Data can (often) be represented as a matrix.

Ex. Data matrix

Columns = variables
rows = observation

$$X = \begin{array}{cc} \begin{bmatrix} 100 & 6.1 & 25 \\ 150 & 5.5 & 45 \\ 326 & 7.3 & 75 \\ 300 & 6 & 30 \end{bmatrix} & \begin{array}{l} \text{per 1} \\ \text{per 2} \\ \text{per 3} \\ \text{per 4} \end{array} \\ \begin{array}{ccc} \text{weight} & \text{height} & \text{age} \end{array} & \end{array}$$

we have $N = 4$ rows and $P = 3$ variables.

View data mtx as collection of rows

$$X = \begin{bmatrix} \text{---} & x_1 & \text{---} \\ \text{---} & x_2 & \text{---} \\ \text{---} & x_3 & \text{---} \\ \text{---} & x_4 & \text{---} \end{bmatrix}$$

x_n is n^{th} observation and $x_n \in \mathbb{R}^P$

Ex. $x_1 = (100, 6.1, 25) \in \mathbb{R}^3$

or as collection of columns

$$X = \begin{bmatrix} | & | & | \\ X_1 & X_2 & X_3 \\ | & | & | \end{bmatrix}$$

X_p is p^{th} variable in $\mathbb{R}^N$

Ex. $X_1 = (100, 150, 320, 300) \in \mathbb{R}^4$

Inner product/Norms (of Vectors)

If $a, b \in \mathbb{R}^p$ then the <u>inner product</u> or <u>dot product</u> is

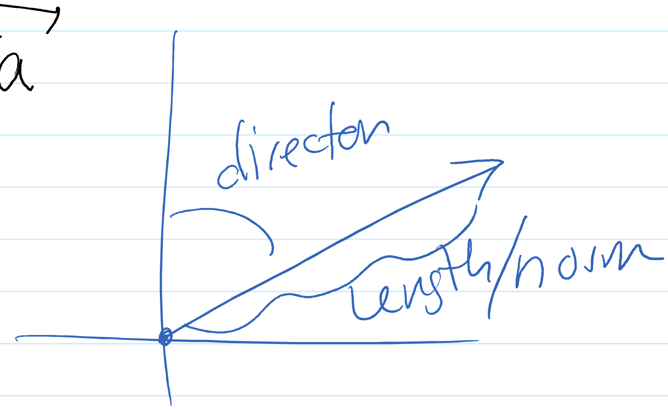
$$a \cdot b = a'b = a^T b = \sum_{k=1}^{p} a_k b_k$$

here $a \in \mathbb{R}^p = \mathbb{R}^{p \times 1}$

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} \text{ and } a^T = [a_1 \; \dots \; a_p]$$

<u>Norms:</u> The norm of a vector is its <u>length</u>

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



<u>What about matrices?</u>

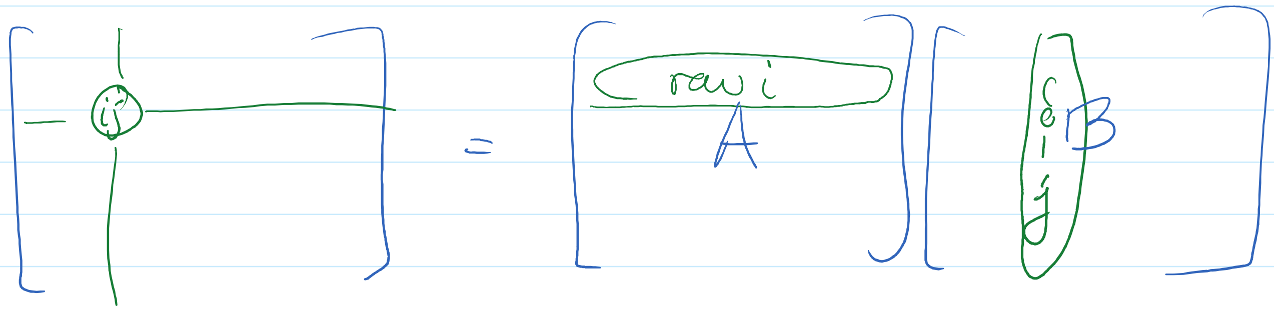
<u>Matrix Products:</u>

inner dimension matches

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

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

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



Another way is if

columns

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

then

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

Matrix Norms

For vecters $\|a\| = \sqrt{a^T a} = \sqrt{\sum_k a_k^2}$

For a matrix $A \in \mathbb{R}^{N \times P}$ then

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

Frobenius Norm

① $A^T \in \mathbb{R}^{P \times N}$ since $A \in \mathbb{R}^{N \times P}$

so $A^T A \in \mathbb{R}^{P \times P}$

② $tr(\cdot)$ is sum of diag elements.

Linear Independence

We say vectors $x_1, \dots, x_p \in \mathbb{R}^N$ are linearly independent if

$$c_1 x_1 + c_2 x_2 + \cdots + c_p x_p = 0$$

then it must be that $c_1 = c_2 = \cdots = c_p = 0$.

Main idea: If vectors are lin. dependent I can write some of them as a LC of others. Not true if vectors are independent.

(independence = no overlapping linear info)

Inverses:

If $A \in \mathbb{R}^{N \times N}$ and $\exists$ a mtx $B \in \mathbb{R}^{N \times N}$ so that

$$AB = BA = I = \mathrm{diag}(1, \dots, 1)$$

then B is called the inverse of A and denoted A^{-1}.

