatorname{diag}\left(\frac{1}{d_1}, \frac{1}{d_2}, \ldots, \frac{1}{d_3}\right)$

③ $Dx = (d_1x_1, d_2x_2, d_3x_3, \ldots, d_nx_n)$

Orthogonal Mtx : Q (square)

Orthogonal mtx : Φ (square)

the cols of Φ are ortho-normal

1. Cols are orthogonal to each other
2. all cols are unit vectors

properties:

1. $\Phi^T\Phi = I = \Phi\Phi^T$
2. $\Phi^{-1} = \Phi^T$
3. Φx = a roto-inversion operation
4. length preserving : $\|\Phi x\| = \|x\|$

Eigenvalues/vectors

$A \in \mathbb{R}^{n \times n}$ then $v \in \mathbb{C}^n$ is an eigenvector assoc. w/ the e-value $\lambda \in \mathbb{C}$ if

$$Av = \lambda v.$$

Eigen Value Decomposition (EVD)

If X is symmetric and $n \times n$

Can show that X has n mutually ortho-normal e-vects / e-val pairs

$$(\lambda_i, v_i) \quad i = 1, \ldots, n$$

So that $\lambda_i \in \mathbb{R}$,

$$v_i^T v_j = 0 \quad i \neq j$$

$$\|v_i\| = 1.$$

If we let Q be the mtx

$$Q = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}$$

and

$$D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

then

$$X = Q D Q^T.$$

This is called the EVD.

<u>Ex.</u> $X = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$

<u>verify:</u> $v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is e-vec

assoc. w/ $\lambda_1 = 8$.

$$v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \lambda_2 = 2.$$

<u>Claim:</u> $X = Q D Q^T$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 8 & 2 \\ 8 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} = X.$$

Can generalize EVD into SVD (singular value decomp.)

For any matrix X $(m \times n)$

I can decompose it as

$$X = UDV^T$$

where

① U is orthog mtx whose cols are the e-vecs of XX^T

symmetric

② V is orthog. and has cols that are e-vecs of X^TX

③ D is "diag"

r = rank(X)

$$D = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right]_{m \times n}$$

$\sigma_i = \sqrt{\lambda_i}$ where λ_i is the corresp. e-val. of X^TX or XX^T

Lec