Linear Algebra

Deep learning AI

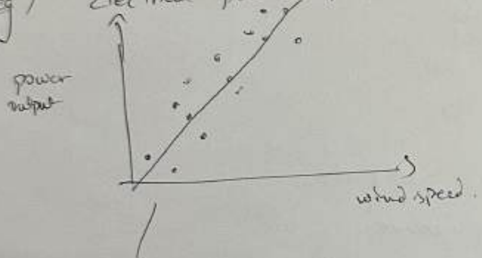
MATHEMATICS FOR ML

LINEAR ALGEBRA

Systems of linear equations

Linear regression → supervised → you have data so you have INPUTS/OUTPUT

eg/ Electrical power o/p from a wind turbine



Clearly theres a pattern
goal of linear regression would be to find the line of best fit for this data.

With a model like this you are ~~estab~~ making the assumption that this relationship is literally linear that it can be modeled by a line.

In other words if you know the wind speed you can multiply it by constant add a second constant & make a reasonable assumption on the expected power o/p

$$m \times \text{wind speed} + b = \text{power output}$$

This model is not perfect but does a good job.

This represents a linear equation.

$y = mx + b$

goal is to find best value of m, b that fit the data

In ML = $y = \underbrace{w}_{\text{weight}} x + \underbrace{b}_{\text{bias}}$

Now if we take two inputs

wind speed, temperature

the equation becomes $y = w_1 \underbrace{x_1}_{\text{wind speed}} + w_2 \underbrace{x_2}_{\text{temp.}} + b$

this is a plane in 3D

Now what if you want to add new features

the concept is the same as with 1 or 2

you simply add a new weight for each new feature.

This model is not perfect but does a good job.

This represents a linear equation.

$$y = mx + b$$

goal is to find best value of m, b that fit the data

in ML = $y = \underset{\text{weight}}{w}x + \underset{\text{bias}}{b}$

Now if we take two inputs

wind speed, temperature

the equation becomes $y = \underset{\text{wind speed}}{w_1 x_1} + \underset{\text{temp.}}{w_2 x_2} + b$

this is a plane in 3D

Now what if you want to add new features

the concept is the same as with 1 or 2

you simply add a new weight for each new feature.

w_1 feature 1 + w_2 feature 2 + ... + b = output.

$w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b = y$ — target

↳ so if you have this for a row for a dataset
• you know w's & y's compute to get target

how will the subqueries look

$$
\begin{aligned}
&— \quad w_1 x_1^{(1)} + w_2 x_2^{(1)} + w_3 x_3^{(1)} + \dots w_n x_n^{(1)} + b^{(1)} = y^{(1)} \\
&— \quad w_2 x_2^{(2)} + w_2 x_2^{(2)} + w_3 x_3^{(2)} + \dots w_n x_n^{(2)} + b^{(2)} = y^{(2)} \\
&— \quad w_1 x_1^{(3)} + w_2 x_2^{(3)} + w_3 x_3^{(3)} + \dots w_n x_n^{(3)} + b^{(3)} = y^{(3)} \\
&— \quad \vdots \\
&— \quad w_1 x_1^{(m)} + w_2 x_2^{(m)} + w_3 x_3^{(m)} + \dots w_n x_n^{(m)} + b^{(m)} = y^{(m)}
\end{aligned}
$$

each row a subquery.

↳ This is a system of linear equations.

x is different, y is different

$w_1, w_2, w_3, \dots w_n$ & b (constant)

This can be represented as

$$[w_1\ w_2\ w_3\ w_4 \ldots w_n] \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & x_4^{(1)} & \ldots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} & x_4^{(2)} & \ldots & x_n^{(2)} \\ x_1^{(3)} & x_2^{(3)} & x_3^{(3)} & x_4^{(3)} & \ldots & x_n^{(3)} \\ \vdots \\ x_1^{(m)} & x_2^{(m)} & x_3^{(m)} & x_4^{(m)} & \ldots & x_n^{(m)} \end{bmatrix} + b \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{bmatrix}$$

W $\qquad X$

$$= [y^{(1)}\ y^{(2)} \ldots y^{(m)}]$$

Y

$$WX + b = \hat{y}$$

Q) Linear Algebra $\to L$

Calculus $\to C$

Prob & Stats $\to P$

$(L+C) - P = 6$

$(L-C) + 2P = 4$

$(4L-2C) + P = 10$

$a + c - p = 6$

$a - c + 2p = 4$

$4a - 2c + p = 10$

~~4a 2c 4p 2 2~~

$(L+C) - P = 6$

$(L-C) + 2P = 4$

$L + \cancel{C} - P = 6$

$L - \cancel{C} + 2P = 4$

$L + P = 10 \quad L = 10 - P$

$\cancel{4a}$ $[a + c = 10]$

System 1

dog black
cat orange

two sentences
two pieces of info.
as much info as sentences.

System 2

dog black
dog black

[Same sentence redundant]

only 1 piece of information

System 3

dog is black
dog is white

sentences contradict each other

contradictory system.

The more info a system carries the more useful it is to you.

redundant & contradictory systems -> singular - not optimal.

complete system -> non-singular.

most informative

Sys 1

dog black
cat orange
bird red

complete

non-singular

2

dog black
dog "
bird red

redundant

singular

3

dog black
dog "
dog "

redundant

singular

4

dog black
dog white
bird red

contradictory.

singular.

Q) ~~apple~~

Price of apple & banana is 10 \$

$a + b = 10$

Day 1 → apple & banana cost 10 \$

" 2 → " " 2 " " 12 \$.

$a + b = 10$

$a + 2b = 12$

$10 - b + 2b = 12$

$\underline{b = 2}$ $a = 8$

Q) Day 1 → apple banana cherry = 10

2 → " 2 " " = 15

3 → " " 2 " = 12

$a + b + c = 10$

$a + 2b + c = 15$

$a + b + 2c = 12$

$a = 10 - b - c$

$10 - b - \cancel{c} + 2b + \cancel{c} = 15$

$\underline{b = 5}$

$a = 5 - c$ $\boxed{a = 3}$

$5 - c + 5 + 2c = 12$

$10 + c = 12$

$\boxed{c = 2\$}$

$a+c-p=6$ $\quad$ $4a-2c+p=10$

$a-c+2p=4$

$[a+p=10]$ $[p=10-a]$

$4a-2c+p \quad -2a+2c-4p=2$

$2a-3p=2$

$2a-(30-3a)=2$

$2a+3a=32$

$5a=32$

$a=32/5 \rightarrow 6.4$

$a=6.4$,

$c-p=-0.4$

$c=p-0.4$

$6.4-p+0.4+2p=4$

$p=4-6.8$

$p=-2.8$

$6.4+c+2.8=6$

$6.4+2.8=9.2$

$[c=-3.2]$

X

Systems of Linear Equations → sentences combine to give meaning.

System of sentences

Q1 $a+b=10$

$2a+2b=20$

$\underline{a=b}$

~~b = 10 - a~~

[redundant]

Not enough info.

Q1 $\underbrace{a+b=10, \quad 2a+2b=24}$

[Contradictory!]

1st Question → $a=8$, $b=2$

non singular.

2nd a, b Infinite solutions [Redundant] singular

3rd contradictory Contradictory

$a+b=5$

Quiz Q1) $a+b+c=10$

$a+b+2c=15$ $c=5$

$a+b+3c=20$.

a & b can be anything that adds upto $\overline{5}$

redundant

Q-2)

$a+b+c=10$) $c=5$

$a+b+2c=15$) $c=7$

$a+b+3c=18$

$a+b=2$ values

contradictory

Q3)

$a+b+c=10$

$2a+2b+2c=20$

$3a+3b+3c=30.$

Infinite

In a linear equation

→ multiply values by scalars

→ add constants

Non Linear

a^2+b^2+10

$\sin(a)+b^5=15$

$2^a-3^b=0$

$ab^2 + \frac{b}{a} - 3/b - \log(c) = 49$

Visualizations

2 variables 2D

3 3D

4 4D

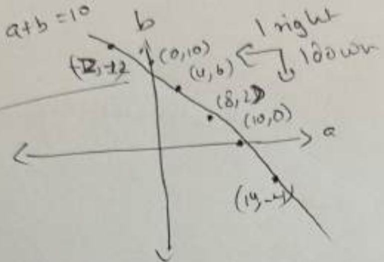
⋮ ⋮

If a linear equation is a line/plane

system of " " is multiple lines/planes.

y intercept = 10

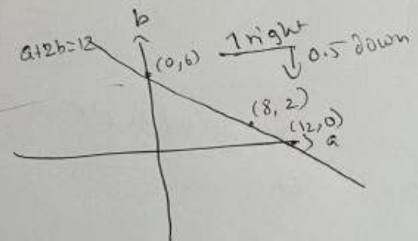
y intercept = value of y when x = 0



→ slope = -1

(because for every 1 movement to right 1 down)

y intercept = 6



slope = -0.5

every 1 right 0.5 down

Now if you merge these to 2-D models

where the intersection happens is the price

$a+b=10$ $\quad a=8$ $\quad$ (8, 2)

$a+2b=12$ $\quad b=2$

If you plot ~ $a+b=10$, $2a+2b=20$

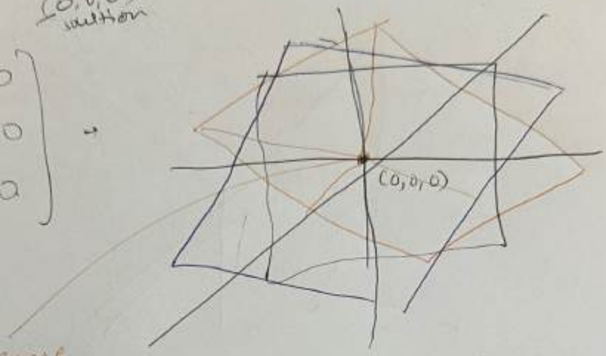
same graph

Now every single point on the two lines which overlap is a solution so we have ∞ solutions.

(0,0,0) solution

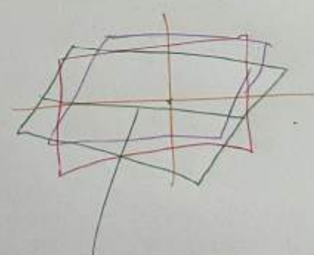
1

$$\begin{bmatrix} a+b+c=0 \\ a+2b+c=0 \\ a+b+2c=0 \end{bmatrix} \rightarrow$$



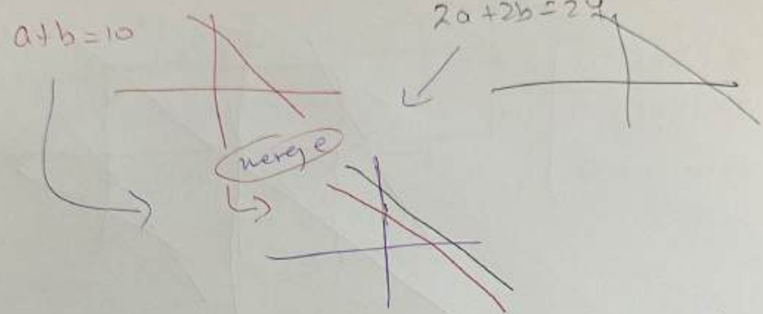
All these 3 planes give one sol = (0, 0, 0) → they intersect there because these equations are giving that unique intersection

$$\begin{bmatrix} a+b+c=0 \\ a+b+2c=0 \\ a+b+3c=0 \end{bmatrix}$$ redundant



These 3 meet on a line they traverse through a line together so multiple solution

↳ ∴ singular.



parallel lines never intersect so no solution.

Quiz!!

$3a+2b=8$

$2a-b=3$

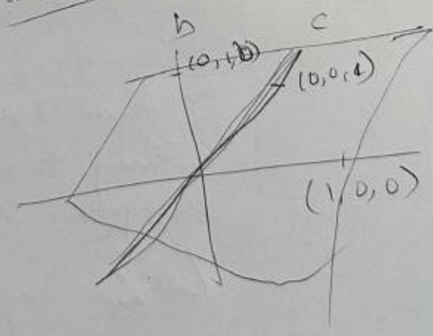
$7a=14$

$\boxed{a=2}$

(a) Non singular

$b=1$

Q2) $a+b+c=1$

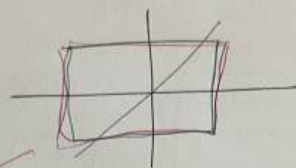


— the entire plane that traverses this is the set of solutions.

