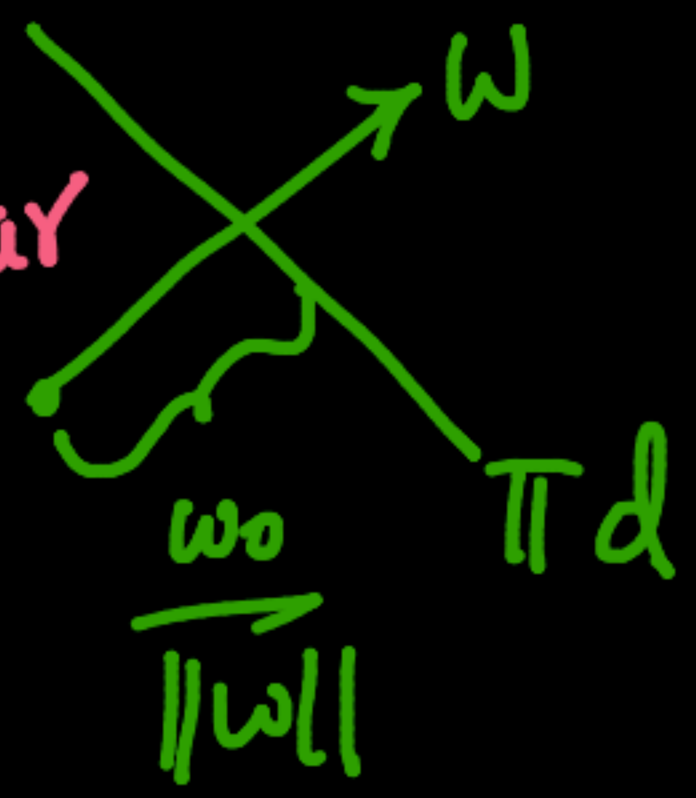
Recap:

