

You are so INCREDIBLY AWESOME

Written → ANAND GOEL (Mathematics)
Reference

5-IN-1-NOTEBOOK

SmileyWorld

HOW TO READ MATRIX

↳ Confusion

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \Rightarrow \begin{cases} (1) \text{ Coefficient of } x_1 \\ (1) \text{ Coefficient of } x_2 \end{cases}$$

{ Then why coefficient of $x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ forms Column space? }

⟹ first column $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, second column $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

↳ Rows tell you equations

↳ Columns tell you directions.

Think this way

↳ $x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

gives,

$(x_1 + x_2) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ } Both Lines Same

Rank = 1





Remember how matrix is written:

$$Ax = b$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} t \\ 2t \end{bmatrix}$$, gives this

This actually gives column space,

$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ⟵ COLUMN-SPACE

1-D ⇒ $x \mapsto ax$

If $a \neq 0$

- image = whole line
- kernel = only zero

If $a = 0$

- image = zero only
- kernel = whole line.

2-D	3-D	4-D .....
Matrix can map plane to point, line, plane	Matrix can map space to point, line, plane, full 3-D	Sam logic

(26) B - Basis

- Set of vectors that span the space
- Linear Independent

It is the smallest complete set of directions needed to describe a space.

Eg. in 2-D

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \in R^2$$

- these can build every vector in plane
- neither is redundant.

⇒ How to find dependent / independent?

vector
- Is one vector just a scaled copy of another?
- Can one vector be written using the others?
- If I remove one vector, do we loose something.





Use in ML (Why Basis)?

⇒ If one feature is basically already contained in the others, the model has trouble deciding how to assign weights cleanly.

⇒ That's why linear independance matter so much.

If we have three vectors

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

Span, Column Space has all of three but $\begin{bmatrix}1\\1\end{bmatrix}$ is redundant. That's why Basis is important. It gives minium number (good calculation) & least value

Basis - How Connected to Rank

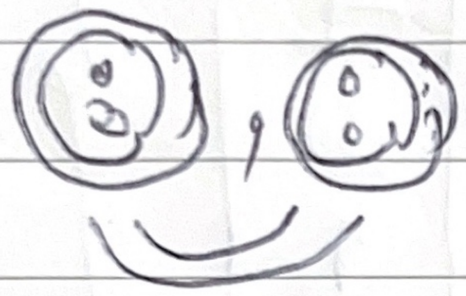
⇒ Rank is the number, Basis is actual set

Basis is for space → Not for just one equation.

Questions what is basis for R^3?

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

Rank => 3

This is simple 

But, Find basis for the set of vectors satisfying
$x+y+z=0$

This is not all of R^3, its subset of R^3
Actually it is space R^2, which is plane on R^3 space, so two directions.

$$\left\{\begin{bmatrix}x\\y\\z\end{bmatrix} : x+y+z=0\right\}$$

$$\Rightarrow x = -y-z$$



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

So basis is $\Rightarrow$

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Dimension - 2
But Rank = 1

$\Rightarrow$ See how 2 (1's) and 1 (0) make 2 directions in 3 value matrix.

Why Rank $\Rightarrow$ 1

Rank is for output, image

$\rightarrow$ Why???

$x + y + z = 0$, it true

⇒ It equation of [redacted] plane, 2-D

⇒ But output is 1-D, point on line.

RanK is found on output (1-D)

Here,

Sno	Object	For $A = [1\ 1\ 1]$	Lives what?
①	A	1×3 matrix	matrix space $R^{1\times3}$
②	input x	$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$	R^3
③	output Ax	$x+y+z$	R^1
④	image	R^1	output side
⑤	column space	R^1	output side
⑥	kernel	$x+y+z=0$, plane	input side
⑦	RanK	1	number, not space





Confusing Symbol

① $x \rightarrow$ plain x

② $\hat{x} \rightarrow$ x-hat $\Rightarrow$ estimate / projection

③ $x^T \rightarrow$ x-transpose $\Rightarrow$ row version of vector

④ $x^T y \rightarrow$ dot product / inner product

⑤ $\dot{x} \rightarrow$ first derivative wrt time

⑥ $\ddot{x} \rightarrow$ second derivative wrt time

⑦ $x' \rightarrow$ first derivative wrt (other than time mostly)

⑧ $x'' \rightarrow$ second derivative wrt (other than time mostly)