3] $a+b+c=0$

$2a+2b+2c=0$

$3a+3b+3c=0$



all three same

so every point on plane is a solution

so ∴ it is singular.

$y = mx + c$

$5 = -5m + c$ $-5 = 5m + c$

~~10~~

Quiz

Q1) $3x + y = 1100$

$x + 3y = 1050$

$y = 1100 - 3x$

$x + 3300 - 9x = 1050$

$3300 - 1050 = 8x$

$8x = 3250$

$x = \frac{\cancel{3250}}{\cancel{8}}$ ~~1625~~ 80 ~~406.25~~ 281.25

~~1218.75~~

$843.75 + y = 1100$

$y = 256.25$

Sys	Plot
3	1
1	2
2	3

How to check for Matrix

Singular

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

row 2 depends on row 1 so dependent

Non

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

row 2 is independent of row 1

$$a = 1$$
$$b = 2$$
$$a + b = 3$$

this can be written as

$$a + 0b + 0c = 1$$
$$+ \quad 0a + b + 0c = 2$$
$$\overline{a + b + 0c = 3}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

row 1 + row 2 = row 3

row 3 depends on rows 1 & 2

[so singular]

$$a + b + c = 0$$
$$2a + 2b + 2c = 0$$
$$3a + 3b + 3c = 0$$

$$a + b + c = 0$$
$$+ \quad 2a + 2b + 2c = 0$$
$$\overline{3a + 3b + 3c = 0}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

row 1 + row 2 = row 3

row 3 depends on rows 1 & 2

rows are linearly dependent.

eg

$$a+b+c=0$$
$$a+b+2c=0$$
$$a+b+3c=0$$

$$\begin{array}{r} a+b+c=0 \\ a+b+3c=0 \\ \hline 2a+2b+4c=0 \end{array}$$

$$\frac{\underbrace{2a+2b+4c=0}}{2} = a+b+2c=0$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \Rightarrow \text{singular}$$

row 2 is the avg of 1^{st} & 3^{rd} row.

Avg of row 1 & 3 = row 2

row 2 depends on rows 1 & 3.

rows are linearly dependent.

eg

$$a+b+c=0$$
$$a+2b+c=0$$
$$a+b+2c=0$$

- no relations b/w equations.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

no relations b/w rows.

rows are linearly independent.

$$a+b=0$$
$$b=0$$
$$3a+2b+3c=0$$

or you can use the determinant

Non	Singular

$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \rightarrow$ singular if $\begin{bmatrix} a & b \end{bmatrix} * k = \begin{bmatrix} c & d \end{bmatrix}$

$$ak = c$$
$$bk = d$$
$$\frac{c}{a} = \frac{d}{b} = k$$
$$ad = bc$$

determinant of matrix $\boxed{ad - bc = 0}$

if $ad - bc = 0$ then singular → dependent.

3X3 → $\begin{bmatrix} a & b & c \\ d & e & f \\ i & g & h \end{bmatrix}$

$a\left(\det\begin{pmatrix} e & f \\ g & h \end{pmatrix}\right)$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\mathrm{DET}(A) = a_{11} \times \mathrm{DET}\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12}\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

~~For~~

~~det(A) = $a_{11}C_{11}$~~ Then

$$a_{11}(a_{33} \times a_{22} - a_{32} \times a_{23}) - a_{12}(a_{21} \times a_{33} + a_{23} \times a_{31}) + a_{13}(a_{21}a_{32} \cancel{a_{22}a_{31}} - a_{22}a_{31})$$

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} + a_{14}C_{14}$$

$$C_{ij} = (-1)^{i+j} \cdot \det(M_{ij})$$

$M_{ij} \rightarrow$ minor matric i.e. the 3×3 matrix left when you remove the i-th row & j-th column from A.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

$$\det(A) = a_{11}\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & . & \\ a_{42} & . & . \end{vmatrix} - a_{12}\begin{vmatrix} & \end{vmatrix} \begin{vmatrix} & \end{vmatrix}$$

$$+ a_{13}\begin{vmatrix} & \end{vmatrix} - a_{4}\begin{vmatrix} & \end{vmatrix}$$

Q-1

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 3 & 3 \end{bmatrix}$$

$$1\begin{pmatrix} 1 & 0 \\ 3 & 3 \end{pmatrix} - 0 + 1\begin{pmatrix} 0 & 1 \\ 3 & 3 \end{pmatrix}$$

$$1(3) + 1(-3)$$

$$3 - 3$$

$$= 0$$

(dependent singular)

$$\begin{bmatrix} 1 & 2 & 5 \\ 0 & 3 & -2 \\ 2 & 4 & 10 \end{bmatrix} \rightarrow (30+8) - 2(4) + 5(6)$$

$$38 - 8 + 30$$

$$60$$

Q-

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & 0 & -1 \end{bmatrix}$$

$$1(-1) + 1(-1) + 0$$

$$-1 \; -1$$

$$= 0$$

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

$$1(6-0)$$

$$0$$

$$3(2-0)$$

$$6$$

Quiz !!

$$A = \begin{bmatrix} 2 & 5 \\ 2 & 4 \end{bmatrix}$$

$8 - 6 = 2$

Q1)

$4x + 3y = 6$

$x - 5y = 8$

$20x + 15y = 30$

$3x - 15y = 24$

$23x = 54$

$\cancel{4x} + 3y = 6$

$-\cancel{4x} + 20y = -32$

$23y = -26$

$y = -26/23$

$x = 54/23$

9.3913

Q2)

$4x + 3y + z = 6$

$x - 5y + 7z = 8$

$5x - 2y + 8z = 14$

$2a + 5b = 46$

$8a + b = 32$

$-8a - 20b = -184$

$-19b = -152$

$b = 38/3$

$2a + 5b = 46$

$8a + b = 32$

$2a + 5b = 46$

$8a + b = 32$

$40a + 5b = 160$

$-2a - 5b = -46$

$38a = 114$

$\boxed{a = 3}$

$5a + b = 11$
$10a + 2b = 22$

$$\begin{cases} 4x + 3y + z = 6 \\ x - 5y + 7z = 8 \\ 5x - 2y + 8z = 14 \end{cases}$$

$$\begin{array}{r} \cancel{4x} + 3y + z = 6 \\ - \cancel{4x} + 20y - 28z = -32 \\ \hline \end{array}$$

$\boxed{23y - 27z = -26}$ $=$ so 1 is redondant.

$$\begin{array}{r} \cancel{5x} - 2y + 8z = 14 \\ - \cancel{5x} + 25y - 35z = -40 \\ \hline \end{array}$$

$\boxed{23y - 27z = -26}$

$23y - 27z = -26 \rightarrow$ ~~zy~~

$x - 5y + 7z = 8$

$y = \frac{-26 + 27z}{23}$ $\rightarrow x - 5\left(\frac{-26 + 27z}{23}\right) + 7z = 8$

$x + \frac{130 - 135z}{23} + 7z = 8$

$23x + 130 - 135z + 161z = 184$

$23x + 130 + 26z = 184$

$23x = 54 - 26z$

$x = \frac{54 - 26z}{23}$

$5a + b = 11$

$10a + 2b = 22$

$$\begin{cases} 4x + 3y + z = 6 \\ x - 5y + 7z = 8 \\ 5x - 2y + 8z = 14 \end{cases}$$

$$\begin{array}{r} \cancel{4x} + 3y + z = 6 \\ - \cancel{4x} + 20y - 28z = -32 \\ \hline \end{array}$$

$\boxed{23y - 27z = -26}$ $=$ so 1 is redundant

$$\begin{array}{r} \cancel{5x} - 2y + 8z = 14 \\ - \cancel{5x} + 25y - 35z = -40 \\ \hline \end{array}$$

$\boxed{23y - 27z = -26}$

$23y - 27z = -26 \rightarrow \cancel{z}y$

$x - 5y + 7z = 8$

$y = \frac{-26 + 27z}{23} \rightarrow x - 5\left(\frac{-26 + 27z}{23}\right) + 7z = 8$

$x + \frac{130 - 135z}{23} + 7z = 8$

$23x + 130 - 135z + 161z = 184$

$23x + 130 + 26z = 184$

$23x = 54 - 26z$

$x = \frac{54 - 26z}{23}$

Systems of Equations to matrices

① Original System

$$5a + b = 17$$
$$4a - 3b = 6$$

original matrix

$$\begin{bmatrix} 5 & 1 \\ 4 & -3 \end{bmatrix}$$

② Intermediate system

$$a + 0.2b = 3.4$$
$$b = 2$$

upper diagnol matrix

$$\begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix}$$

[row echelon form]

③ Solved system

$$a = 3$$
$$b = 2$$

Diagnol matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

[reduced row echelon form]

Original system

$$a + b = 10$$
$$2a + 2b = 20$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Intermediate

$$a + b = 10$$
$$0a + 0b = 0$$

Row Echelon Form

$$\begin{bmatrix} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The main diagnol should be a bunch of 1's followed by bunch of 0's it can be all 1's or all 0's on the left of diagnol everything must be zero.

on the right of the 1's (*) the *'s can be any number & right of 0's only 0's

In 2X2 Eg

$$\begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & * \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Row Operations to maintain singularity,

$A = \begin{bmatrix} 5 & 1 \\ 4 & 3 \end{bmatrix}$ Det (A) => 15 - 4 -> 11 Non-Singular

1st operation:

Switch rows

$\begin{bmatrix} 5 & 1 \\ 4 & 3 \end{bmatrix} \rightarrow B = \begin{bmatrix} 4 & 3 \\ 5 & 1 \end{bmatrix}$ Det (B) => 4 - 15 = -11 $\neq 0$

multiply row by a (non-zero) scaler

$\begin{bmatrix} 5 & 1 \\ 4 & 3 \end{bmatrix}$ $\begin{bmatrix} 5 & 1 \end{bmatrix} \times 10 = \begin{bmatrix} 50 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 50 & 10 \\ 4 & 3 \end{bmatrix}$

det => 110 $\neq 0$

(Scaler should be non-zero.)

Add a row to another row.

$$\begin{array}{|c|c|}\hline 5 & 1 \\ \hline 4 & 3 \\ \hline\end{array} + \begin{array}{|c|c|}\hline 5 & 1 \\ \hline 4 & 3 \\ \hline\end{array} \to \begin{array}{|c|c|}\hline 9 & 4 \\ \hline 4 & 3 \\ \hline\end{array}$$ Det's 11 $\neq 0$

REF

A matrix is in row echelon form if it looks like a staircase going down to the right.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 5 & 6 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

- Every leading number (first non zero no. in each row) is to the right of the one above it.
- All entries below each leading number are zero.
- Any row of all zeroes, if it exists is at the bottom.

What it looks like:

$$1x + 2y + 3z = \text{something}$$
$$0x + 1y + 5z = \text{''}$$
$$0x + 0y + 1z = \text{''}$$

[leading 1 step to right
zero beneath each leading 1]

Why?
Solve for last variable.
plug z into second to get y
plug y & z into first to get x.

How Gaussian elimination works.

RREF

matrix is in RREF if:

1) Should be in REF
2) Every leading entry is 1 (first non-zero no. in a row)
3) Every ~~lea~~ leading 1 is the only non-zero value in it's column.
4) any row of all 0's is at the bottom.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \end{array}\right]$$

(So like storing fully solved linear equations)

Quiz!!