Fish sorting problem / Binary Classification

1. d-dim Sep $\rightarrow \Pi_d$ $\quad w^T x + w_0 = 0$

   $w \rightarrow$ d-dim, $\quad w_0 \rightarrow$ scalar

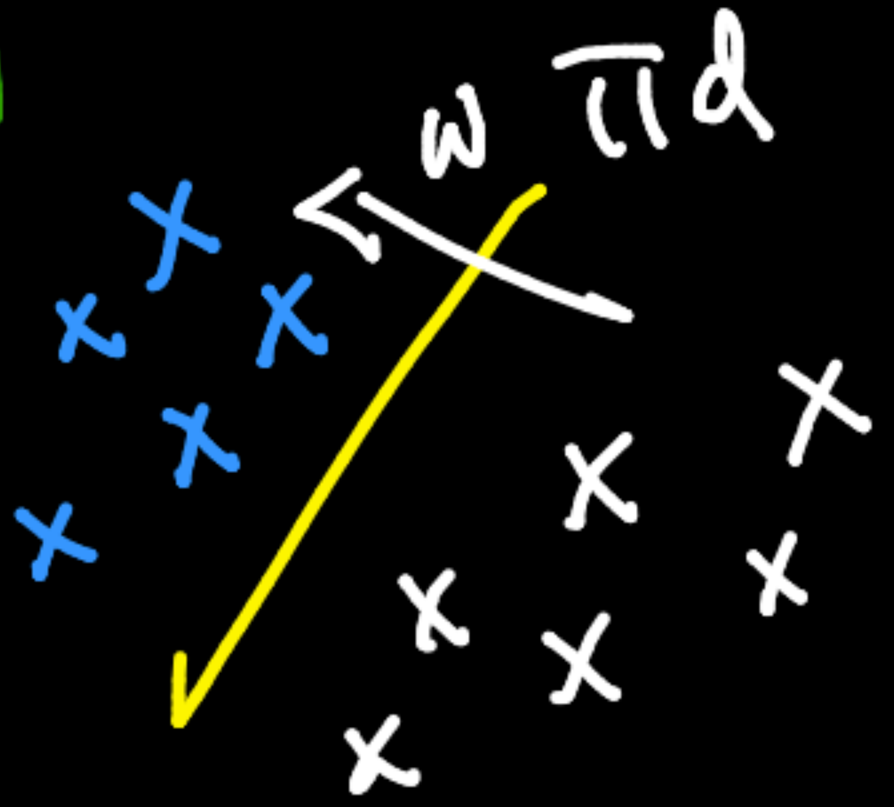


2. [find the sep. Π_d

   brute force

   calculus

3. Half-spaces

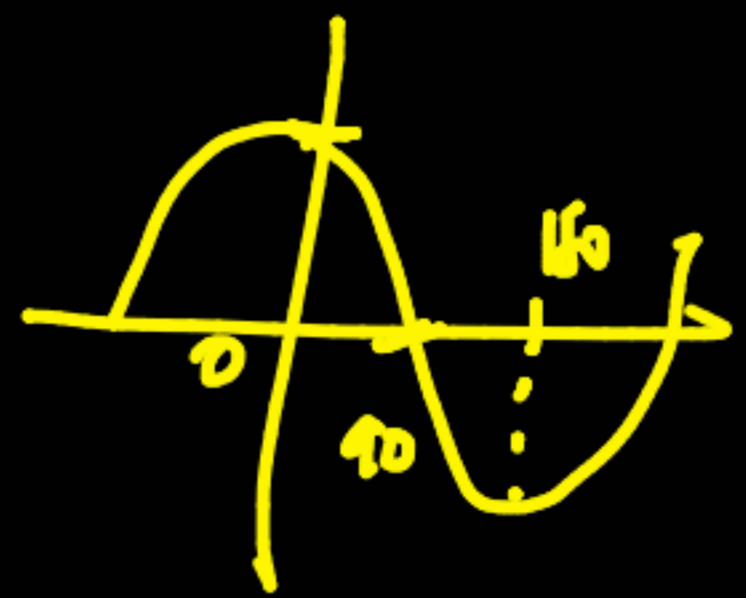


$w^T x + w_0 > 0$ $\quad$ $w^T x + w_0 < 0$

4 $\text{dist}(\underline{\underline{x_q}}, \pi_d) \rightarrow$ confidence

d-dim

$$\frac{w^T x_q + w_0}{||w||}$$



x_1 lies on π_d

$w^T x_1 + w_0 = 0$

$-\|w\|\|x_1\|\cos\theta + w_0 = 0$

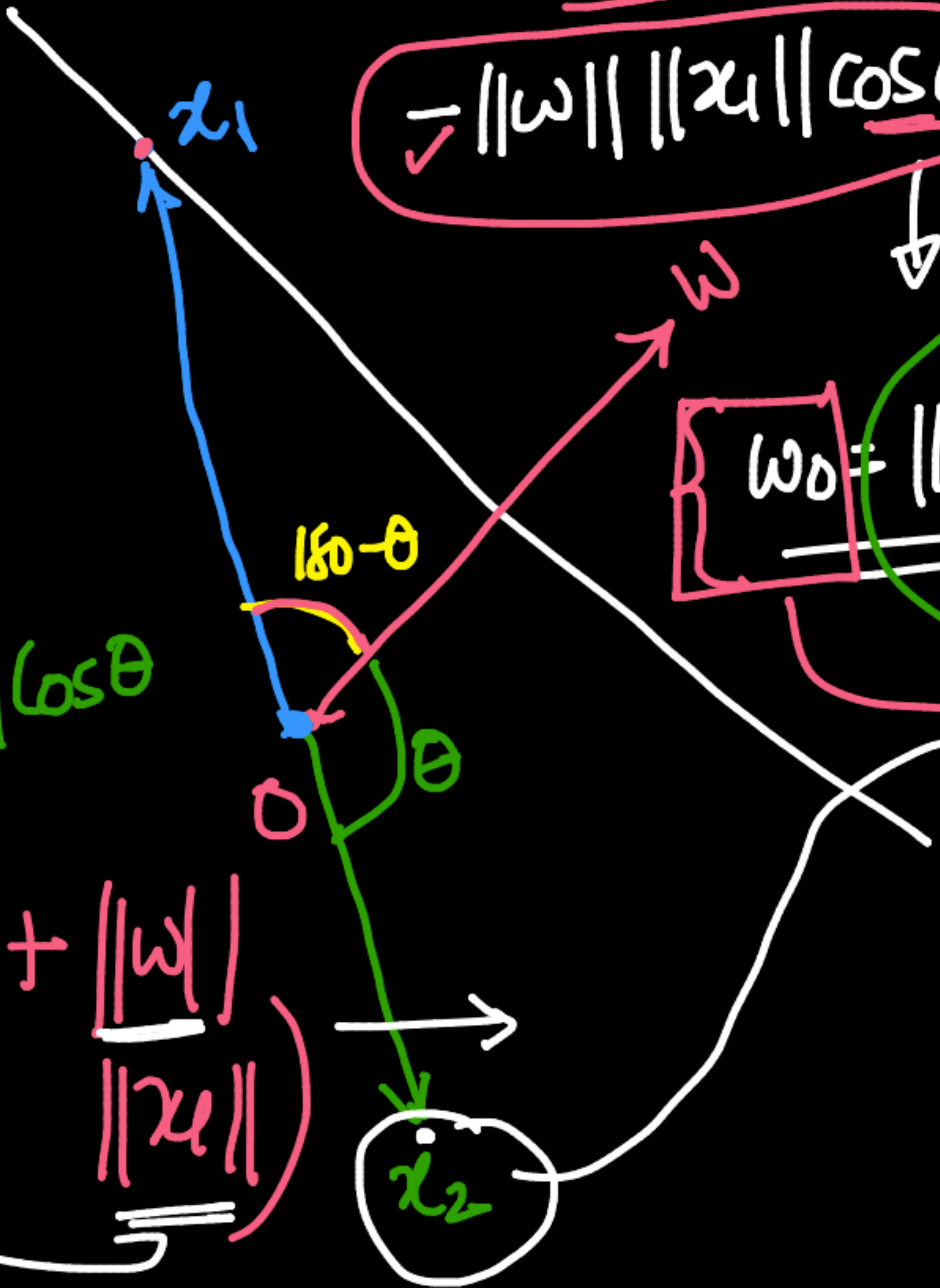
Anwesha

(geom + algebra)

$180^\circ > \theta > 90^\circ$

$\cos\theta < 0$

$\cos(180-\theta) = -\cos(\theta)$



$w_0 = \|w\|\|x_1\|\cos\theta$ (1)

$\pi_d : w^T x + w_0 = 0$

$\rightarrow w^T x_2 + w_0$ (2)

<0

$\rightarrow \|w\|\|x_2\|\cos\theta$

$+ \|w\|\|x_1\|\cos\theta$

$\cos\theta\left(\|w\|\|x_2\| + \frac{\|w\|}{\|x_1\|}\right)$

<0

70

(Q) * SVM

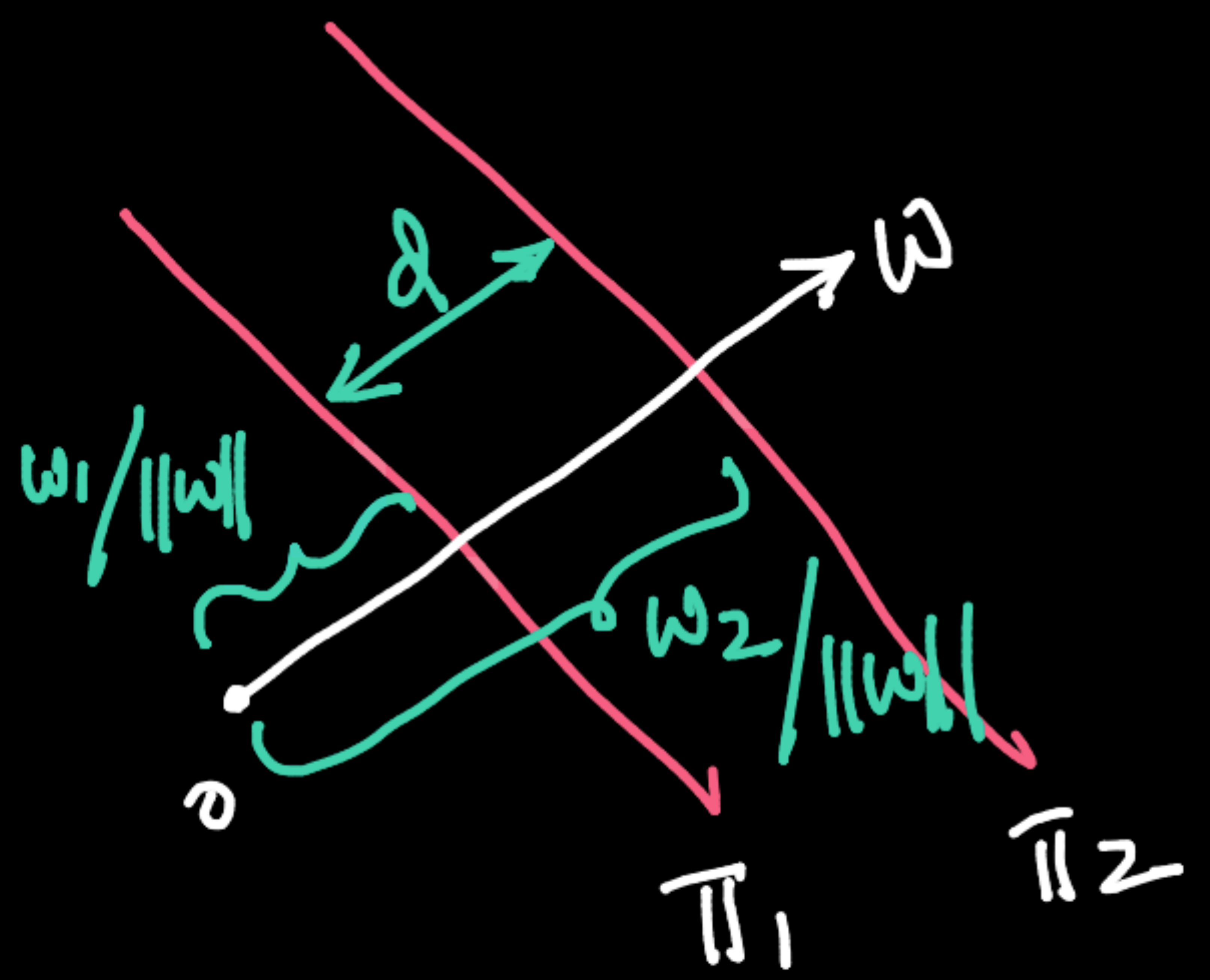
normal

$$\text{d-dim} \begin{cases} \pi_1 : \underline{w}^T x + w_1 = 0 \\ \pi_2 : \underline{w}^T x + w_2 = 0 \end{cases}$$