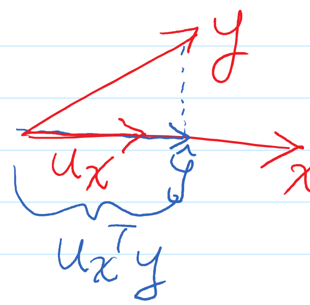
Fact: A has an inverse $\Leftrightarrow$ cols of A are lin. independent.

Projections:

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

$$\hat{y} = \underbrace{u_x}_{\substack{\text{vect. in} \\ \text{dir. of } x}} \underbrace{u_x^T y}_{\substack{\text{amt } y \text{ is} \\ \text{in dir. of } x}}$$

when u_x is a unit vector in direction of x.



Of course:

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

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

So

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

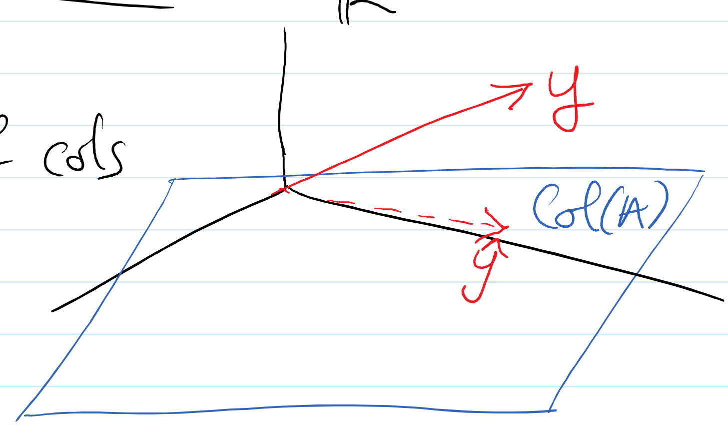
Seems complicated...

What about matrices?

$A \in \mathbb{R}^{N \times P}$

$Col(A)$ = span of cols of A

$\subset \mathbb{R}^N$

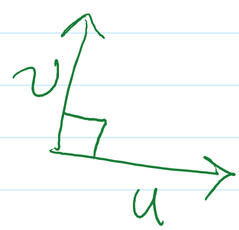


What is the proj. of y onto $Col(A)$?

$$\boxed{\hat{y} = A\underbrace{(A^TA)^{-1}A^T}_{\text{called the projector onto Col}(A)}y}$$

<u>Unit Vectors</u>: $\|u\| = 1$ $u, v \in \mathbb{R}^N$

<u>Orthogonal</u>: $u^Tv = 0$ then $u \perp v$



<u>Orthogonal Matrix</u>:

A matrix $Q \in \mathbb{R}^{N\times N}$ is orthogonal if

→ (1) all cols of Q are unit vectors

→ (2) Cols are mutually orthogonal

i.e. $Q_i^TQ_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

<u>or</u> $Q^TQ = I$ <u>or</u> $Q^{-1} = Q^T$

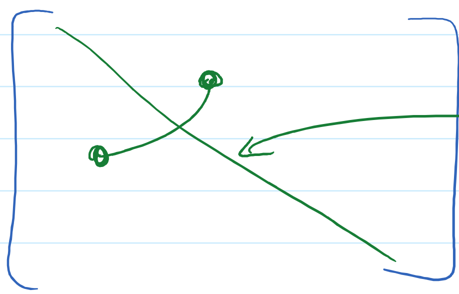
<u>Eigenvectors / Eigenvalues</u>

If $A \in \mathbb{R}^{N\times N}$ then v is an e-vector assoc.

w/ the e-value λ if

$$\boxed{Av = \lambda v.}$$

We say a matrix is symmetric if $\underline{A_{ij} = A_{ji}}$



reflection/ Symmetry about main diag

Ex.

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & 3 \\ 0 & 3 & 7 \end{bmatrix}$$

Fact: If $A \in \mathbb{R}^{N \times N}$ symmetric matrix then it has N eigen-pairs

$$(v_1, \lambda_1), (v_2, \lambda_2), \ldots, (v_N, \lambda_N)$$

and the $\lambda_i \in \mathbb{R}$ and the $\{v_i\}$ are mutually orthogonal. (v_is form an orthogonal basis of $\mathbb{R}^N$)

Spectral Decomposition of a matrix

Spectral Decomposition of a matrix

eigen-pairs

If A is symmetric let

$$\Phi = [v_1 \ \cdots \ v_N] \in \mathbb{R}^{N\times N} = \text{mtx of e-vectors (assume they are unit vectors)}$$

So Φ is orthogonal mtx.

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

$$= \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix} \in \mathbb{R}^{N\times N}$$

Fact: (spectral decomp).

$$\boxed{A = \Phi \cdot D \Phi^T.}$$

Generalize: Singular Value Decomposition (SVD)

Let $A \in \mathbb{R}^{m\times n}$ then

T

$$\underset{m\times n}{A} = \underbrace{U}_{n\times n}\underbrace{D}_{m\times n}\underbrace{V'}_{n\times n}$$

U is orthogonal and its columns are called the left singular vectors and are a basis of the Col space of A

V is orthog., cols are basis of row space of A and called right singular vectors

D is "diag"

$$D = \left[\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \end{array}\right]$$

d_i are called singlar values

where $r = \text{Rank}(A)$

Fact: Want to get U, D, V

(1) cols of U are e-vectors of AA^T

(2) col of V " A^TA

(3) $d_i = \sqrt{\lambda_i}$ e-values of A^TA or AA^T