1) $x + y = 4$
$-6x + 2y = 16$
$(y = 5)$

$-6x + \cancel{2y} = 16$
$-2x - \cancel{2y} = -8$

$-8x = 8$
$(x = -1)$

2) $A = \begin{bmatrix} 4 & -3 \\ 7 & -8 \end{bmatrix}$ $\det(A) \Rightarrow -32 + 21$
$\Rightarrow -11$

3) $\begin{bmatrix} -3 & 8 & 1 \\ 2 & 2 & -1 \\ 5 & 6 & 2 \end{bmatrix} \Rightarrow -3(4+6) - 8(4-5) + 1(12+10)$

$-30 + 8 + 22 = 0$

4) $\begin{bmatrix} a & b & c \\ d & e & f \\ 2a-d & 2b-e & 2c-f \end{bmatrix}$

Row echelon Form 3D - 4D

$$\begin{aligned} a + b + 2c &= 12 \\ 3a - 3b - c &= 3 \\ 2a - b + 6c &= 24 \end{aligned} \longrightarrow \begin{aligned} a + b + 2c &= 12 \\ -6b - 7c &= -33 \\ 6c &= 18 \end{aligned}$$

MATRIX

1	1	2
3	-3	-1
2	-1	6

$\longrightarrow$

R.E.F

1	1	2
0	-6	-7
0	0	6

$\downarrow$

$$\begin{matrix} 1 & 1 & 2 \\ 1 & -1 & -1/3 \\ 1 & -1/2 & \end{matrix}$$

These are valid REF's (Both) $\Rightarrow$

the 3 non zero pivots are on the right of row above

3	*	*	*	*
0	0	1	*	*
0	0	0	4	*
0	0	0	0	0
0	0	0	0	0

you can also divide $\rightarrow$ by 1st non zero coefficient of row.

Rank for matrices

Matrix 1 | matrix 2 | matrix 3 | matrix 4

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

$\Downarrow$ REF

~~Subtract First row~~

$$\begin{bmatrix}1&1&1\\0&1&0\\0&0&1\end{bmatrix} \quad \begin{bmatrix}1&1&1\\0&0&1\\0&0&0\end{bmatrix} \quad \begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix} \quad \begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

REF

Matrix 1

$$\begin{bmatrix}1&1&1\\1&2&1\\1&1&2\end{bmatrix} \xrightarrow{\text{subtract 1st from 2nd \& 3rd}} \begin{bmatrix}1&1&1\\0&1&0\\0&0&1\end{bmatrix}$$

2:

$$\begin{bmatrix}1&1&1\\1&1&2\\1&1&3\end{bmatrix} \xrightarrow{\text{subtract 1st from 2nd \& 3rd}} \begin{bmatrix}1&1&1\\0&0&1\\0&0&2\end{bmatrix} \xrightarrow{\text{subtract 2* second row from 3rd}} \begin{bmatrix}1&1&1\\0&0&1\\0&0&0\end{bmatrix}$$

3)

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

So Again → Rank = how many 1's or non zeros in the diagonal / or how many **pivots**!!!

3 pivots rank=3 $\begin{bmatrix}1&1&1\\0&1&0\\0&0&1\end{bmatrix}$

$\begin{bmatrix}1&1&1\\0&0&1\\0&0&0\end{bmatrix}$ – 2 pivots rank = 3

$\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}$ 1 pivot rank = 1

Reduced Row Echelon Form:

Original system	Intermediate system	Solved system
$5a + b = 17$	$a + 0.2b = 3.4$	$a = 3$
$4a - 3b = 6$	$b = 2$	$b = 2$

matrix

$$\begin{bmatrix} 5 & 1 \\ 4 & -3 \end{bmatrix}$$

upper diagonal matrix

$$\begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix}$$

row echelon form

Diagonal matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$1a + 0b = 3$

$0a + 1b = 2$

reduced row echelon form

How to go from REF to RREF

$$\begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix} \text{ row1} - \text{row2} \times 0.2 \Rightarrow \begin{array}{rr} 1 & 0.2 \\ -\ 0 & 0.2 \\ \hline 1 & 0 \end{array} \Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

[Left most coefficient of row = pivot!!]

RREF General Form:

1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

=5

1	[illegible]	0	0	[illegible]
0	0	1	0	[illegible]
0	0	0	1	[illegible]
0	0	0	0	0
0	0	0	0	0

=3

- It has to be in row echelon form
- It's pivots have to be 1
- Any number above pivot has to be 0.
- rank = no. of pivots

General method to go from REF to RREF.

$$\begin{bmatrix} 3 & * & * & * & * \\ 0 & 0 & 2 & * & * \\ 0 & 0 & 0 & 4 & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Divide each row by pivot coefficient}} \begin{bmatrix} 1 & * & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

REF

Now you use each pivot's 1 to clear out any number above it

eg/

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{remove 2}} \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{remove 4 \& 5}} \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

multiply 3rd row by 5 & add to 1st
multiply 3rd row by 4 & subtract from 2nd.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The Gaussian Elimination ~~Proc~~ Algorithm

$$\left.\begin{aligned} 2a - b + c &= 1 \\ 2a + 2b + 4c &= -2 \\ 4a + b &= -1 \end{aligned}\right]$$

In the matrices before the we did not consider these values.
to actually solve a system of equations we must pay attention to them.

Augmented matrix

$$\begin{bmatrix} 2 & -1 & 1 & 1 \\ 2 & 2 & 4 & -2 \\ 4 & 1 & 0 & -1 \end{bmatrix}$$

To start change coefficient in 1st row to 1 so multiply by $\frac{1}{2}$

$$\left[\begin{array}{ccc|c} 2 & -1 & 1 & 1 \\ 2 & 2 & 4 & -2 \\ 4 & 1 & 0 & -1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -1/2 & 1/2 & 1/2 \\ 2 & 2 & 4 & -2 \\ 4 & 1 & 0 & -1 \end{array}\right] \Rightarrow \begin{array}{l} R_2 \leftarrow R_2 - 2R_1 \\ R_3 \leftarrow R_3 - 4R_1 \end{array}$$

$$\Rightarrow \begin{array}{cccc} 1 & -1/2 & 1/2 & 1/2 \\ 0 & 3 & 3 & -3 \\ 0 & 3 & -2 & -3 \end{array} \rightarrow \begin{array}{cccc} 1 & -1/2 & 1/2 & 1/2 \\ 0 & 1 & 1 & -1 \\ 0 & 3 & -2 & -3 \end{array} \quad R_3 \leftarrow R_3 - 3R_2$$

$$\begin{array}{cccc} 1 & -1/2 & 1/2 & 1/2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -5 & 0 \end{array} \quad R_3 \leftarrow R_3/5 \rightarrow \begin{array}{cccc} 1 & -1/2 & 1/2 & 1/2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{array}$$

REF

Now $R_2 \leftarrow R_2 - R_3$ &

$$\begin{array}{cccc} 1 & -1/2 & 1/2 & 1/2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array} \xrightarrow{R_1 \leftarrow R_1 + 1/2 R_2} \begin{array}{cccc} 1 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}$$

$R_1 \leftarrow R_1 - 1/2 R_3$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

RREF

$a = 0,\ b = -1$

$c = 0$

In an augmented with only 1's in the diagonal

This is called identity matrix

After gaussian reduction if one row is zero then it's singular.

Week 3 Let's recap how this works in ML →

You have

Feature 1	feature 2	.....	Feature n		target
$w_1 x_1^{(1)}$ +	$w_2 x_2^{(1)}$ +	....	+ $w_n x_n^{(1)}$	+ b^1 =	$y^{(1)}$
$w_1 x_1^{(2)}$ +	$w_2 x_2^{(2)}$ +	.....	+ $w_n x_n^{(2)}$	+ b^2 =	$y^{(2)}$
⋮	⋮		⋮		⋮
m					

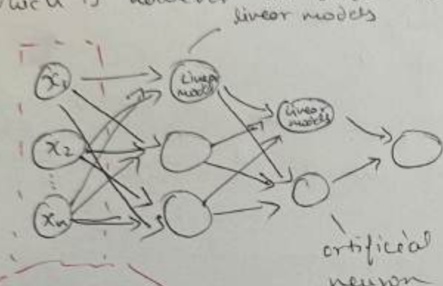
Where you have $x_1, x_2, \ldots x_n$ & you need to figure out w values & values of b.

This becomes →

$$W \cdot X + b = \hat{y}$$

You can only get so far with this because as it turns out most features consist of non linear relationships

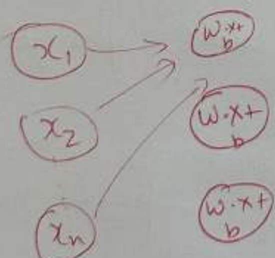
One of the most powerful ML non linear model is Neural Network which is however all under the hood are linear models



inputs to network/ features

so you send each of these x values to each individual neuron in the next layer.

so all x's go to first neuron they will have a w value & b value associated with it.



but actually it has a whole system of linear equations that is represented by $wx + b$.

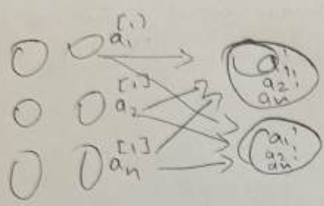
↓

What happens in the neuron the whole $wx+b$ is placed in an activation function → $\sigma(w \cdot x+b)$

this generates an output → a vector a_1

Each neuron does this & generates unique vectors → $a_1, a_2, a_3 \ldots$

Then you pass a values to the next layer.



this causes x's to be replaced by a's
so now ~~...~~ it has a matrico of all a's of previous layer

then you multiply all that with a different weight & adding a difference bias value

this gives

$$\sigma(W_1^{[2]} \cdot A^{[1]} + b_1^{[2]}) \to a_1^{[2]}$$

$$\sigma(W_2^{[2]} \cdot A^{[1]} + b_2^{[2]}) \to a_2^{[2]}$$

this then keeps happening until you get desired output.

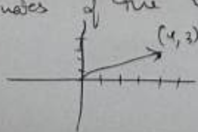
Vectors

single array of numbers, arrows in a plane.

simply a tuple of numbers -> It could be 2 or 3 numbers.

the coordinates of the vector & the dimension in which it lives.

for eg

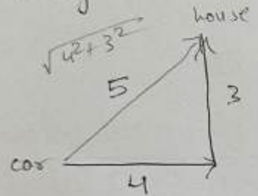


could be in more D.

Magnitude/size & direction → important components

Magnitude / Size

This emulates the distances you use in real life.
basically the displacement.



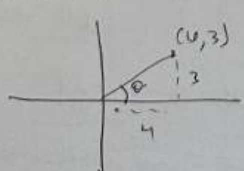
$4+3=7$ → taxicab distance or L1-norm.

L1-norm $= |(a,b)|_1 = |a| + |b|$ (taxicab)

L2-norm $= |(a,b)|_2 = \sqrt{a^2+b^2}$ (vector)

We use L2 because precisely it is the length of the arrow.

direction can be deduced from coordinates.



$\tan\theta = 3/4$ $\theta = \arctan(3/4) = 0.64$
$= 36.87°$

Vector Notation → row vector → $x = [x_1, x_2 \ldots x_n]$

Column vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Now lets see L1 & L2 norms with any vector of n total components

$$x = (x_1, x_2 \ldots x_n)$$

L1 norm

$$||x||_1 = |x_1| + |x_2| + \ldots + |x_n|$$

L2 Norm

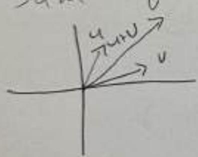
$$||x||_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

Vector Operations

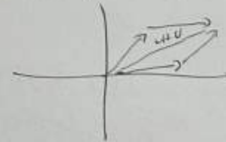
Sum of vectors.

$u = 1,3$
$v = 4,1$

$u+v = (5,4)$

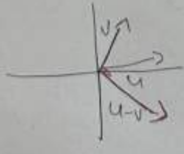


u+v is the diagnol of the parallelogram

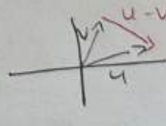


Difference

$u - v = u - v$



If you move u-v up



matches precisely with the vector obtained by joining (1,3) & (4,1) or u & v

General definition

lets say $x = [x_1\ x_2\ x_3 \ldots x_n]$ $y = [y_1\ y_2\ y_3 \ldots y_n]$

For both sum & Difference x & y must have same number of components.

Sum | Difference
Sum component by component | subtract component by component

Distance b/w two vectors.

$v = (1,5)$ $u = (6,2)$

L_1 distance → $|u-v|_1 = |5| + |-3| = 8$

L_2 distance = $|u-v|_2 = \sqrt{5^2+3^2} = 5.83.$

Multiply vector by a scalar. → General definition

$x = (x_1, x_2, x_3 \dots x_n)$

$\lambda x = (\lambda x_1, \lambda x_2 \dots \lambda x_n)$

$u = (1,2)$ $\lambda = 3$

$\lambda u = (3,6)$

Scalar negative?