① $\pi_1 \parallel \pi_2$

② $\text{dist}(\pi_1, \pi_2)$

2-min

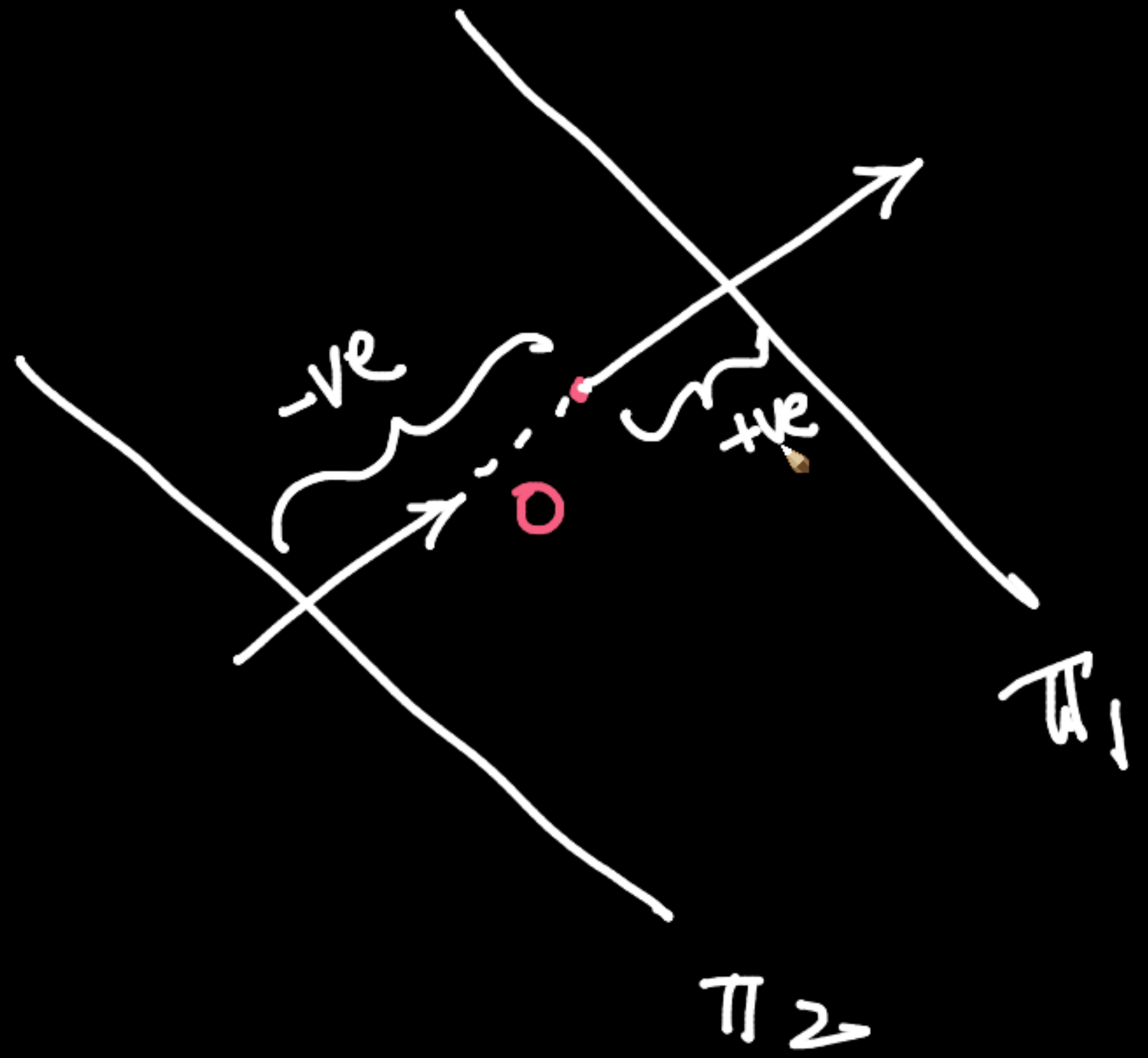


$$d = \frac{|w_2 - w_1|}{\|w\|}$$ ← abs

→ $w_2 - w_1$

→ $|w_2 - w_1|$

→ $\dfrac{|w_2 - w_1|}{\|w\|}$ ✓


-ve
+ve
O
π1
π2

3 Types $\longrightarrow$ T_1, T_2 & T_3

simplest way

π_1 $M_1 \rightarrow T_1$ or not

π_2 $M_2 \rightarrow T_2$ or not

π_3 $M_3 \rightarrow T_3$ or not

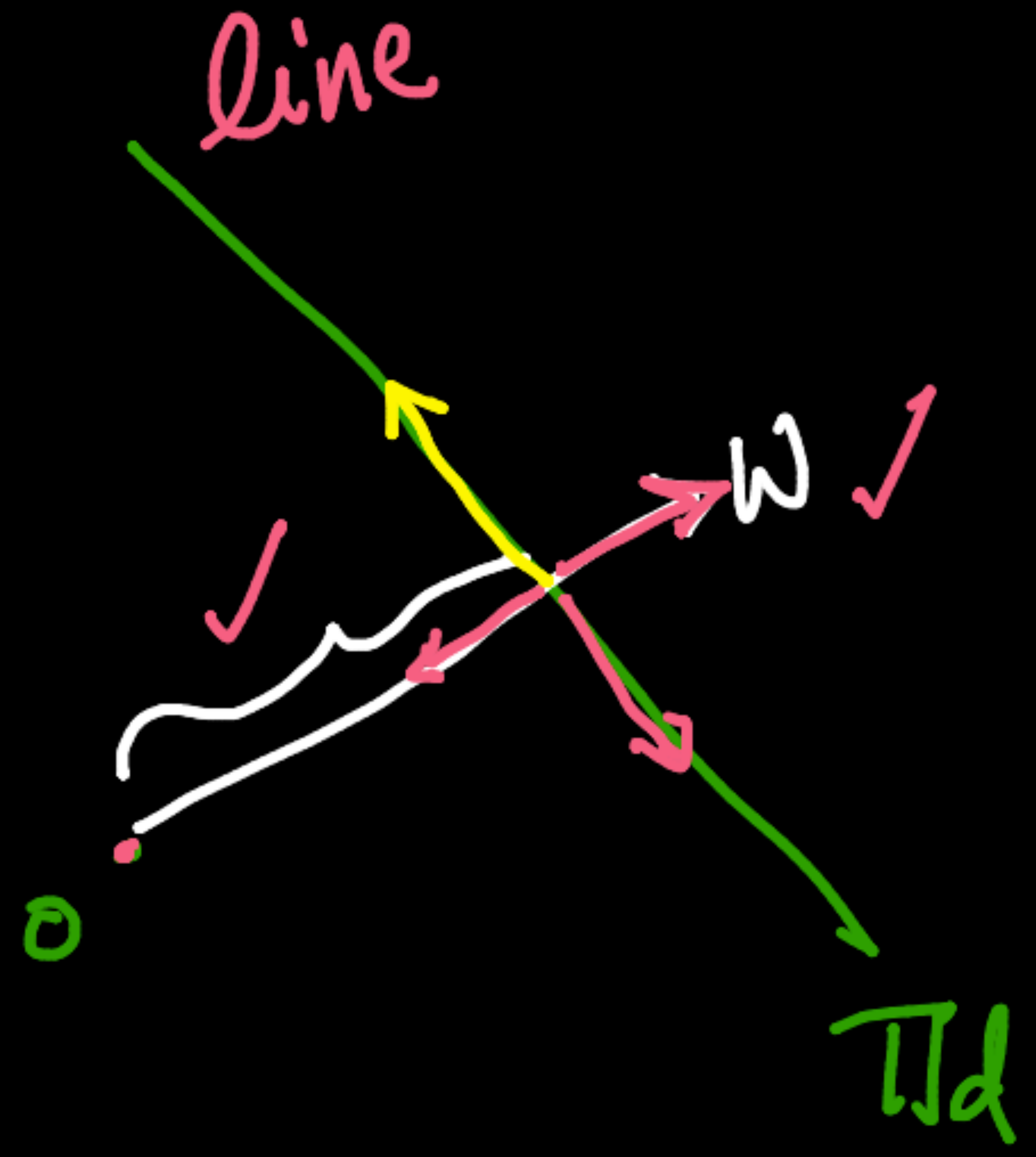
one vs rest

(Q)