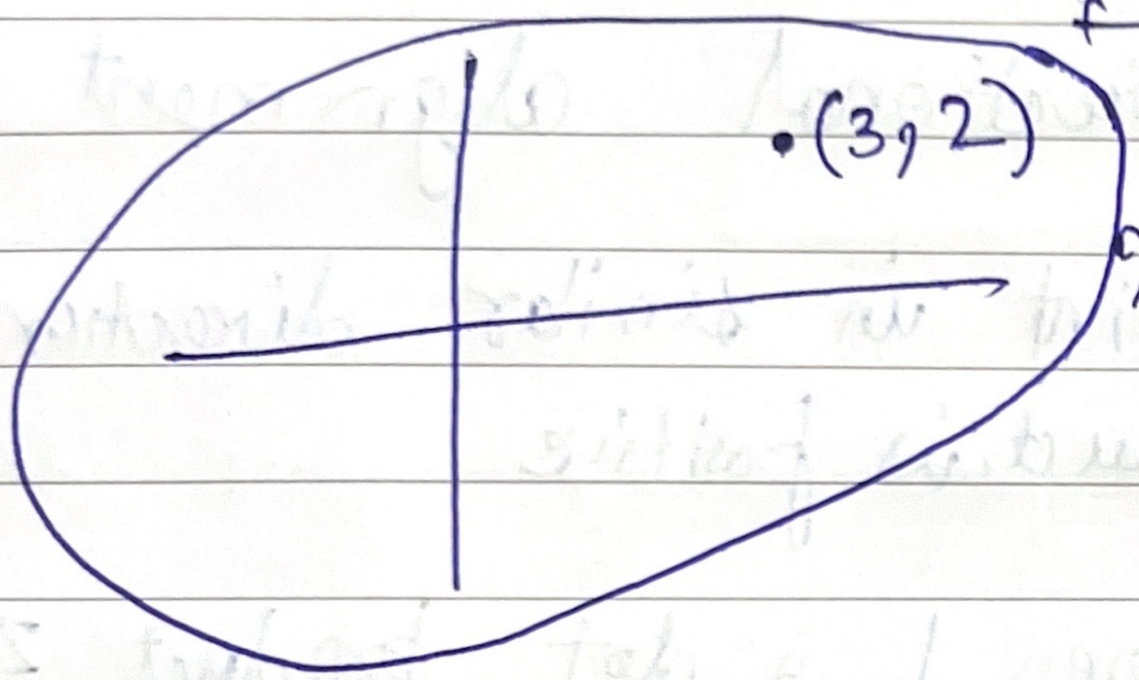
Till now we saw $Ax = b$, exact targetting.

APPROXIMATION :- If we cannot hit the target exactly, what is closet thing we can hit

(21) $U^{\perp}$ — Orthogonal Complement of U

All directions that are perfectly perpendicular to every direction of U.

→ Original vector = explainable part + leftover part



If $U = \text{span}\left(\begin{bmatrix}1\\0\end{bmatrix}\right)$

⇓

This is x-axis

⇓

Then what is $U^{\perp}$

Purpose of Orthogonal Complement

→ Find closent point from given point into space

Point (3,2) is closet to (3,0) as U is x-axis as we assumed. what is

leftover $= \begin{bmatrix}0\\2\end{bmatrix}$ because,

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





→ Remember dot product & why it has geometric meaning

$$a^T b = \|a\| \|b\| \cos\theta$$

where θ is angle between a & v

It measures directional alignment

- vectors point in similar direction dot product is positive.
- vector are ⊥, dot product zero
- → If they point opposite-ish, dot product is negative.

{{ Pythogoreans theoram is valid for all of them. }}

This means,

$$\cos\theta = \frac{a^T b}{\|a\| \|b\|}$$

If $\perp_g$ then $\cos 90 = 0$ so,

$$\frac{a^T b}{\|a\| \|b\|} = 0 \iff a \perp b \iff a^T b = 0$$

Now,

The column space of A is the set of all possible output of $(Ax = b)$

$$Col(A) = \{Ax : x \in R^n\}$$

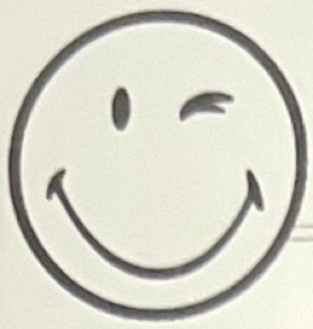
Since every Ax is a linear combination of the columns, column space is:

$\Rightarrow$ {all vectors you can build from the columns of A}

If target vector b is inside $Col(A)$, then there exists some x such that

$$Ax = b$$





If b is not inside col (A), then there exists some x such that

$$Ax = b$$

$\Rightarrow$ { If b is not inside column space, then exact solution is impossible.