$u = 1,2)$ $\lambda = -2$

$\lambda u = -2, -4$

Dot product

Quantities
2 apples
4 bananas
1 cherry

Prices
apples → 3$
bananas → 5$
cherries → 2$

total price

vector $\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}$ × vector $\begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix}$ ⇒

$2 \times 3 = 6$
$4 \times 5 = 20$ = $6+20+2 = 28$
$1 \times 2 = 2$

This is the dot product.

$$\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix} = 2\cdot3 + 4\cdot5 + 1\cdot2 = 28$$

⇓ more common

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

Vector transpose

It converts columns to rows ^ vice versa, how we did before

$$\begin{matrix}2\\4\\1\end{matrix} = \begin{matrix}2 & 4 & 1\end{matrix} \qquad \begin{matrix}2 & 4 & 1\end{matrix} \rightarrow \begin{matrix}2\\4\\1\end{matrix}$$

Matrix transpose

$$\begin{bmatrix}2 & 5\\4 & 7\\1 & 3\end{bmatrix} \qquad \begin{bmatrix}2 & 4 & 1\\5 & 7 & 3\end{bmatrix}$$

$3\times2 \rightarrow 2\times3$

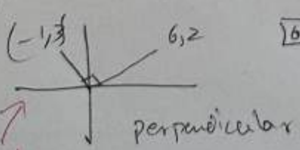
General Definition "DOT PRODUCT"

$x = (x_1\ x_2 \dots x_n) \qquad y = (y_1\ y_2 \dots y_n)$

$$x \cdot y = (x_1 \times y_1) + (x_2 \times y_2) + \dots + (x_n \times y_n)$$

Geometric Dot Product

Angles are important. There is a relation b/w angles & dot product



Known as Orthogonal Vectors

$\begin{bmatrix}6 & 2\end{bmatrix} \begin{bmatrix}-1\\3\end{bmatrix} = \boxed{0} \qquad \langle u, v \rangle = 0$

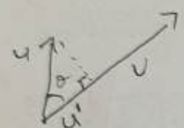
dot product notation

If dot product of vector = 0 Orthogonal Vectors

$\langle u, u \rangle = |u|^2$ (dot product of same vector)

$\langle u, v \rangle = 0$

so what about the dot product of



$\langle u, v \rangle = |u'| \cdot |v|$

$= |u||v| \cos\theta$

vector $(6,2)$, $(-1,3)$

$\rightarrow$ 6 2 -1 3 $\rightarrow 0$

multiply Matrix & vector

$a + b + c = 10$

$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 10$

$a + 2b + c = 15$

$\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 15$

$a + b + 2c = 12$

$\begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 12$

Colum Vector is the same

So this can be represented as : $\underset{3\times3}{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}} \cdot \underset{3}{\begin{bmatrix} a \\ b \\ c \end{bmatrix}} = \begin{bmatrix} 10 \\ 15 \\ 12 \end{bmatrix}$ # colum must equal the length of vector!!!

$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 12 \\ 13 \end{bmatrix}$ Length = 4 $\rightarrow$ this is valid

Quiz!!! Q1) $u = 1,3$ $\quad v = 6,2$ $\quad u+v = 7,5$

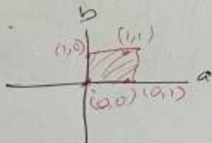
2) $u - v = -5,1$

Q4) $\vec{a} = \begin{bmatrix} -1 \\ 5 \\ 2 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} -3 \\ 6 \\ -4 \end{bmatrix}$ $\vec{a}.\vec{b} \to 3 + 30 + -8 \to 25$

Matrices as linear transformation

A way to send a point in the plane to another in a very structured way.

$$\begin{array}{cc} a & b \\ \end{array}$$
$$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$



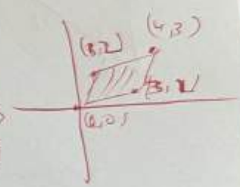
$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = (0,0)$

$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \to (1,2)$

$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \to (3,1)$

$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \to (4,3)$

$\to$



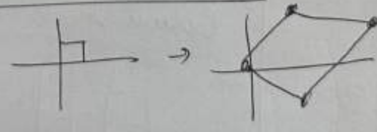
So the square on the left becomes the parallelogram on right

theses are called bases

This can be expanded across the whole plane.

Linear transformation as matrix

$$\begin{array}{cc} a & b \\ \end{array}$$
$$\begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$$



$(0,0) \to (0,0)$

$(1,0) \to (3,-1)$ $\Big\}$ these 2 suffice

$(0,1) \to (2,3)$

$(1,1) \to (5,2)$

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

$\begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{matrix} 3 \\ -1 \end{matrix}$

the (0,1) is the b value → $b = \boxed{\begin{matrix}2\\3\end{matrix}}$

& (1,0) is the a value $a = \boxed{\begin{matrix}3\\-1\end{matrix}}$

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

Matrix multiplication

combining two linear transformation into a 3rd one.

$$\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 \\ +1 & 2 \end{bmatrix} \rightarrow$$

~~2×3+-1×-1 2×2+-1×3~~
~~0×3+2×-1 0×2+2×3~~

$$2\times3 + -1\times1 \qquad 2\times1 + -1\times2$$
$$0\times3 + 2\times1 \qquad 0\times1 + 2\times2$$

$$\rightarrow \begin{bmatrix} 5 & 0 \\ 2 & 4 \end{bmatrix}$$

eg

$$\begin{bmatrix} 3 & 1 & 4 \\ 2 & -1 & 2 \end{bmatrix} \quad \begin{bmatrix} 3 & 0 & 1 & -2 \\ 1 & 5 & 2 & 0 \\ -2 & 1 & 4 & 0 \end{bmatrix} \rightarrow$$

$$3\times3 + 1\times1 + 4+-2 \qquad 3\times0+1\times5+-1\times1 \qquad 3\times1+1\times2+4\times4 \qquad 3\times-2+1\times0+0$$
$$2\times3+-1\times1+2\times-2 \qquad 2\times0+-1\times5+2\times1 \qquad 2\times1+-1\times2+2\times4 \qquad 2\times-2+0$$

$$9+1-8 \qquad 5-1 \qquad 3+2+16 \qquad -6$$
$$6+-1+-4 \qquad -5+2 \qquad 2-2+8 \qquad -4$$

$$\begin{bmatrix} 2 & 4 & 21 & -6 \\ 1 & -3 & 8 & -4 \end{bmatrix}$$

Identity Matrix

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}a\\b\\c\\d\end{bmatrix}=\begin{bmatrix}a\\b\\c\\d\end{bmatrix}$$

Dot product with identity matrix gives same result (basically 1)

Matrix Inverse

$\text{inverse}(x) \cdot x = 1$

$\text{Inverse}(x) = 1/x$ so $\text{inverse}(2) = 1/2$

so the inverse of a matrix is the matrix that if multiplied by it gives the inverse matrix

$$\begin{bmatrix}3&1\\1&2\end{bmatrix}\cdot\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix}3&1\\1&2\end{bmatrix}^{-1} \rightarrow \begin{matrix}3a+c=1\\ \cancel{3b+d=0}\end{matrix}$$

Q) $\begin{bmatrix}5&2\\1&2\end{bmatrix}$ Inverse? $\begin{bmatrix}5&2\\1&2\end{bmatrix}\cdot\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$

$a = 2/8$

$\boxed{a = 1/4}$

$\begin{bmatrix}1/4&-1/4\\-1/8&5/8\end{bmatrix}$

$5a + 2c = 1$

$5b + 2d = 0$

$1 - 5/4 = -1/4$

$a + 2c = 0$

$b + 2d = 1$

$b = -1/4$

$b = 1 - 2d$

$5 - 10d + 2d = 0$

$5 = 8d$

$d = 5/8$

$a = -2c$

$-10c + 2c = 1$

$-8c = 1$

$\boxed{c = -1/8}$

Q1 $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow$ $a + c = 1$, $2a + 2c = 0$, $b + d = 0$, $2b + 2d = 1$, $a = c - 1$, $b = -d$

No Inverse.

So which matrices have an Inverse.?

$5^{-1} = 0.2$ $8^{-1} = 0.125$ $0^{-1} = ????$

Non-singular matrices always an inverse so they are invertible

singular -> non-invertible.

Check by determinants.

Neural Networks On Matrices

-> Natural language processing.

'Lottery' & 'win' appear to be in spam ~~table~~ emails so you build this table.

Classifier -> you assign score to lottery & win

Spam	Lottery	win
	1	1
yes	2	1
yes	0	0
No	0	2
yes	0	1
No	1	0
yes	2	2
yes	2	0
yes	1	2

Scores

Lottery: — points

win: — points

lets say 3 & 2

email says win win the lottery
 2 2 3

total = 7 points.

Rule -> If the number of points > —

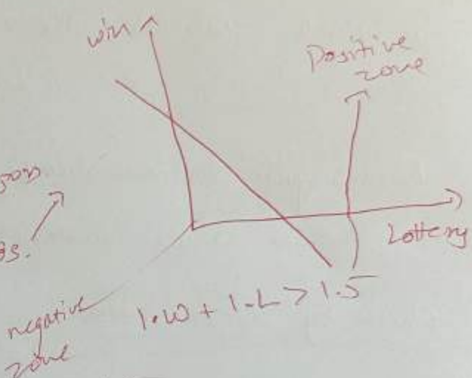
then email spam.

using the table in previous page.
What should be the course distribution.
Lottery → 1 Win → 1 threshold → 1.5

Score	>1.5?
2	yes
3	yes
0	no
2	yes
1	no
1	no
4	yes
2	yes
3	yes

→ matches result So classifier did a good job.

This is NLP as input was words.



$1 \cdot W + 1 \cdot L > 1.5$

$1 \cdot W + 1 L < 1.5$

This is a Neural Network with 1 layer.

Spam	Lottery	Win
Y	1	1
Y	2	1
N	0	0
Y	0	2
N	0	1
N	1	0
Y	2	2
Y	2	0
Y	1	2

Model $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ check > 1.5?

① take second row

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

$3 > 1.5$

$\boxed{\text{so spam}}$

② take 5th row

$\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1$

Not spam.

Product of matrix & vector gives us →

$$\begin{array}{c} \text{Pro} \\ \begin{bmatrix} 2 \\ 3 \\ 0 \\ 2 \\ 1 \\ 1 \\ 4 \\ 2 \\ 3 \end{bmatrix} \end{array}$$

Check >1.5?

until now we were checking

$1 \cdot W + 1 \cdot L > 1.5$

threshold. $\underbrace{1 \cdot W + 1 \cdot L - 1.5}_{\text{bias}} > 0$

$1 \cdot W + \ldots$

$\Downarrow$

How do we check with bias. Introduce new column

Spam Lottery Win Bias

1

1

1

1

1

1

1

model → $\begin{bmatrix} 1 \\ 1 \\ -1.5 \end{bmatrix}$

If you multiply it's same.

Q) The AND operator

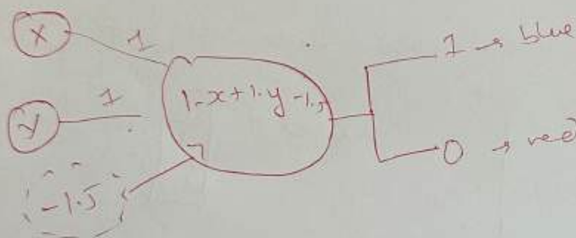
AND	x	y
No	0	0
No	1	0
No	0	1
Yes	1	1

model $\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{array}{c} \text{Dot prod} \\ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} \end{array}$ check >1.5 ⇒ $\begin{bmatrix} \text{No} \\ \text{No} \\ \text{No} \\ \text{Yes} \end{bmatrix}$

So this can be modeled as a layer in a neural N

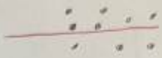
Same graph

$1 \cdot x + 1 \cdot y = 1.5$



Principal Component Analysis (PCA)

Now if we look from the pov of the line, it looks like something like this



• dataset

now you wanna simplify your db a bit - perhaps all points a close & lie along the line

Now PCA would be at the line, & what it would do is to simplify the dataset by imagining this line to be the space projecting all the points on the line

So PCA allows you to go from a 2-D dataset to a 1-D dataset that carried almost the same amount of information. That's why it's called dimensionality reduction algorithm.

TLDR

So PCA is a technique used in data science applications to reduce the dimensions of a dataset while losing as little data as possible.