Plane → ✓ normal or $\perp$ to πd (w)

→ $\text{dist}(0, \pi d) : \dfrac{w_0}{\|w\|}$

2D



$w \perp \pi d$

can we define a πd

→ $d(0, \pi d)$ &

→ one direction that is $\parallel$ to the πd

$$x_1 \cdot x_2 = \|x_1\| \ \|x_2\| \cos 0^\circ$$

$$= \|x_1\| \ \|x_2\|$$

Vectr: lines

$w, w_0 \rightarrow \pi_d$

unique $\perp$ directin

3D

d ✓

$w \perp w_1$

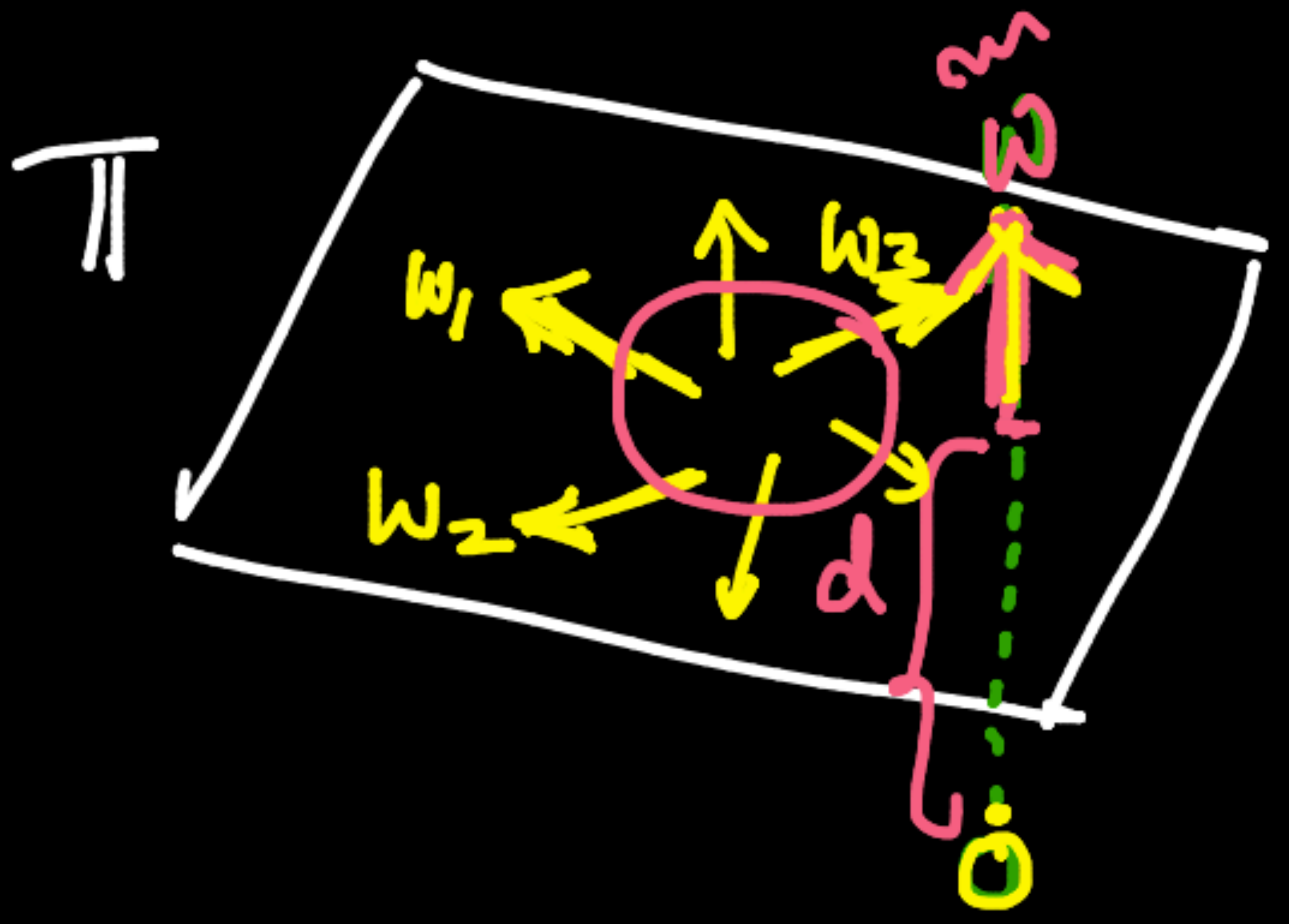
$w \perp w_2$

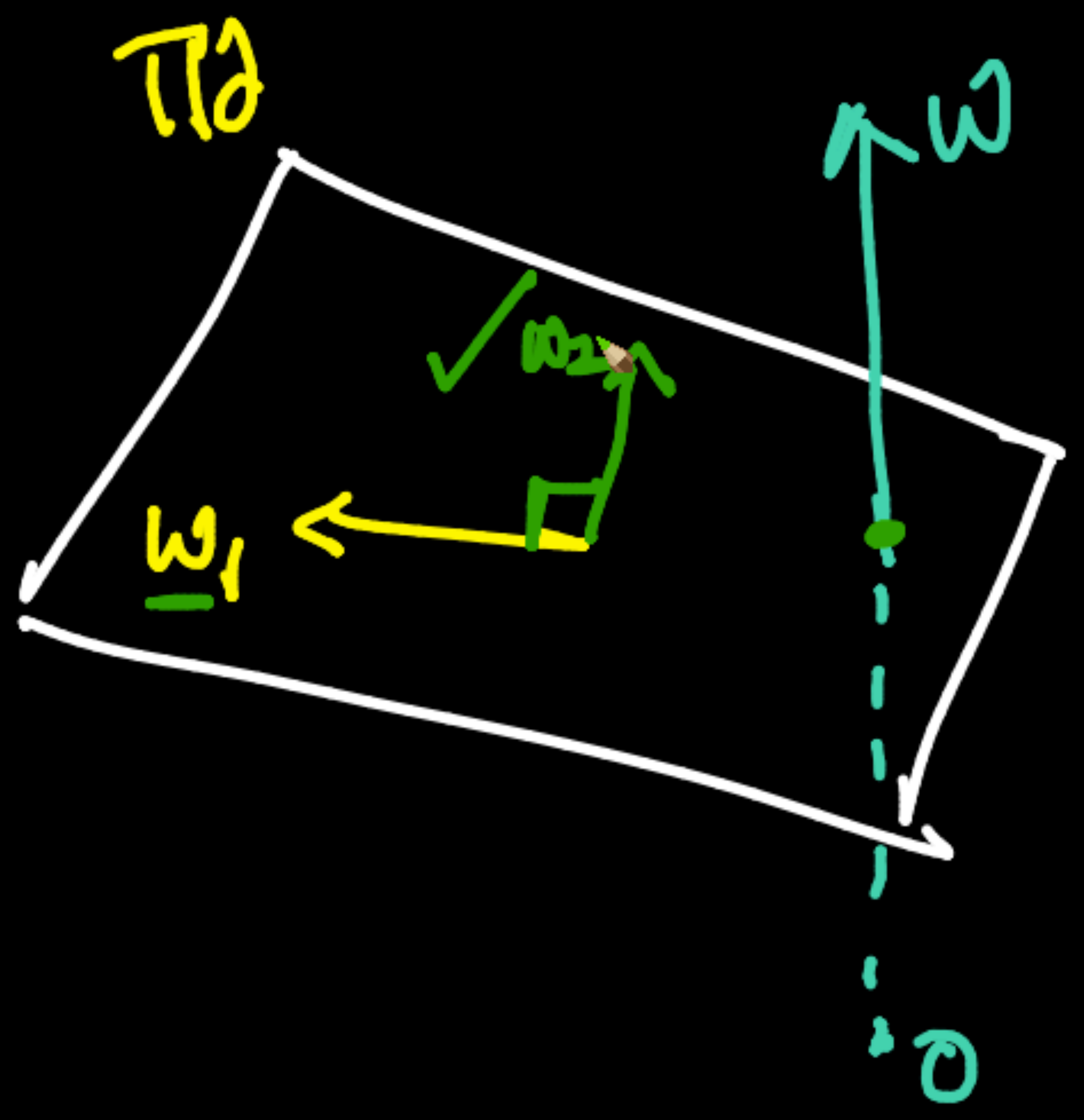
$\vdots$

$\rightarrow$ on the plane

direction

X parall to π_d ✓





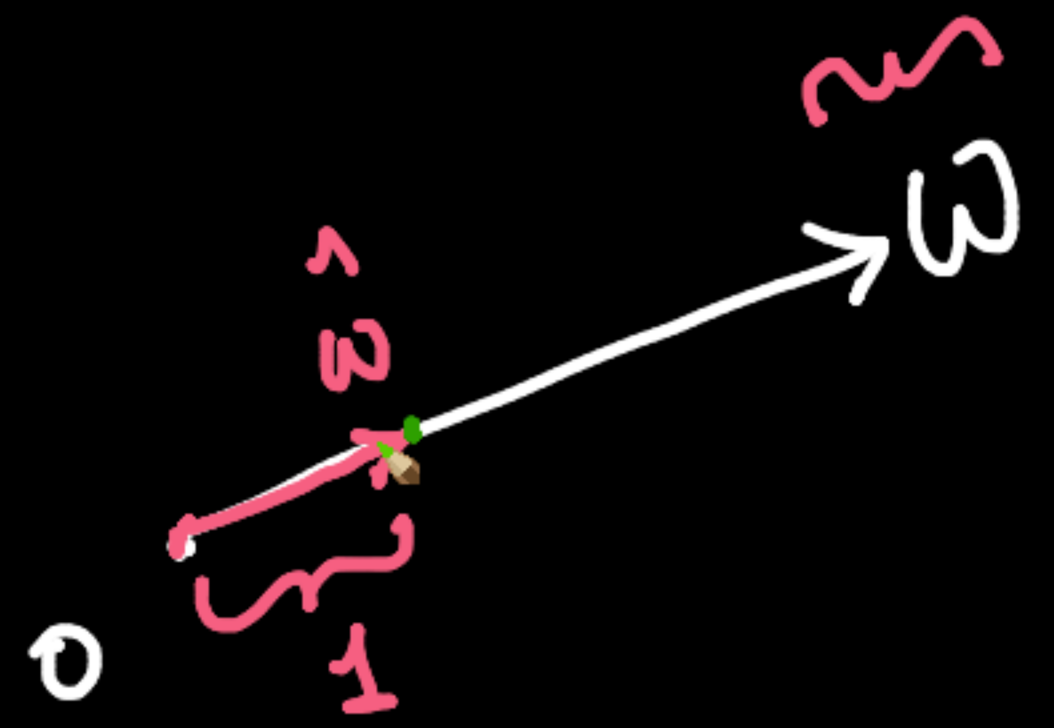
$\omega \perp \Pi\alpha$

$\omega_1 \parallel \Pi\alpha$

$\omega_2 \perp \omega_1$

$\omega_2 \not\perp \Pi\alpha$

Unit -vector



$\|w\| = 10$ (let)

$\hat{w}$ has same dir as w

