Lecture 1: Introduction

Data can be rep. as a mtx

rows = obs.
cols = variable

E.g.

$$X = \begin{array}{c} \begin{bmatrix} 6.1 & 100 & 10 \\ 5.5 & 150 & 20 \\ 7.3 & 200 & 25 \\ 6 & 250 & 75 \end{bmatrix} \\ \begin{array}{ccc} \text{height} & \text{weight} & \text{age} \end{array} \end{array}$$

$N \times P$

N = # obs. (4)

P = # vars. (3)

Can view mtx as a collection of rows

$$X = \begin{bmatrix} \text{---} & x_1 & \text{---} \\ \text{---} & x_2 & \text{---} \\ & \vdots & \\ \text{---} & x_N & \text{---} \end{bmatrix}$$

$x_n \in \mathbb{R}^P$

(lower-case = obs.)

$x_1 = (6.1, 100, 10)$

or as a collection of cols

$$X = \begin{bmatrix} | & | & & | \\ X_1 & X_2 & \cdots & X_P \\ | & | & & | \end{bmatrix}$$

$X_p \in \mathbb{R}^N$

(upper-case = var.)

X_1 = height
$= (6.1, 5.5, 7.3, 6)$

Inner Products

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

$$a \cdot b = a'b = a^T b = \sum_{k=1}^{p} a_k b_k \in \mathbb{R}$$

Norm: the norm/length

$$\|a\| = \sqrt{\sum_{k=1}^{p} a_k^2} = \sqrt{a^T a}$$

What about matrices? must be the same

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

4 ways to define AB:

Sum of Inner products

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

$$= \text{row } i \text{ of } A \cdot \text{col } j \text{ of } B$$

② Linear Combn of cols of A

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

then

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

AB_i = LC of cols of A

③ LC of rows of B

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

④ Sum of Outer Products

$a^T b$ = inner product ← R

ab^T = outer product

mtx

$$AB = \sum_{k=1}^{n} \overbrace{A_k}^{k^{th} \text{ col of } A} \underbrace{b_k^T}_{k^{th} \text{ row of } B}$$

$$\underbrace{m \times 1 \quad 1 \times p}_{m \times p}$$

---

tr = sum of diag. elements

Matrix norm?

Vector: $\|a\| = \sqrt{a^T a} = \sqrt{\sum_{k=1}^{p} a_k^2}$

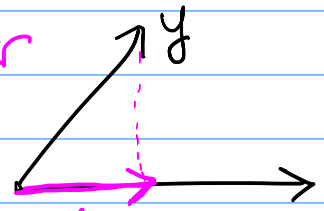
Matrix: $\|A\|_F = \sqrt{\sum_i \sum_j A_{ij}^2} = \sqrt{tr(A^T A)} = \sqrt{tr(AA^T)}$

Projection

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

$$\hat{y} = \underbrace{u_x}\ \underbrace{u_x^T y}_{\mathbb{R} = \text{mag. } y \text{ in dir } x}$$

u_x = unit. vector in dir of x



$\hat{y}$ = proj. y onto x

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

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

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

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

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

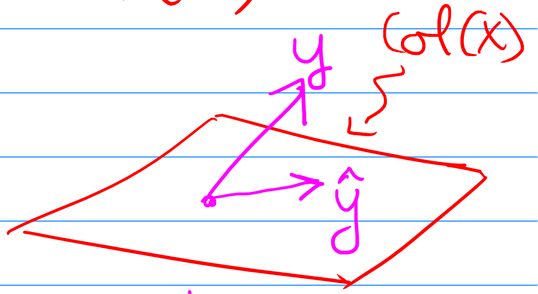
$$= x x^T y / x^T x$$

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

what about matrices?

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

X is a mtx



$$\hat{y} = \underbrace{X (X^T X)^{-1} X^T}_{P_X} y$$

$\hat{y}$ = closest vec to y that is in Col X

Orthoganality

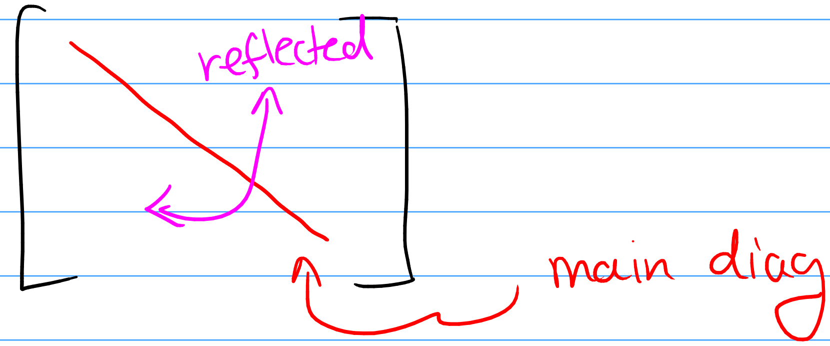
Unit vector v : $\|v\| = 1$

Orthogonal: $u^T v = 0$

both = ortho-normal

Special Matrices

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

$X =$ 

E.g. $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is symmetric

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

Diagonal Matrix:

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

zero everywhere off main diag

$$= \mathrm{diag}(d_1, d_2, \ldots, d_n)$$

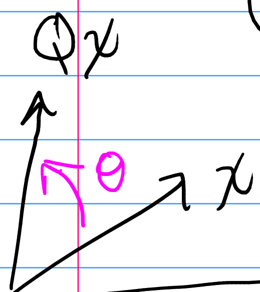
properties:

1. symmetric $D = D^T$
2. $D^{-1} = \begin{pmatrix} 1/d_1 & & & 0 \\ & 1/d_2 & & \\ & & \ddots & \\ 0 & & & 1/d_n \end{pmatrix}$
3. $Dx = (d_1 x_1, d_2 x_2, \dots, d_n x_n)$

Orthogonal Matrix : Q (square)

properties: 1. cols of Q are mutually ortho-normal

2. $Q^T Q = I = Q Q^T$
3. $Q^{-1} = Q^T$
4. Qx = roto-inversion operation



5. length preserving: $\|Qx\| = \|x\|$

Eigen-Values / Eigen-Vectors

If $A \in \mathbb{R}^{n \times n}$ then $v \in \mathbb{C}^n$ is an eigen vector associated w/ eigen-value $\lambda \in \mathbb{C}$ if

$$Av = \lambda v$$

Eigen-Value Decomposition: (EVD)

If X is symmetric and $n \times n$

we can show that X has n mutually ortho-normal e-pairs

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

So that $v_i \in \mathbb{R}^n$; $\lambda_i \in \mathbb{R}$, $\|v_i\| = 1$; $v_i^T v_j = 0$

If

$$Q = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}$$ is $n \times n$, orthogonal

$$D = \text{diag}(\lambda_1, \dots, \lambda_n)$$

then

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

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

verify: $v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is e-vec assoc. w/ $\lambda_1 = 8$

$v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is e-vec. assoc. w/ $\lambda_2 = 2$

Claim $X = QDQ^T$

$$= \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}}_{D} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{Q^T}$$

$$= \left(\frac{1}{2}\right)\begin{bmatrix} 8 & 2 \\ 8 & -2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \left(\frac{1}{2}\right)\begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix}$$

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