You can imagine a dataset of many dimensions of data, overcomplicating & difficult to visualise.

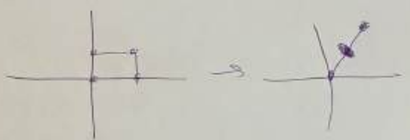
PCA intelligently reduces the no. of columns

Singularity & rank of linear transformation

Singular

$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \to \begin{matrix} 0 \\ 0 \end{matrix}$ $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} (1,2)$ $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \to (1,2)$ $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \to (2,4)$

So the linear transformation would look like:



this is a degenerate parallelogram → this is also called as a line segment.

dimension → 1 rank = 1

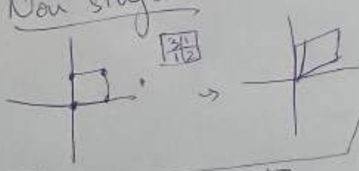
The grid above (boxes) does not ~~get~~ go to a parallelogram.
Its actually a line segment.
This does not cover the entire plane on the right.

A parallelogram that is thin can still cover the entire plane however not if it is flat then it only covers a line segment.

$$\boxed{\begin{array}{c|c} 0 & 0 \\ \hline 0 & 0 \end{array}} \rightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

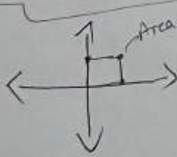
dimension 0 → rank 0

Non singular



→ 2 dimension → rank = 2

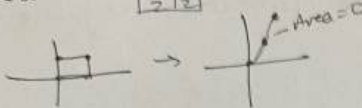
Determinant as an area



$$\begin{array}{|c|c|} \hline 3 & 1 \\ \hline 1 & 2 \\ \hline \end{array}$$

Det = 3·2 − 1·1
Det = 5

Now for $\begin{bmatrix}1&1\\2&2\end{bmatrix}$

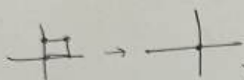


Det = 1.2 − 1.2

Det = 0.

Area of line = 0

$\begin{bmatrix}0&0\\0&0\end{bmatrix}$



Det = 0.

→ the area of a point = 0

Now for $\begin{bmatrix}1&3\\2&1\end{bmatrix}$ → det = 1.1 − 3.2 = $\underline{\underline{-5}}$

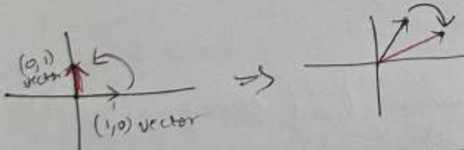
that means

$\begin{bmatrix}1&3\\2&1\end{bmatrix}$ ~~(0,0) → (0,0)~~ ~~(0,1) → 1×0+3×1~~

$\begin{bmatrix}1&3\\2&1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} \rightarrow \begin{bmatrix}1\times1+3\times0\\2\times1+1\times0\end{bmatrix} \rightarrow \begin{bmatrix}1\\2\end{bmatrix}$

$\begin{bmatrix}1&3\\2&1\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} \rightarrow \begin{bmatrix}1\times0+3\times1\\2\times0+1\times1\end{bmatrix} \rightarrow \begin{bmatrix}3\\1\end{bmatrix}$

So (1,0) → (1,2)
(0,1) → (3,1)



⇒ The vectors have gone from ↺ to ↻ so switched along the axis & hence −ve

Determinant of a product.

$\begin{bmatrix}3&1\\1&2\end{bmatrix} \cdot \begin{bmatrix}1&1\\-2&1\end{bmatrix} \rightarrow \begin{bmatrix}3\times1+1\times-2 & 3\times1+1\times1\\1\times1+2\times-2 & 1\times1+2\times1\end{bmatrix}$

det = 5 det = 3

$\rightarrow \begin{bmatrix}1&4\\-3&3\end{bmatrix}$ → det → 15 ← 5×3 = 15 ?

coincidence ???

No!! this is always true

det (AB) = det (A). det (B)

If one singular ~~zero~~ ~~also~~ product singular.

area $\times 0 = 0$.

Determinant of Inverse

Quize $\det \begin{bmatrix} 0.4 & -0.2 \\ -0.2 & 0.6 \end{bmatrix} \rightarrow 0.6 \times 0.4 - 0.04$

$0.24 - 0.04$

$\rightarrow 0.2$

$\det \begin{bmatrix} 0.25 & -0.25 \\ -0.125 & 0.625 \end{bmatrix}$ $0.15625 - 0.03125$

$\rightarrow 0.125$

this $\Rightarrow \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$ the det of this

$\rightarrow = 5$

$5^{-1} = 0.2$

$\begin{bmatrix} 5 & 2 \\ 1 & 2 \end{bmatrix} \Rightarrow \det = 8$

$8^{-1} = 1/8 = 0.125$

$$\det(A^{-1}) = \frac{1}{\det(A)}$$ [If matrix is invertable]

$\det(AB) = \det(A)\det(B)$

$\det(AA^{-1}) = \det(A)\det(A^{-1})$

[AA^{-1} = Identity matrix] $\det(I) = \det(A) \cdot \det(A^{-1})$

[Identity matrix = 1] $1 = \det(A) \cdot \det(A^{-1})$

$\det(A) = 1/\det(A^{-1})$

Quiz!

1) $\begin{bmatrix}1 & 0\\2 & 3\end{bmatrix}$ rank = ? $= 2$

2) T maps (1,0), (0,1)

$T(0,1) = (2,5)$

$T(1,0) = (3,1) \rightarrow \begin{bmatrix}3 & 2\\1 & 5\end{bmatrix}$

13

3) $(-1)\cdot(-2)\cdot(-2) \rightarrow 4$

5) $\begin{bmatrix}0 & 0 & 1\\2 & 2 & 1\\1 & 0 & 0\end{bmatrix} \rightarrow M$ find $\det(M^{-1})$ $\quad \det(M^{-1}) = \frac{1}{\det(M)}$

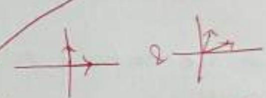
100%

$\det(M) = 0\left(\det\begin{pmatrix}2 & 1\\0 & 0\end{pmatrix}\right) - 0\left(\det\begin{pmatrix}2 & 1\\1 & 0\end{pmatrix}\right) + 1\left(\det\begin{pmatrix}2 & 2\\1 & 0\end{pmatrix}\right)$

$= 0 - 0 + (1)\cdot(-2)$

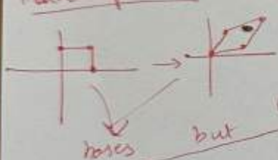
$= -2 \rightarrow |-2| = \underline{-2}$

$\det(M^{-1}) = \frac{1}{-2}$



Bases

Plural of "basis"



bases but why? It's because of these! the two vectors that arise from the origin important

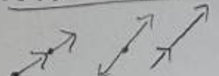
main property: Every point in the space can be expressed as a linear combination of elements in the basis.

So

this means you can reach any point in the space with this vectors

Pretty much any two vectors form a basis. You can make microscopial moves or even move backwards in that direction.

What is NOT a basis?



Anything that has two vectors in same or opposite directions on the same line cannot be a basis.

Determinants & Eigen vectors (Span in linear Algebra)

What we saw was basis of a plane, can something be a bases for another space not a plane

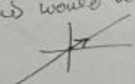
"Span" → this is a set of vectors is simply the set of points that can be reached by walking in the direction of these vectors in any combination.

→ spans whole plane → span is whole line they fall on.

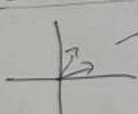
→ this is not a basis however → NO → basis needs to be a minimal spanning set.

Here there are too many vectors. any of these two vectors span the line so that's one too many.

→ this would be a basis (for the line)



let's see in 2D



they span the plane

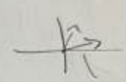
span the plane but one too many vectors as 2 do the job

No. of elements in the basis is the dimension

Linearly Independent & dependent vectors

Independent vectors → none of the vectors in the group can be obtained as a linear combination of the others.

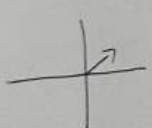
1 vector → always linearly independent

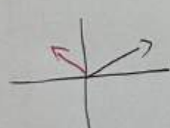


Independent
they point in different directions. so you can't get any as a linear combination of the other.

The new vector in red is twice the one in black in the same direction.
so since one vector can be obtained as a linear combination of the others.

Notice, the span remains the same when the red one is added





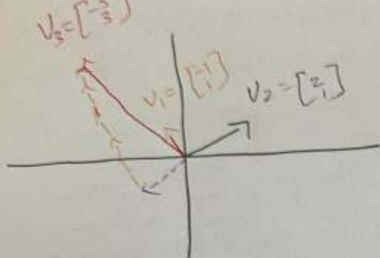
Independent → same

Not lmao. why though???? omg

Because the green vector can be obtained by the linear combination of the other two vectors silly. This results in general.

If you have more vectors & dimension of the span you are trying to span, you will always have a linearly dependent group

How to Check Boo (linear dependence)



$\alpha \cdot v_1 + \beta v_2 = v_3$

In other words you are looking for the coefficients α & β which give you v_3.

Q $\alpha \begin{bmatrix}-1\\1\end{bmatrix} + \beta \begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}-5\\3\end{bmatrix}$

$-\alpha + 2\beta = -5$

$\alpha + \beta = 3$

$\alpha = 3 - \beta$

$-3 + \beta + 2\beta = -5$

$3\beta = -2$

$\beta = -2/3$

$\alpha = 3 - \beta$

$\alpha = 11/3$

Q) $\begin{bmatrix}1\\-1\\0\end{bmatrix}$ $\begin{bmatrix}0\\-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\\-1\end{bmatrix}$

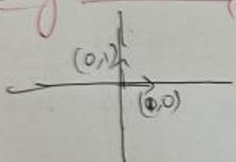
$\begin{bmatrix}1\\-1\\0\end{bmatrix} + -1 \begin{bmatrix}0\\-1\\1\end{bmatrix} = \begin{bmatrix}1\\0\\-1\end{bmatrix}$

$\alpha = 1, \ \beta = -1$

Basis: Formal definition

A basis is a set of vectors that:
- Spans a vector space &
- Is linearly independent.

Eigenbasis →



⇒

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

$2\times1-1\cdot0$
$0\times1-3\times0$

→ $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \to \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

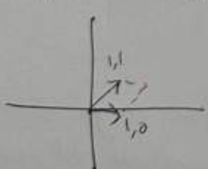
$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow$

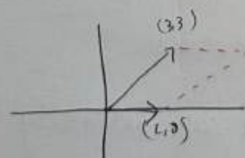
$2\times1+1\cdot1$
$0\times1+3\times1$ $\Rightarrow \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

→ this is from square coordinates on left to the parallelogram on right.

Instead of taking $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ let's take $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

So (1,0) & (1,1) now use $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ for transformation → (1,0) goes to (2,0)
(1,1) " " (3,3)





⇒ So one parallelogram goes to the other parallelogram

The sides of the two parallelogram are parallel.

We are stretching horizontal by ×2
diagnol by ×3] eigen values

vectors → eigen vectors.

eg $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 \\ 0 \end{bmatrix} \Rightarrow 2\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ } eigen vectors.

$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 \\ 3 \end{bmatrix} \quad 3\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

this helps scale up vectors

eigen vector $\rightarrow v_1$ matrix $\rightarrow A$

so $A v_1 = \lambda_1 v_1$

for special values of v_1 this equation holds true

these are eigen vectors.

$$A_2 v_2 = \lambda_2 v_2$$

What's important in this.

matrix multiplication | scalar multiplication.
more work | less work

$$A_1 v_1 = \lambda_1 v_1$$
$$A_2 v_2 = \lambda_2 v_2$$

8 multiplications | 2 multiplications

I get it now!! so you can literally transform any vector into a linear combination of the eigen vectors & transform all matrix multiplications into scalar.