$\|\hat{w}\| = 1$

$$w \rightarrow \underbrace{d\text{-dim}} = [w_1 \ w_2 \ \cdots \ w_d]^T$$

Vector $\rightarrow \hat{w}$

$$\hat{w} = \frac{w}{\|w\|} = \frac{[w_1 \ w_2 \ \cdots \ w_d]^T}{l}$$

$$= \left[\frac{w_1}{l}, \frac{w_2}{l}, \cdots \frac{w_d}{l}\right]^T$$

$$l = \sqrt{w_1^2 + w_2^2 + \cdots w_d^2}$$

$$\|\hat{w}\| = \sqrt{\left(\frac{w_1}{l}\right)^2 + \left(\frac{w_2}{l}\right)^2 + \cdots \left(\frac{w_d}{l}\right)^2}$$

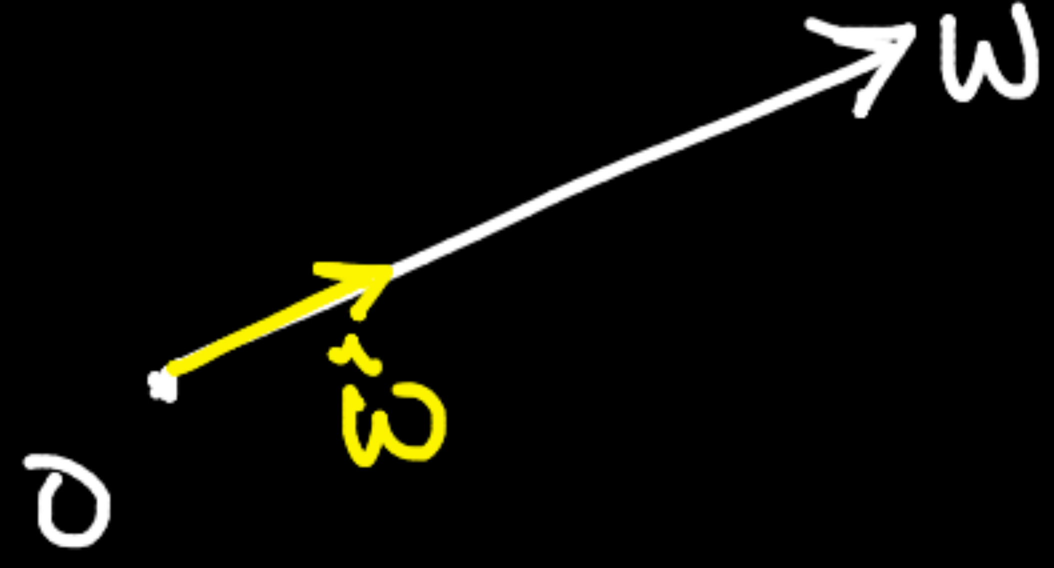
$$= \sqrt{\frac{w_1^2}{w_1^2+w_2^2+\cdots} + \frac{w_2^2}{w_1^2+w_2^2} + \cdots}$$

$$= \sqrt{1} = 1$$

Q

$$w = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\hat{w} = \begin{bmatrix} \frac{1}{\sqrt{14}} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$$