BEST - We - CAN - DO

$\Rightarrow$ { Find the point in column space closest to b.

THIS IS PROJECTION

If V is a subspace and x is any vector, the projection of x ~~onto~~ onto V, usually written as $\hat{x}$ of means:

"the vector in V, closest to x"

$$x = \hat{x} + (x - \hat{x})$$

$\hat{x}$ = explained / captured / representable part

$(x - \hat{x})$ = leftover / residual / unexplained part

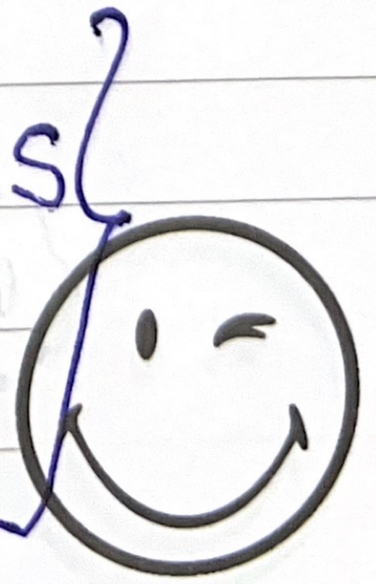
FORMULAE

$$\hat{x} = \arg\min_{y \in S} \|x - y\|$$

This is very dense line. It contains many ideas packed together

Read above formulae as ⇒

"$\hat{x}$ is the point, y inside the set S that makes the distance from x to y as small as possible"





Here,

$$\min_{y \in S} \|x - y\|$$

means,

among all points y in S, take smallest distance.

But,

$$\operatorname{argmin}_{y \in S} \|x - y\|$$

means, the input y where the minimum happens.

So,

min, gives the smallest distance

argmin, gives the point achieving that smallest distance.

Example, Suppose set S is the x-axis

$x = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, then closest point on the x-axis is $\hat{x} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$

Distance from point to x-axis

$$\left\| \begin{bmatrix} 3 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\| = 2$$

So,

$$\min_{y \in S} \|x - y\| = 2$$

This is number.

$$\arg\min \|x - y\| = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

This is point/vector. Thats difference.





Calculus / Differenciation will be covered later but whats important to understand and solve projection questions / see navies.

$$f(c) = c^2$$

If c changes, $f(c)$ also changes

$$\left\{ f'(c) \text{ or } \frac{d}{dc} f(c) \right\} \text{ Differenciation}$$

So $f'(c) = 0$ (minimum point)

THIS HELPS FINDING PROJECTIONS

POWER RULE OF DIFFERENTIATION

$$\frac{d}{dc}(c^2) = 2c$$

why power of c comes down?

Let c increase h which is very little

$$(c+h)^2 = c^2 + 2ch + h^2$$

So the change is:

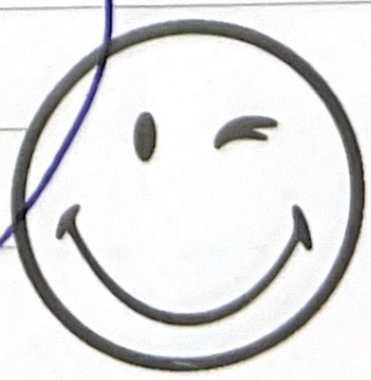
$$(c+h)^2 - c^2 = 2ch + h^2$$

Divide by h:

$$\frac{(c+h)^2 - c^2}{h} = 2c + h$$

as h change is small $h = dc$ & $h \to 0$

so $$\frac{d(c^2)}{dc} = 2c$$ $\left(\text{as } h \to 0 \text{ & } h = dc\right)$