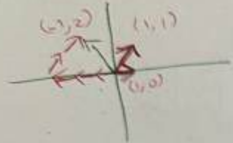
eg/ $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ we can transform this into a linear combination - $\begin{bmatrix} -1 \\ 2 \end{bmatrix} \Rightarrow -3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

so now $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \left[-3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right] \Rightarrow -3\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Done by scalar multiplication

$\Rightarrow -3\left[2\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right] + 2\left[3\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right] \Rightarrow \begin{bmatrix} -6 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$

This will look like this:



so $(-1,2) \Rightarrow -3(1,0) + 2(1,1)$

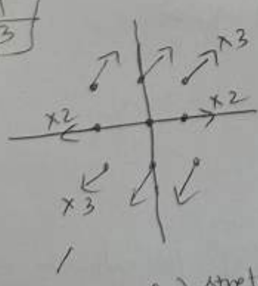
So linear transformation is done without any matrix multiplication.

Main takeaways

- $Av = \lambda v$ for each eigenvector / eigenvalue.
- Eigenvectors: direction of stretch
- Eigenvalues: how much stretch.
- Eigen Basis: the set of a matrix's eigenvectors, can be arranged as a matrix with one eigenvector in each column.
- Save work & help characterise a linear transformation.

Calculating Eigenvalues & Eigenvectors

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

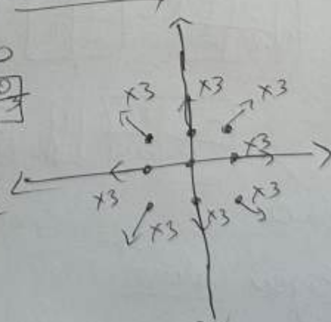


the horizontal is stretched by 2 & diagnols by 3.

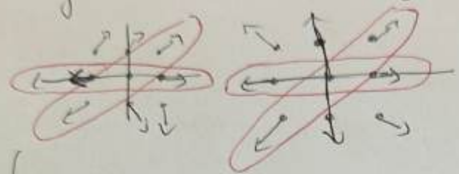
now

$$\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

really 3×3 Identity matrix



These two transformations are not the same transformation but they do coincide on many points.



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

They act the exact same way for infinitely many points. All the points on these lines.

So lets take this first. Since the two transformations act the same they can be represented as:

$$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$ for infinitely many points

That is strange. transformations should only match at one point that is (0,0) [non singular]. When they match at infinitely many points that is singular

$$\left(\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\right)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ at infinitely many points

$$\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ – singular

$\det = 0$

Finding Eigen values

if λ is an eigen value $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ for infinitely many x, y

$$\begin{bmatrix} 2-\lambda & 1 \\ 0 & 3-\lambda \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Now we saw det is 0 as it is a singular

$$\det \begin{vmatrix} 2-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} = 0$$

$(2-\lambda)(3-\lambda) - 10 = 0$

$\lambda = 2$

$\lambda = 3$ → eigenvalues.

Now save equations

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

$2x + y = 2x$ $\quad x = 1$

$0x + 3y = 2y.$ $\quad y = 0$ → $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ one eigen vector

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

$2x + y = 3x$ $\quad x = 1$

$0x + 3y = 3y$ $\quad y = 1$ $\quad \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Q)

$$\begin{bmatrix} 9 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{bmatrix} 9-\lambda & 4 \\ 4 & 3-\lambda \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

So the determinant ← this is singular

of this is 0.

$(9-\lambda)(3-\lambda) - 16 = 0$

$27 - 9\lambda - 3\lambda + \lambda^2 - 16 = 0$

$\lambda^2 - 12\lambda + 11 = 0$

$\lambda^2 + -11\lambda + -\lambda + 11 = 0$

$\lambda(\lambda - 11) - (\lambda - 11)$

$(\lambda - 1)(\lambda - 11)$

$$\boxed{\begin{array}{l} \lambda = 1 \\ \lambda = 11 \end{array}}$$

$\lambda = 1$

$$\begin{bmatrix} 9 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$9x + 4y = x$ $\quad y = -2x$

$4x + 3y = y$ $\quad \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$4x - 6x = 0$

$x = 0$ $\quad y = 0$

Q)

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & -3 \\ -1 & -3 & 0 \end{bmatrix} \qquad \lambda I = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$\det(A - \lambda I) = 0$$

$$\det\left(\begin{bmatrix} 2-\lambda & 1 & -1 \\ 1 & -\lambda & -3 \\ -1 & -3 & -\lambda \end{bmatrix}\right) = 0$$

lets start with $\lambda = 4$

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & -3 \\ -1 & -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$(2-\lambda)\lambda^2 + 3 + 3 - 9(2-\lambda) + \lambda + \lambda$$

$$\Rightarrow \lambda^3 + 2\lambda^2 + 11\lambda - 12$$

$$\Rightarrow \cancel{\lambda(\lambda^2 + 2\lambda + 11) + 12}$$

$$\Rightarrow \cancel{\lambda^3 + 2\lambda^2 +}$$

$$-(\lambda+3)(\lambda-1)(\lambda-4) = 0$$

$$\underline{\lambda = -3, 1, 4}$$

Non square matrices do not have eigen vector or eigen value.

this has 3 eigen vectors & 3 eigen values.

Do all 3X3 have 3 eigens?

$$A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 4 & 0.5 \\ 0 & 0 & 2 \end{bmatrix}$$

now $\det(A - \lambda I) = \det$

$$(2-\lambda)^2(4-\lambda) + 0 + 0 - 0 - 0 - 0 = 0$$

$$\lambda = \underbrace{2, 2}_{\text{repeated eigen value.}}, 4$$

Now we saw det is 0 as it is a singular

$$\det \begin{bmatrix} 2-\lambda & 1 \\ 0 & 3-\lambda \end{bmatrix} = 0$$

$(2-\lambda)(3-\lambda) - 10 = 0$

$\lambda = 2$
$\lambda = 3$ → eigenvalues.

Now solve equations

$$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} x \\ y \end{bmatrix} \qquad \begin{aligned} 2x + y &= 2x \\ 0x + 3y &= 2y. \end{aligned} \qquad \begin{aligned} x &= 1 \\ y &= 0 \end{aligned} \rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

one eigen vector

$$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 3 \begin{bmatrix} x \\ y \end{bmatrix} \qquad \begin{aligned} 2x + y &= 3x \\ 0x + 3y &= 3y \end{aligned} \qquad \begin{aligned} x &= 1 \\ y &= 1 \end{aligned} \qquad \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Q)

$$\begin{bmatrix} 9 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{bmatrix} 9-\lambda & 4 \\ 4 & 3-\lambda \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

← this is singular

So the determinant of this is 0.

$(9-\lambda)(3-\lambda) - 16 = 0$

$27 - 9\lambda - 3\lambda + \lambda^2 - 16 = 0$

$\lambda^2 - 12\lambda + 11 = 0$

$\lambda^2 + -11\lambda + -\lambda + 11 = 0$

$\lambda(\lambda - 11) - 1(\lambda - 11)$

$(\lambda - 1)(\lambda - 11)$

$$\boxed{\begin{aligned} \lambda &= 1 \\ \lambda &= 11 \end{aligned}}$$

$\lambda = 1$

$$\begin{bmatrix} 9 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$9x + 4y = x$

$4x + 3y = y$

$4x - 6x = 0$

$x = 0 \quad y = 0$

$y = -2x$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Q)

$A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & -3 \\ -1 & -3 & 0 \end{bmatrix}$ $\lambda I = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$

$\det(A - \lambda I) = 0$

$\det\left(\begin{bmatrix} 2-\lambda & 1 & -1 \\ 1 & -\lambda & -3 \\ -1 & -3 & -\lambda \end{bmatrix} \right) = 0$

$(2-\lambda)\lambda^2 + 3 + 3 - 9(2-\lambda) + \lambda + \lambda$

$\Rightarrow \lambda^3 + 2\lambda^2 + 11\lambda - 12$

$\Rightarrow \cancel{\lambda(\lambda^2 + 2\lambda + 11) + 12}$

$\Rightarrow \cancel{\lambda^3 + 2\lambda^2 +}$

$-(\lambda + 3)(\lambda - 1)(\lambda - 4) = 0$

$\underline{\lambda = -3, 1, 4}$

lets start with $\lambda = 4$

$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & -3 \\ -1 & -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Non square matrices do not have eigen vector or eigen value.

this has 3 eigen vectors & 3 eigen values.

Do all 3X3 have 3 eigens?

$A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 4 & 0.5 \\ 0 & 0 & 2 \end{bmatrix}$

now $\det(A - \lambda I) = \det$

$(2-\lambda)^2(4-\lambda) + 0 + 0 - 0 - 0 - 0 = 0$

$\lambda = \underbrace{2, 2}_{\text{repeated eigen value.}}, 4$

Let's start with 4.

$Av = 4v$

$$\begin{matrix} 2 & 0 & 0 \\ -1 & 4 & -0.5 \\ 0 & 0 & 2 \end{matrix} \quad \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} = 4 \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} = \begin{bmatrix} 4x_1 \\ 4x_2 \\ 4x_3 \end{bmatrix}$$

$x_1 = 0$

$x_2 =$ any number $\rightarrow \boxed{\begin{matrix} 0 \\ 1 \\ 0 \end{matrix}}$

$x_3 = 0$

$$= \begin{bmatrix} 2x_1 \\ -x + 4x_2 - 0.5x_3 \\ 2x_3 \end{bmatrix}$$

$2x_1 = 4x_2$

$-x_1 + 4x_2 - 0.5x_3 = 4x_2$

$2x_3 = 4x_3$

$-2x_1 = 0$

$-x - 0.5x_3 = 0 \rightarrow$ No x_2.

$-2x_3 = 0$

Now with 2.

$2x_1 = 2x_1$

$-x_1 + 4x_2 - 0.5x_3 = 2x_2$

$2x_3 = 2x_3$

$0 = 0$

$-1 + 2x_2 - 0.5x_3 = 0$

$0 = 0$

$x_1 = 2x_2 - 0.5x_3 \rightarrow$ infinitely many solutions.

let's say $x_1 = 2$, $x_2 = 1$, $x_3 = 0$ $x_1 = 1$, $x_2 = 1$, $x_3 = 2$

these point in different directions. so different eigen vectors.

So

$\lambda = 4$ $\boxed{\begin{matrix} 0 \\ 1 \\ 0 \end{matrix}}$ $\lambda = 2$ $\boxed{\begin{matrix} 2 \\ 1 \\ 0 \end{matrix}}$ $\lambda = 2$ $\boxed{\begin{matrix} 1 \\ 1 \\ 2 \end{matrix}}$ $\rightarrow$ still 3

So ~~for~~ I was confusing eigen vector with eigen basis.

eigen vectors $\rightarrow \infty$, eigen basis $\Rightarrow$ dimension or <

Summary

$$\begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array}$$

eigen values $\rightarrow \lambda_1, \lambda_2$

if $\lambda_1 \neq \lambda_2$ 2 eigen vectors different directions

if $\lambda_1 = \lambda_2$ — 1 eigen vector 1 direction / 2 eigen vectors 2 directions

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline d & e & f \\ \hline g & h & i \\ \hline \end{array}$$

$\lambda_1, \lambda_2, \lambda_3$

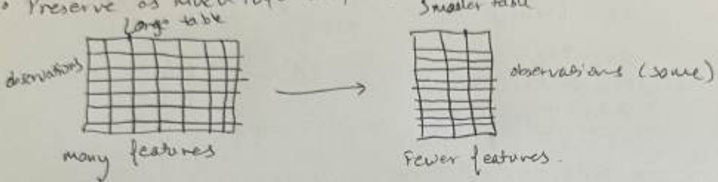
if $\lambda_1 \neq \lambda_2 \neq \lambda_3$ 3 eigenvectors

$\lambda_1 = \lambda_2 \neq \lambda_3$ — 2 eigen 2 / 3 eigen 3 direction

$\lambda_1 = \lambda_2 = \lambda_3$ — 1 / 2 / 3

PCA → Dimensionality Reduction

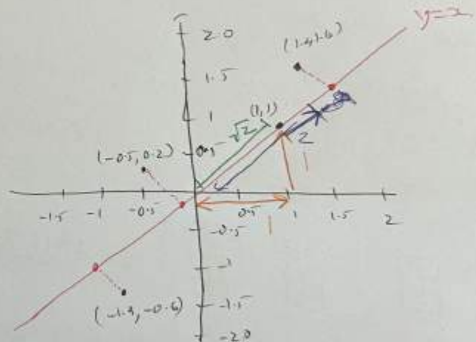
- Reduce dimensions (# of columns) of dataset
- Preserve as much info as possible



Projections

x	y
1	1
1.2	1.6
-0.5	0.2
-1.3	-0.6

looks like →



Now imagine you want to project your data on the line $y=x$

All points move perpendicularly towards the line except (1,1) which is already on the line

Where do these points end up? They're however 2-D coordinates we can now start (1,1) → didn't move, 1,1 is however 2-D coordinates. We can now give it 1 coordinate i.e. distance along the line in the coordinate. That is the distance $\sqrt{2}$.