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

$$\|w\| = \sqrt{1+4+9} = \sqrt{14}$$

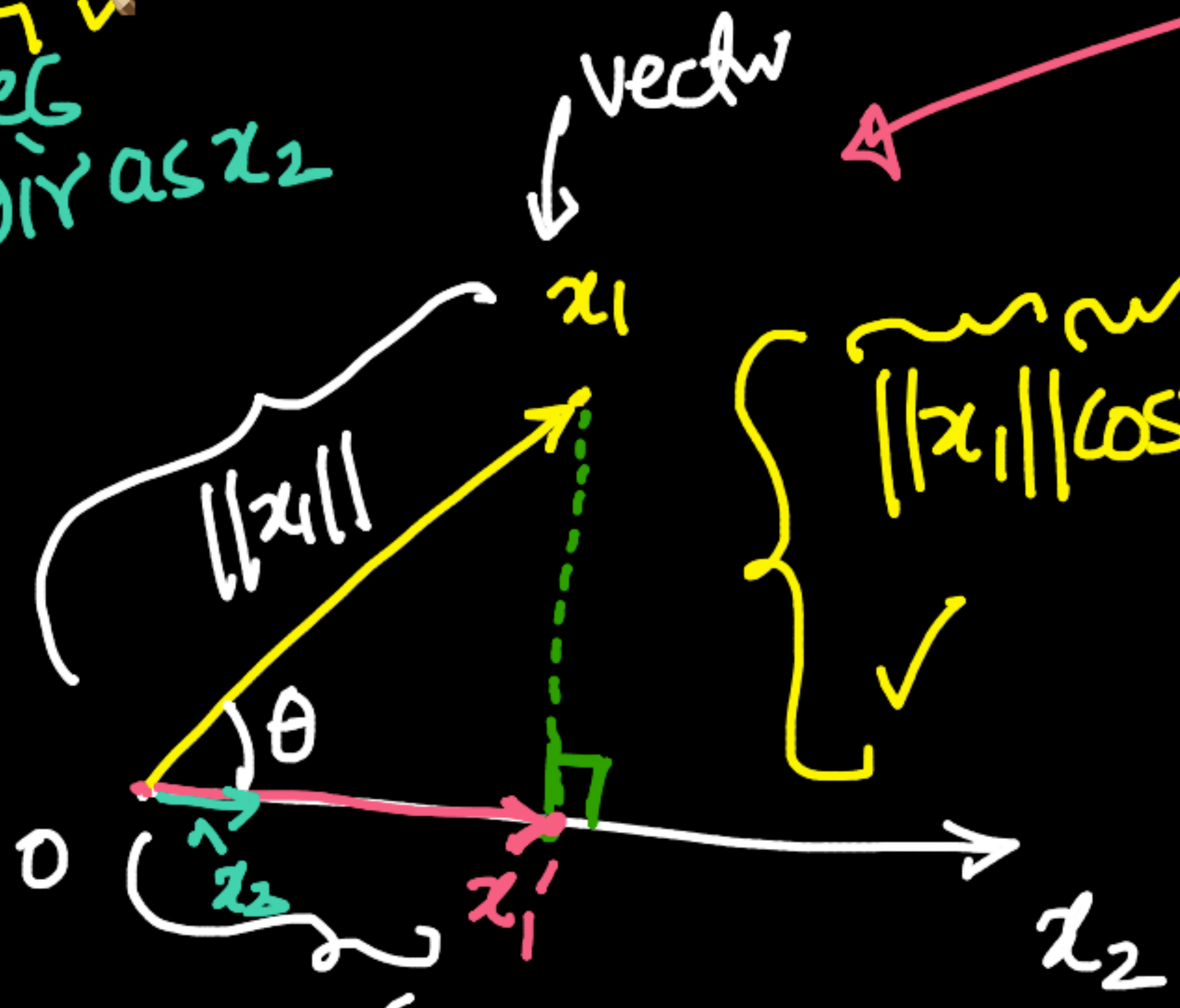
Projection

$\boxed{(PCA)}$

$\hat{x}_2 = $ unit-vec, same-dir as x_2

$\downarrow$

$\frac{x_2}{\|x_2\|}$



$\|x_1\|\cos\theta \cdot \frac{x_2}{\|x_2\|}$ (vector)

$\text{Proj}_{x_2}(x_1) = x_1'$

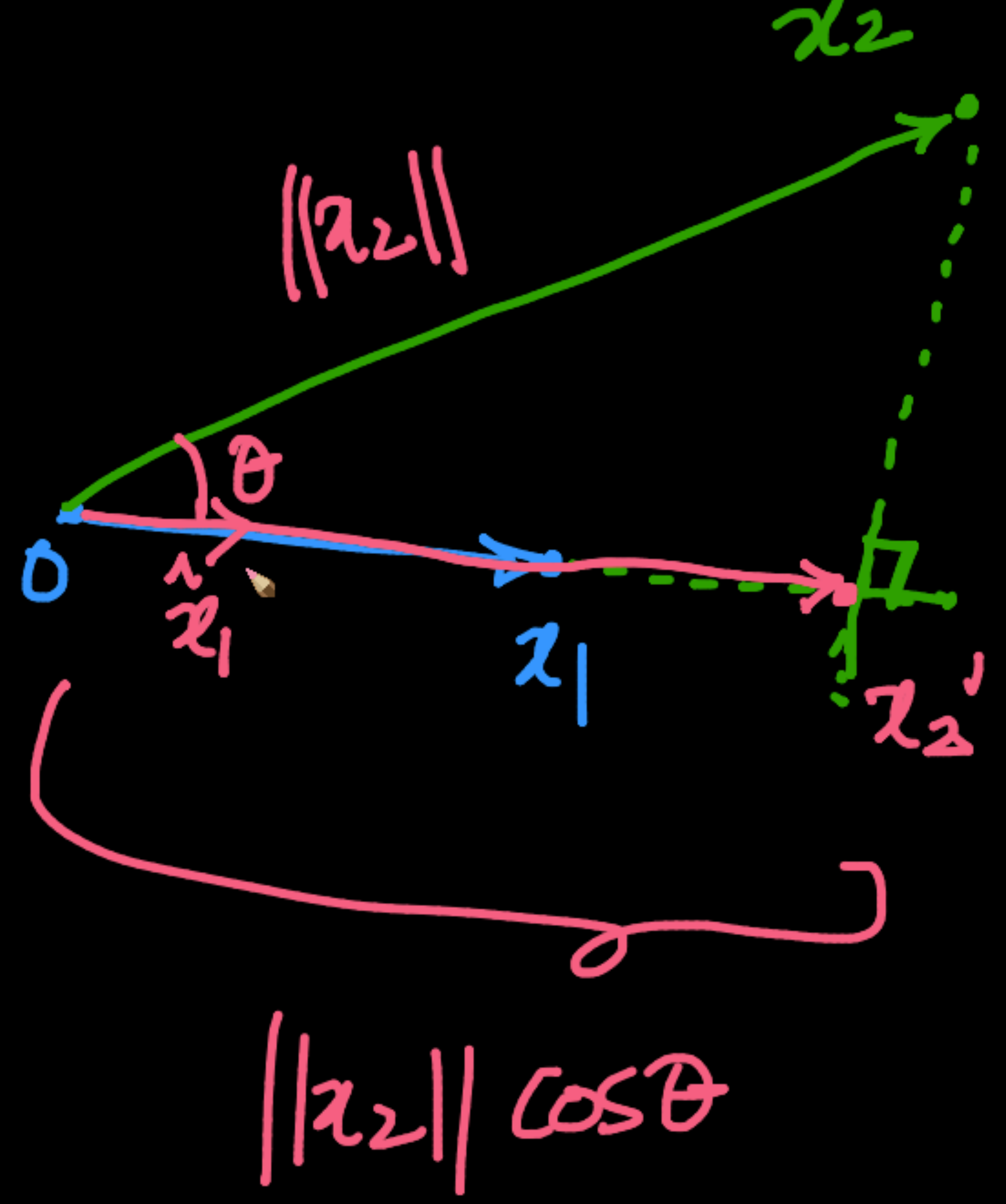
vector

1. Same direction as x_2
2. $\|x_1'\| = \|x_1\|\cos\theta$

$x_1' = \|x_1\|\cos\theta \cdot \hat{x}_2$

$\hat{x}_2 \rightarrow x_2 / \|x_2\|$

(Q)



$$x_2' = \text{Proj}_{x_1}(x_2)$$

$$= \text{len} \times \text{direction}$$

$$= \|x_2\| \cos\theta \cdot \frac{\vec{x_1}}{\|x_1\|}$$

$x_1 + x_2$

$$\begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1d} \end{bmatrix} + \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2d} \end{bmatrix} = \begin{bmatrix} x_{11} + x_{21} \\ x_{12} + x_{22} \\ \vdots \\ x_{1d} + x_{2d} \end{bmatrix}$$