Question :- $x = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ & a line $= c \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

we want closest point on that line to x

Line $\Rightarrow y = cu = \begin{bmatrix} c \\ c \end{bmatrix}$

Now, $\|x-y\|$

Since $y = \begin{bmatrix} c \\ c \end{bmatrix}$, we get $(x-y) = \begin{bmatrix} 4 \\ 2 \end{bmatrix} - \begin{bmatrix} c \\ c \end{bmatrix}$

$$\begin{bmatrix} 4-c \\ 2-c \end{bmatrix} = (x-y)$$

$$\|x-y\|^2 = (4-c)^2 + (2-c)^2$$

$$f(c) = \|x-y\|^2 = 20 - 12c + 2c^2$$

$$f(c) = 20 - 12c + 2c^2$$

$$\frac{d f(c)}{dc} = 4c - 12$$

WHY - DIFFERENTIATION

Now for minimum point,

$$\frac{d\,(f(c))}{dc} = 0$$

LHS | RHS

$$4c - 12 = 0$$

$$c = 3$$

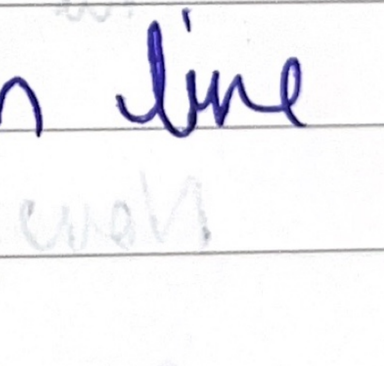
, At this point we get minimum

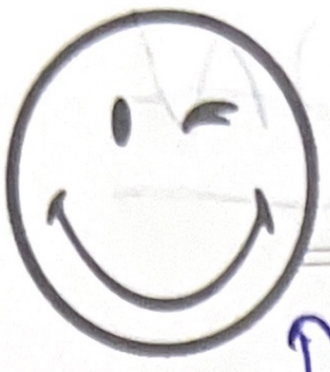
$$\hat{x} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

For point $(4, 2)$, nearest point on line

$y = x$ is $(3, 3)$

CHECK - THIS - OUT





Derivation of the line projection formula

Let

- x = vector to project
- $u \neq 0$ = direction vector of line

Every point on the line is: $y = cu$

$$f(c) = \|x - cu\|^2 = (x - cu)^T (x - cu)$$

$$= x^T x - c x^T u - c u^T x + c^2 u^T u$$

So, $x^T u = u^T x$

this becomes, $f(c) = x^T x - 2 c x^T u + c^2 u^T u$

Now this is quadratic in c.

$$\left\{ f'(c) \Rightarrow -2 x^T u + 2 c u^T u = 0 \right.$$

$$\Rightarrow 2 c u^T u = 2 x^T u$$

$$\left. \Rightarrow c = \frac{x^T u}{u^T u} \right\}$$

So the projection is,

$$\hat{x} = cu = \frac{x^T u}{u^T u}\, u$$

Theory means :-

→ $x^T u$ → how much x aligns with u (dot product)

→ then divide by the size of u, that is $u^T u$ (NORM)

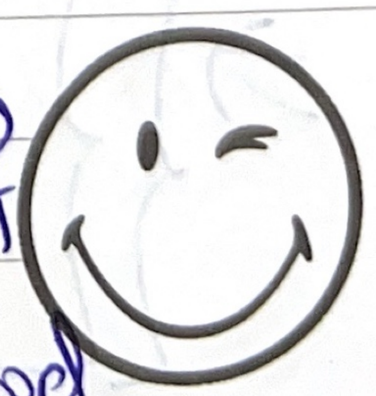
→ then multiply by u to get vector on the line

(22) $proj_w(x)$

└ Projection of x onto direction w

⇒ $v^{\perp}$ is what can be on any space / part of space

⇒ Because it is projection in direction of direction is always line hence its good.





$\Rightarrow$ { do,

$$\boxed{proj_w(x) = \frac{x^T w}{w^T w} w}$$

$\Rightarrow$ { also if $x, w \in R^D$, then

$$\boxed{proj_w(x) \in R^D}$$

Orthogonal Component of $U \Rightarrow U^{\perp}$

- If U is line, $U^{\perp}$ is a plane
- If U is plane, $U^{\perp}$ is a line

$\Rightarrow$ {{{ $x = u + v$ with $u \in U$ & $v \in U^{\perp}$
or or
$\hat{x}$ Residual }}}

$\Rightarrow$ {{{ $\hat{x}$ if on a line is nothing but
$proj_x(w)$ }}}

(23) $\pi(x)$ → Important

Here π is not 3.14, Something.

↳ It is projection onto a subspace U

This is what we are talking about.

This symbol

$$\pi(x) \neq proj_W(x)$$

where $W \Rightarrow$ direction = line

$$\pi(x) = proj_V(x)$$

where $V \Rightarrow$ Universe $\Rightarrow$ Can be line/plane/sphere etc.