$\sqrt{1^2+1^2} = \sqrt{2}$ → $\sqrt{2}$ can be written as $1+1/\sqrt{2}$

here the coordinates are showing up

Note: Dot Product is the projection of one matrix or vector onto another.

It leads to a value not a vector.

Matrix multiplication → linear transformation of matrices.

So we use Dot product in PCA

$x = 1,\ y = 1$ $\quad$ $y = x$ so we use $(1, 1)$ as it's the same & it spans to it.

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow 1 + 1 = \underline{\underline{2}}$$

Look in the previous page In this colour on the graph to see what this calculates

So it takes 1 → (horizontal) & 1 (vertical)

However this length = 2 but it's supposed to be $\sqrt{2}$

the blue is 2, so it overshot by a factor of $\sqrt{2}$ so divide by $\sqrt{2}$

Norm of 1

$\cancel{\begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{\sqrt{2}} = (1+1)/\sqrt{2}}$

x	y
1.0	1.0
1.2	1.6
-0.5	0.2
-1.3	-0.6

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{\sqrt{2}} = (1+1)/\sqrt{2}$$

this $= \dfrac{1}{\text{norm} \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$ $\qquad$ - norm = magnitude / length.

main idea for projection → multiplying by the vector projects the points along that vector & dividing by the vector's norm ensures that there is no stretching introduced

For second row.

$\begin{bmatrix} 1.2 & 1.6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ → $(1.2+1.6)$ → too far if you divide by $\sqrt{2}$ will get you too the desired point.

Final coordinates → $\begin{bmatrix} 1.4142 \\ 1.9799 \\ -0.2121 \\ -1.344 \end{bmatrix}$

those points are now reflected on a 1-D line

In general.

to Project a matrix A onto a vector V.

$$A \cdot \frac{V}{\|V\|_2}$$

→ norm of vector v to scale it back a bit
"If I back it up is it far enough".

$$\underset{r\times 1}{A_p} = \underset{r\times c}{A} \underset{c\times 1}{\frac{V}{\|V\|_2}}$$

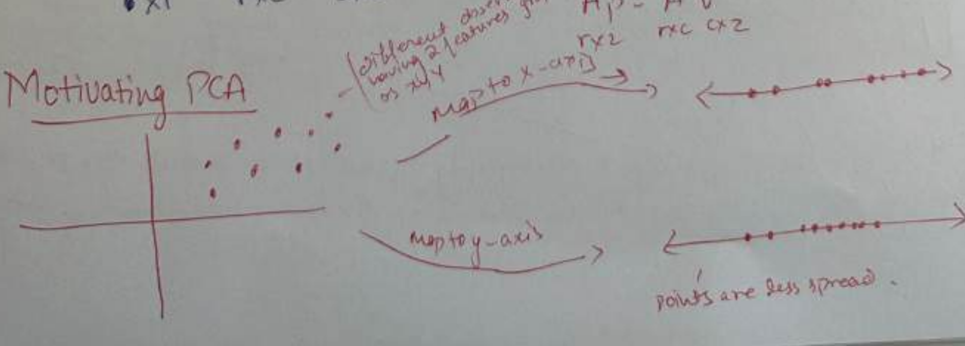
To project a matrix A onto vectors V_1 & V_2

$$\underset{r\times 2}{A_p} = \underset{r\times c}{A} \left[\frac{V_1}{\|V_1\|} \frac{V_2}{\|V_2\|}\right] - c\times 2$$

this can be represented as

$$\underset{r\times 2}{A_p} = \underset{r\times c}{A} \underset{c\times 2}{V}$$

Motivating PCA



You can project onto any line. Find different projections & sort them according to ~~highest~~ lowest to highest spread or vice versa

more spread = more info

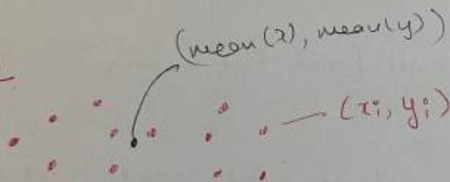
less spread = preserve least info.

So PCA has to find the lower dimension that preserves the maximal possible spread in your data.

Benefits

- Easier to manage datasets.
- PCA reduces dimensions while minimizing information loss.
- Simpler visualization.

Mean



mean = average of data

$$\text{mean}(x) = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{mean}(y) = \frac{1}{n}\sum_{i=1}^{n} y_i$$

The middle coordinate is the mean.

<u>Variance</u> → To see how much data is spread out.

look at them along horizontal & vertical axis

Dataset with no spread has a variance of 0.

" " large " " " large variance

Variance:

$$\text{Variance}(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \text{Mean}(x))^2$$

	x_i	$x_i - \text{Mean}(x)$	$(x_i - \text{Mean}(x))^2$
1	10	1	1
2	4	-5	25
3	11	2	4
4	14	5	25
5	6	-3	9

Sum = 64

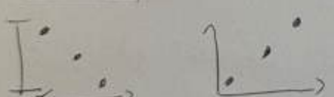
mean(x) = 9.

$$\text{Variance}(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \text{Mean}(x))^2$$

$$\text{Var}(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu_x)^2$$

"The average square distance from the mean"

Problem



→ These both have same spread on x & y - axis but patterns are significantly different.

We use covariance → it helps measures how two features of a dataset varies with respect to one another.

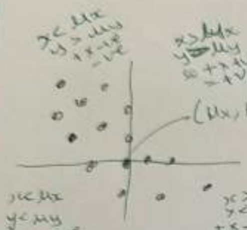
→ In left pattern → down & to the right $x\uparrow y\downarrow$

In right " → up & to the right $x\uparrow y\uparrow$

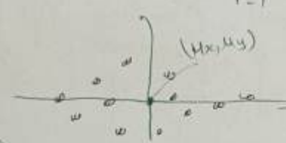
Covariance quantifies the data set to the left having negative covariance & right having positive.

Covariance $\quad cov(x, y) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y)$

$$Var(x) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu_x)^2$$



negative covariance



covariance zero
(or very small)

Positive
covariance.

Subtracting μ_x from each x & μ_y from each y essentially recenters the data around its center to this point splitting it into 4 quadrants each point depending on what quadrant they are in contribut (+ or −ve) to the covariance

$(x_i - \mu_x)(y_i - \mu_y) \rightarrow$ results in +ve or −ve

$\frac{1}{n-1} \sum_{i=1}^{n}$ → takes sum & divides by number of values −1

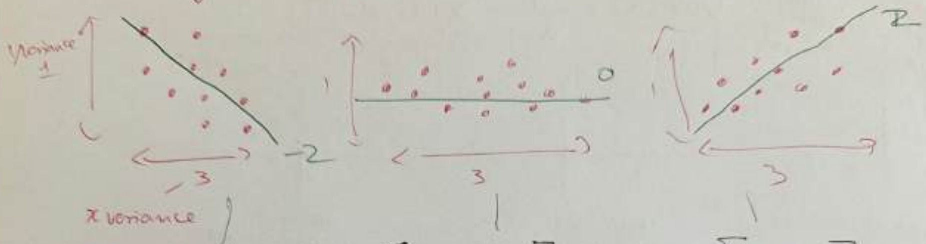
↳ other words averaging the product

So basically saying on average more dots +ve or −ve

Measures the direction of the relationship b/w two variables

Covariance Matrix

Compact way of storing your relationship b/w pairs of variables in your dataset



$$\begin{bmatrix} 3 & -2 \\ -2 & 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$$

First get $\underset{x}{var(x)}$, $\underset{y}{var(y)}$ & $cov(x,y)$

$$C = \begin{array}{c} x \\ y \end{array} \begin{bmatrix} var(x) & cov(x,y) \\ cov(y,x) & var(y) \end{bmatrix}$$

$cov(y,x) = cov(x,y)$

$$\boxed{cov(x,x) = var(x)}$$

$$\text{So } C = \begin{bmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix}$$

$$A = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_x & \mu_y \\ \mu_x & \mu_y \\ \vdots & \vdots \\ \mu_x & \mu_y \end{bmatrix}$$ - same n length.

$$C = \frac{1}{n-1} (A-\mu)^T (A-\mu)$$

$$\Rightarrow \frac{1}{n-1}\left[\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} - \begin{bmatrix} \mu_x & \mu_y \\ \mu_x & \mu_y \\ \vdots & \vdots \\ \mu_x & \mu_y \end{bmatrix}\right]^T \left(\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} - \begin{bmatrix} \mu_x & \mu_y \\ \mu_x & \mu_y \\ \vdots & \vdots \\ \mu_x & \mu_y \end{bmatrix}\right)$$

$$= \frac{1}{n-1}\begin{bmatrix} x_1-\mu_x & y_1-\mu_y \\ x_2-\mu_x & y_2-\mu_y \\ \vdots & \vdots \\ x_n-\mu_x & y_n-\mu_y \end{bmatrix}^T \begin{bmatrix} x_1-\mu_x & y_1-\mu_y \\ x_2-\mu_x & y_2-\mu_y \\ \vdots & \vdots \\ x_n-\mu_x & y_n-\mu_y \end{bmatrix}$$

$$= \frac{1}{n-1}\underbrace{\begin{bmatrix} x_1-\mu_x & x_2-\mu_x & \cdots & x_n-\mu_x \\ y_1-\mu_y & y_2-\mu_y & \cdots & y_n-\mu_y \end{bmatrix}}_{2\times n}\underbrace{\begin{bmatrix} x_1-\mu_x & y_1-\mu_y \\ x_2-\mu_x & y_2-\mu_y \\ \vdots & \vdots \\ x_n-\mu_x & y_n-\mu_y \end{bmatrix}}_{n\times 2}$$

$$\begin{bmatrix} Var(x) & cov(x,y) \\ cov(y,x) & cov(y) \end{bmatrix}$$

$2\times n \cdot n\times 2$

$\rightarrow 2\times 2 \rightarrow$ which is exactly the size of covariance matrix.

$$(x_1-\mu_x)(x_1-\mu_x) + (x_2-\mu_x)(x_2-\mu_x) + \cdots (x_n-\mu_x)(x_n-\mu_x)$$

this gives: $\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu_x)^2 \rightarrow Var(x)$

2nd column $\rightarrow \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu_x)(y_i-\mu_x) \rightarrow Cov(x,y)$

eg

x	y
10	3
12	9
⋮	4
⋮	⋮

A

x	y
10	5
12	3
6	9
6	4
5	11
14	2
8	1
3	13

$\mu_x = 8 \quad \mu_y = 6$

$A - \mu$

$x-\mu_x$	$y-\mu_y$
2	-1
4	-3
-2	3
-2	-2
-3	5
6	-4
0	-5
-5	8

$$C = \frac{1}{n-1}(A-\mu)^T(A-\mu)$$

$(A-\mu)^T$

$\frac{1}{7}$

$x-\mu_x$	2	4	-2	-2	-3	6	0	-5
$y-\mu_y$	-1	-3	3	-2	5	-4	-5	8

$x-\mu_x$	$y-\mu_y$
2	-1
4	-3
-2	3
-2	-2
-3	5
6	-4
0	-5
-5	8

$$C = \begin{bmatrix} 14 & -11.86 \\ -11.86 & 19.71 \end{bmatrix}$$

Matrix Formula:

$$C = \frac{1}{n-1}(A-\mu)^T(A-\mu) \qquad A = \begin{bmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix}$$

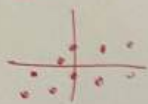
1) Arrange data with different feature in each column.
2) Calculate column averages.
3) Subtract each average from their respective column to generate $A-\mu$
4) $\frac{1}{n-1}(A-\mu)^T(A-\mu)$ gives covariance matrix C

Eigen value / vector Recap

Let's say you have a transformation matrix A. It stretches, rotates, reflects or squashes space.

Eigenvectors are the directions in space that don't rotate under this transformation. They just get scaled. Eigen values are the scaling factor applied to the eigen vectors.

PCA → combines projections, eigen, covariance matrix.



let's say $C = \begin{bmatrix} 9 & 4 \\ 4 & 3 \end{bmatrix}$ So you find the eigenvector & eigen values

direction

magnitude

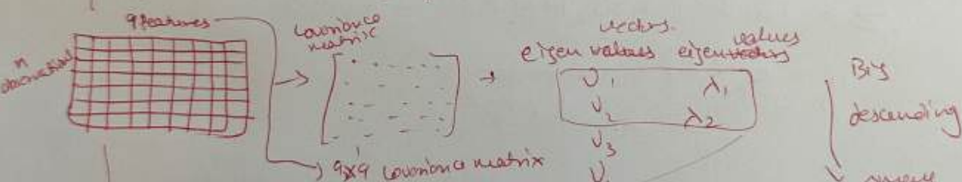
eigen vector = $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$

$\lambda =$ 11 1

Every covariance matrix is symmetric so the eigen vectors will be orthogonal. you can project your data among one of the two eigen vectors but which has more spread. → $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ the vector with larger eigen value will always have higher variance.

11 →

project data on $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

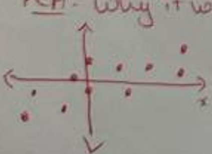


take two biggest to project onto.

finally

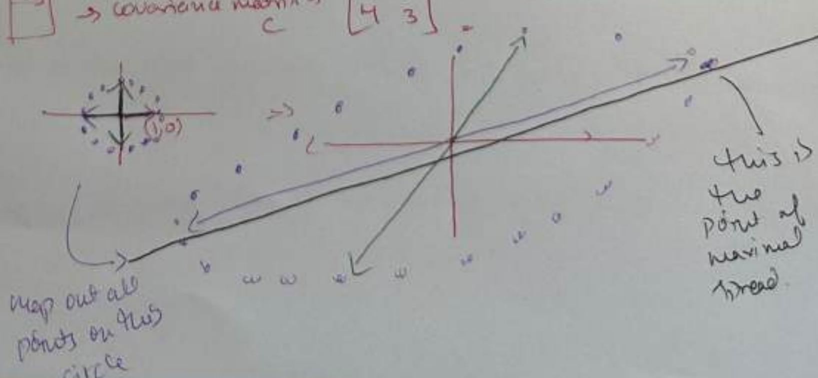
large table × $\begin{bmatrix} \frac{v_1}{\|v_1\|} & \frac{v_2}{\|v_2\|} \end{bmatrix}$ → [2xn] - dataset with 2 features.

PCA: Why it works?



$\boxed{x\ y}$ → covariance matrix C → $\begin{bmatrix} 9 & 4 \\ 4 & 3 \end{bmatrix}$

let's visualise



So a circle of radius one is transformed into an elipse.

heres where eigen values & vectors come in. there are 2 eigen vectors for covariance matrix.

$$\lambda = 11 \quad \lambda = 1$$

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$ → this makes eigen basis.
& from the perspective of this eigenbasis the transformation C only stretches the plane

The greatest spread if you map out is the same as the highest eigen value eigen vector.

This is because when transformation occurs one eigenvector gets ~~stretched~~ by 11, the other by 1 & any other point by a value b/w 1 & 11.

PCA Mathematical Formula

say you have "n" observations of 5 variables (x_1, x_2, x_3, x_4, x_5)

Goal: Reduce to 2 variables.

① Creat Matrix

$$\begin{array}{c c} \text{n observations} & \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{15} \\ x_{21} & x_{22} & \cdots & x_{25} \\ \vdots & & & \\ x_{n1} & x_{n2} & \cdots & x_{n5} \end{bmatrix} \\ & \text{5 variables} \end{array}$$

② Center the data.

$$x - \mu = \begin{bmatrix} x_{11}-\mu_1 & x_{12}-\mu_2 & \cdots & x_{15}-\mu_5 \\ x_{21}-\mu_1 & x_{22}-\mu_2 & \cdots & x_{25}-\mu_5 \\ \vdots & \vdots & & \vdots \\ x_{n1}-\mu_1 & x_{n2}-\mu_2 & \cdots & x_{n5}-\mu_5 \end{bmatrix}$$

③ calculate Covariance Matrix.

$$C = \frac{1}{n-1}(X-\mu)^T(X-\mu) = \begin{bmatrix} Var(X_1) & Cov(X_1,X_2) & Cov(X_1,X_3) & \cdots \\ Cov(X_1,X_2) & Var(X_2) & & \\ Cov(X_1,X_3) & Cov & Var(X_3) & \\ Cov(X_1,X_4) & & & Var(X_4) \\ Cov(X_1,X_5) & & & & Var(X_5) \end{bmatrix}$$

look at slides

Goal reduce to 2 variables.

4) Calculate Eigenvalues & Eigenvectors.

big

$\lambda_1 v_1$

$\lambda_2 v_2$

$\lambda_3 v_3$

$\lambda_4 v_4$

$\lambda_5 v_5$

small

$$V = \left[\frac{v_1}{\|v_1\|_2} \; \frac{v_2}{\|v_2\|}\right], \quad X_{PCA} = (X-\mu)V$$

Rank of Matrix

measures how much info the matrix or it's corresponding linear equations have

~~For~~ one application is image compression.
For eg → you have an image of high quality means more pixels which uses more memory.
you could have the same image of slightly lower quality & save a lot of space.

Systems of Information

System 1	System 2	System 3
dog black	dog black	dog
cat orange	dog black	dog
L 2 sentences & carries 2 pieces of information	2 sentences but only 1 piece of info.	2 sentences 0 info.
2 info	1 info	0 info
2 rank	1 rank	0 rank

So the amount of info a system carries is called the rank of a system

Systems of equations.

1	2	3
$a+b=0$	$a+b=0$	$0a+0b=0$
$a+2b=0$	$2a+2b=0$	$0a+0b=0$
2 equations. each equation has new information.	2 equations but second equation same as 1st. so only 1 piece of info	2 equations → 0 information rank = 0

So

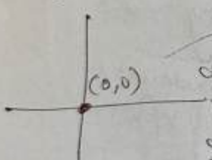
$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ 2 $\qquad \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ 1 $\qquad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ 0

Special relation b/w rank of matrix & it's solution space : ~~Special~~

Solution space -> The set of solutions to the system of equations when the constants are 0.

So for

$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$

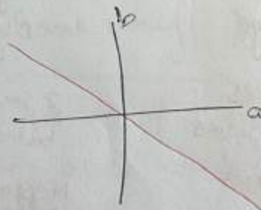


this is the solution

the solution is one point so the dimension of a point = 0

so rank = 2

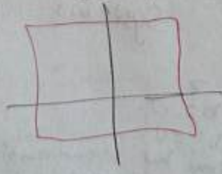
$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$



Solution is a line the dimension of a line is 1

so rank = 1

$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$



Solution is a plane so dimension = 2

rank = 0

(2) Rank = (No. of rows in matrix) – (Dimension of solution space)

If Rank = No. of rows → Non-Singular

System 1   2   3   4

$a+b+c=0$

$a +$

For a matrix you find rank using REF

$$\begin{bmatrix} 5 & 1 \\ 4 & -3 \end{bmatrix}$$

The idea is to get rid of 4 in bottom left & 1st divide each row by their first non-zero coefficient

$$\begin{bmatrix} 1 & 0.2 \\ 1 & -0.75 \end{bmatrix}$$

→ Now subtract 1st row from bottom to get 0 on bottom left row.

$$\Rightarrow \begin{bmatrix} 1 & 0.2 \\ 0 & -0.95 \end{bmatrix}$$

Now again step 1

divide 2nd row by the left most non-zero coefficient
so divide by −0.95

This gives → $\begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix}$ → REF

REF for Singular

$\begin{bmatrix} 5 & 1 \\ 10 & 2 \end{bmatrix}$ divide by left most

⇓

$\begin{bmatrix} 1 & 0.2 \\ 1 & 0.2 \end{bmatrix}$ take bottom row subtract top row

⇓

$\begin{bmatrix} 1 & 0.2 \\ 0 & 0 \end{bmatrix}$ Now divide bottom row by leftmost non-zero coefficient

Not possible.

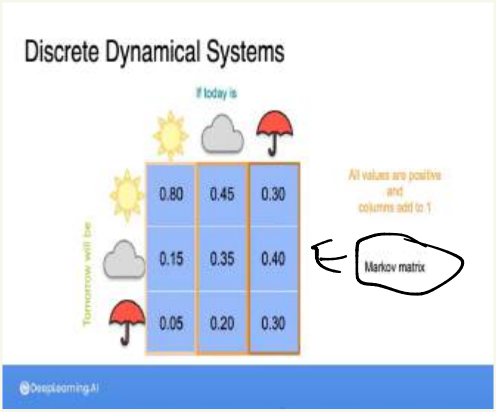
Let this be REF

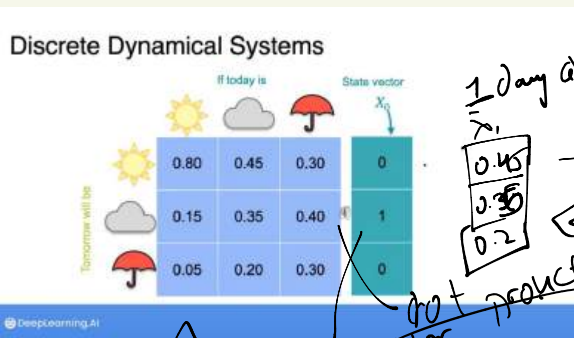
(3)

$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ → ~~this~~ this is the REF

Compare 3 REF's

MATRIX		REF	
$\begin{bmatrix} 5 & 1 \\ -4 & -3 \end{bmatrix}$	→	$\begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix}$	— ~~2 1's~~ there's 2 1's so rank=2
$\begin{bmatrix} 5 & 1 \\ 10 & 2 \end{bmatrix}$	→	$\begin{bmatrix} 1 & 0.2 \\ 0 & 0 \end{bmatrix}$	→ 1 one for diagnol so rank=1
$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	→	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	→ 0 one for diagnol so rank=0





1 day ahead

x_1

$$\begin{bmatrix} 0.45 \\ 0.35 \\ 0.2 \end{bmatrix}$$

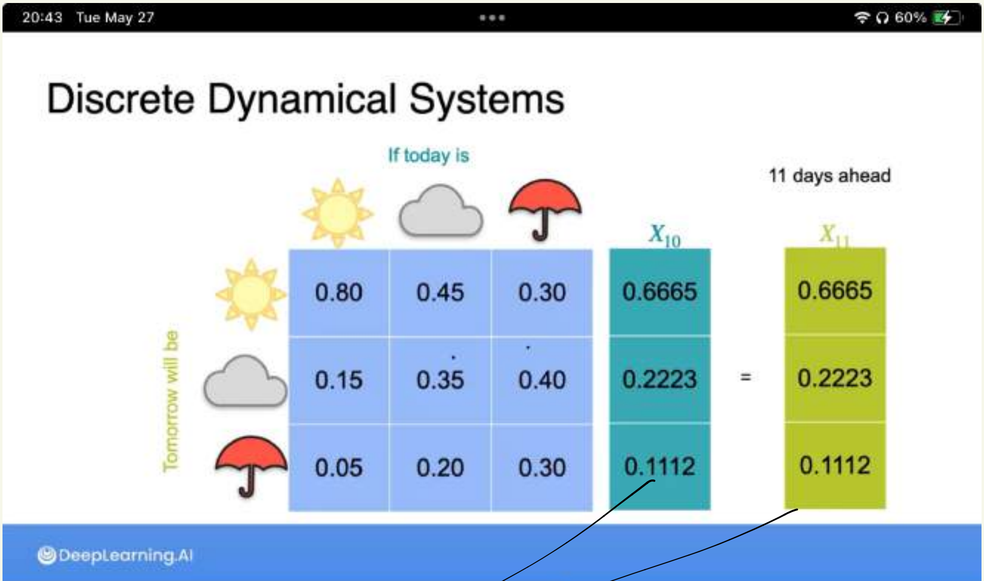
dot proudct

state vector that it is a cloudy day

for day 1 this dot product with the

Transformation matrix

So when you do this you will notice that when you start finding probability 10-11 days ahead the changes will be minute.



The difference is miniscule. This means this is an eigenvector with a eigen value of 1. It cannot be changed no matter how many transformations happen