addition/sub $\rightarrow$ component-wise

$x_1 \in \mathbb{R}^d$

$k \neq d$

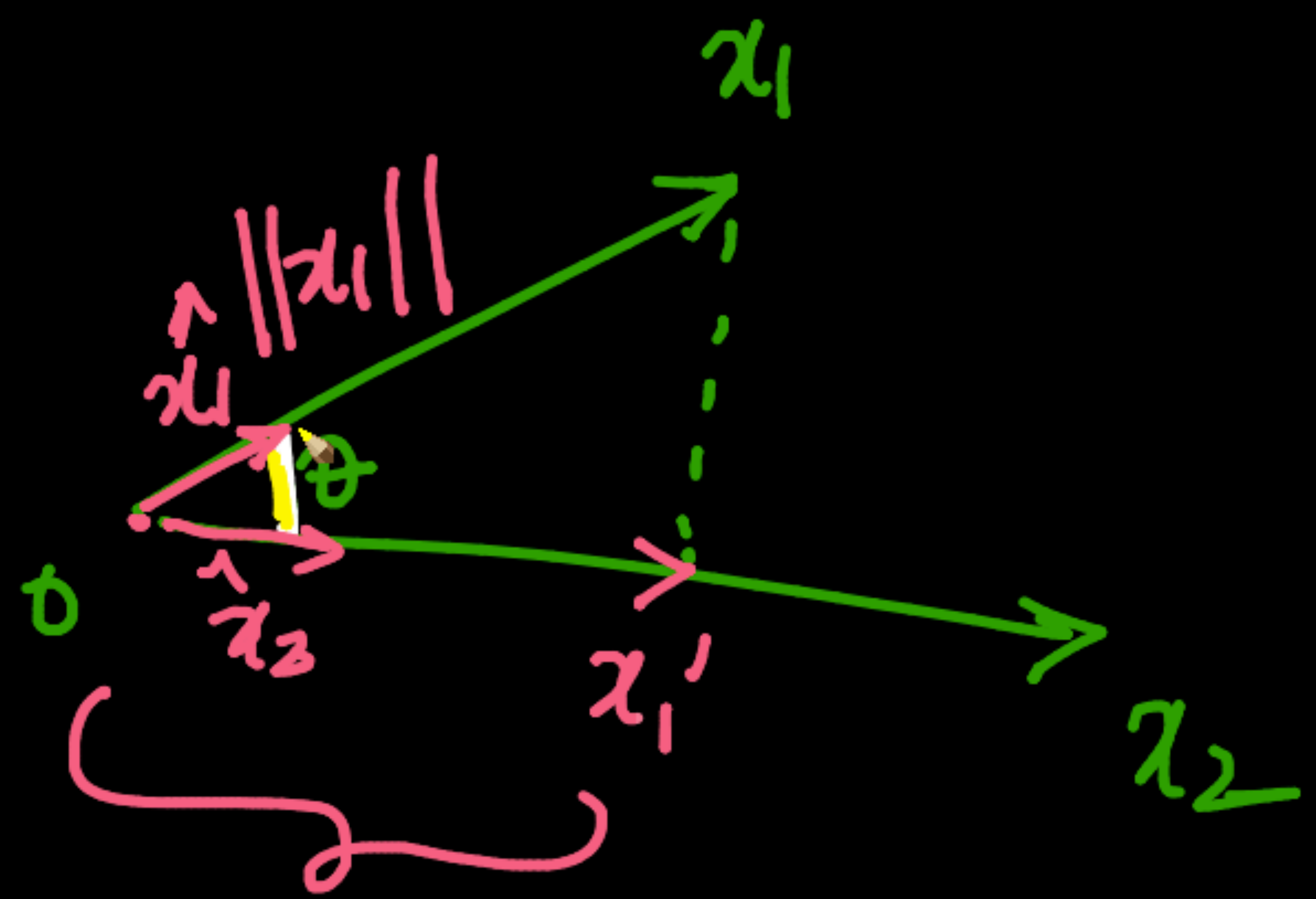
$x_2 \in \mathbb{R}^k$

$x_1 + x_2 = ?$

dot-product

len of proj of x_1 onto x_2

$$l = \|x_1\| \cos\theta = \frac{x_1 \cdot x_2}{\|x_2\|}$$



$$\cos\theta = \frac{x_1 \cdot x_2}{\|x_1\| \, \|x_2\|}$$

2 unit-vec

Geom: angle between 2-vec l

$$\hat{x}_1 \cdot \hat{x}_2 = \cos\theta$$

$$\theta = \cos^{-1}(\hat{x}_1 \cdot \hat{x}_2)$$ ✓

ML → planes; dir; dist ✓

planes (π_d) → simplest sep in d-dim space

Sep → circles/spheres/hyperspheres

$\boxed{S_d}$

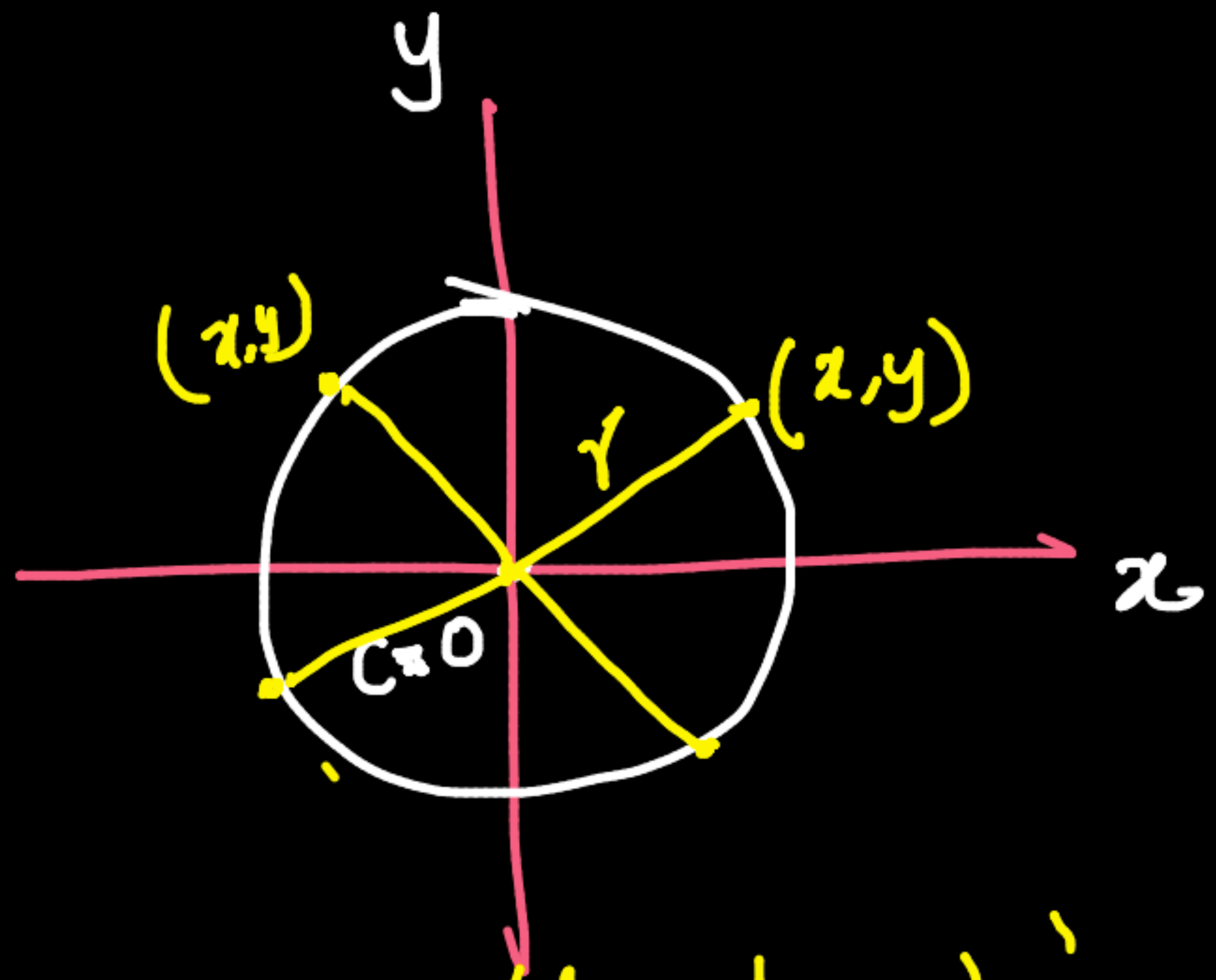
circle:

(2D)

not often used in ML

→ planes with transformations

(log - transform)



circle through origin

$$\sqrt{x^2+y^2} = r$$

$$x^2+y^2 = r^2$$

d-dim:

Sd:

Center $\in \mathbb{R}^d$

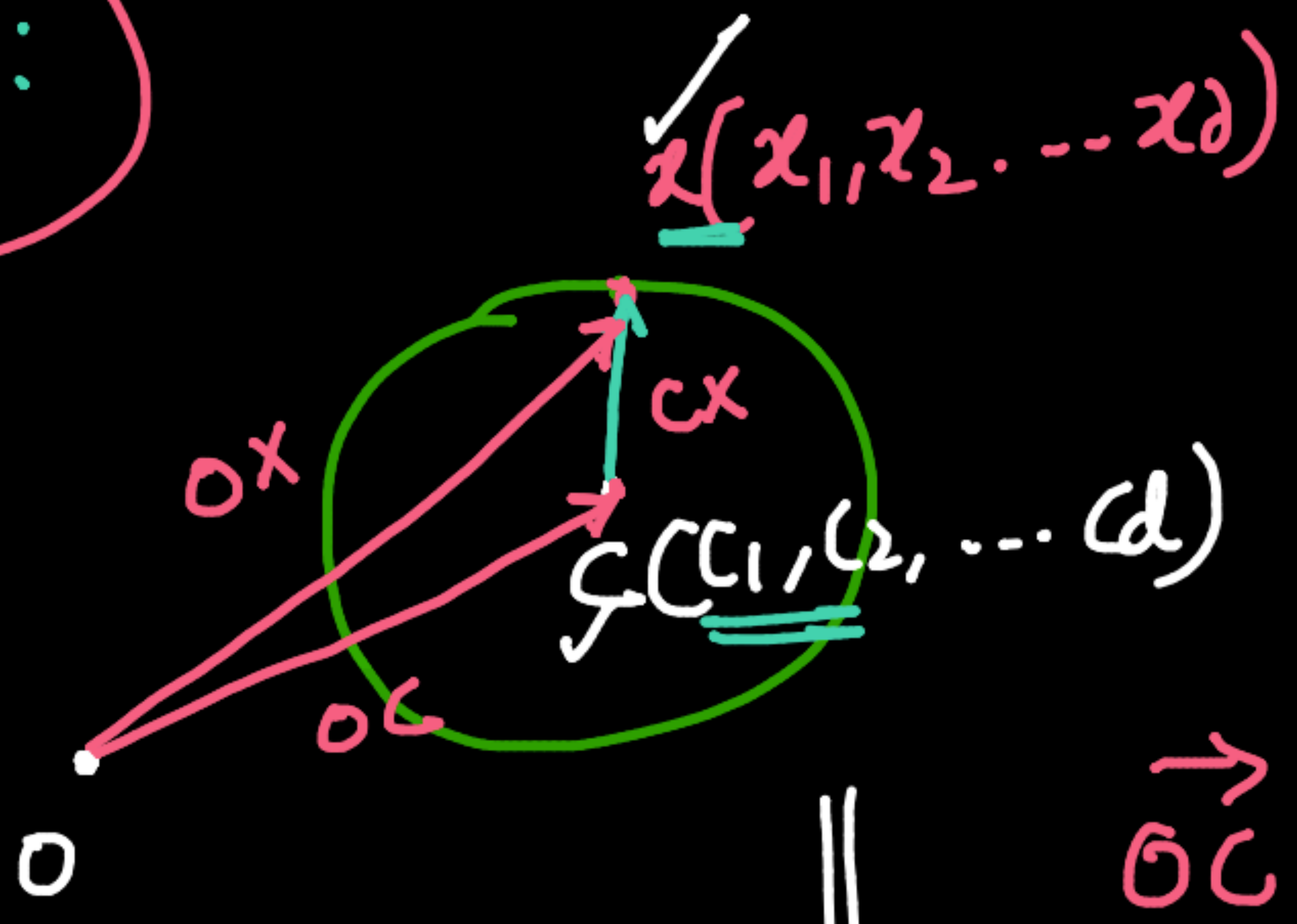
radius $= r$

any x on the Sd

$x \in \mathbb{R}^d$

$c \in \mathbb{R}^d$

$O = [0, 0, \dots 0]^T$



$$S_d : \overbrace{\|\vec{Cx}\| = r}$$

$$\vec{OC} + \vec{CX} = \vec{OX}$$

$$\vec{CX} = \vec{OX} - \vec{OC}$$

$$\|\vec{CX}\| = \|\vec{OX} - \vec{OC}\| = r$$

$$\|\vec{CX}\| = \left\| \begin{bmatrix} x_1 - c_1 \\ x_2 - c_2 \\ \vdots \\ x_d - c_d \end{bmatrix} \right\| = r$$

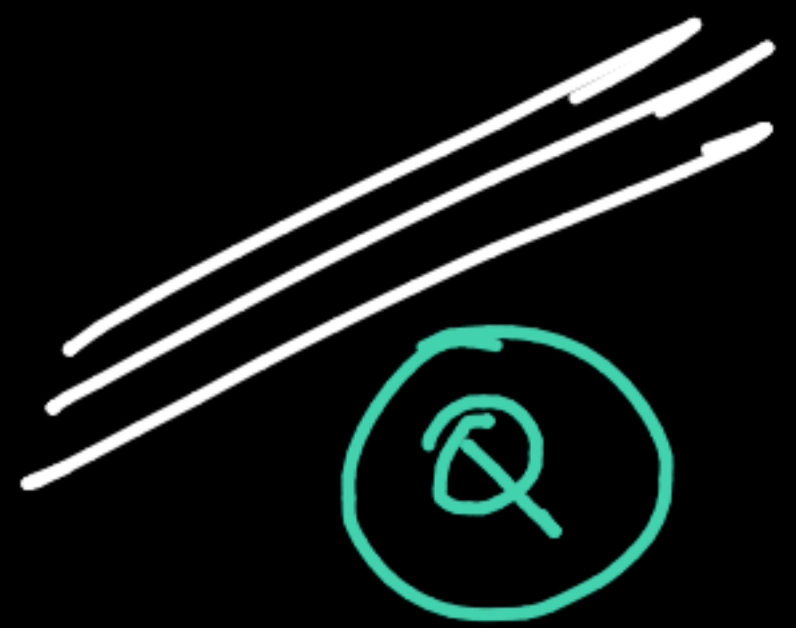
$$\left\| \begin{bmatrix} x_1 - c_1 \\ x_2 - c_2 \\ \vdots \\ x_d - c_d \end{bmatrix} \right\| = r$$

$[x_1 \cdots x_d]^T$ on S^d

$$\Rightarrow \boxed{(x_1 - c_1)^2 + (x_2 - c_2)^2 + \cdots + (x_d - c_d)^2 = r^2}$$

S^d

$\Rightarrow$



Sd: Seperator → inside / outside

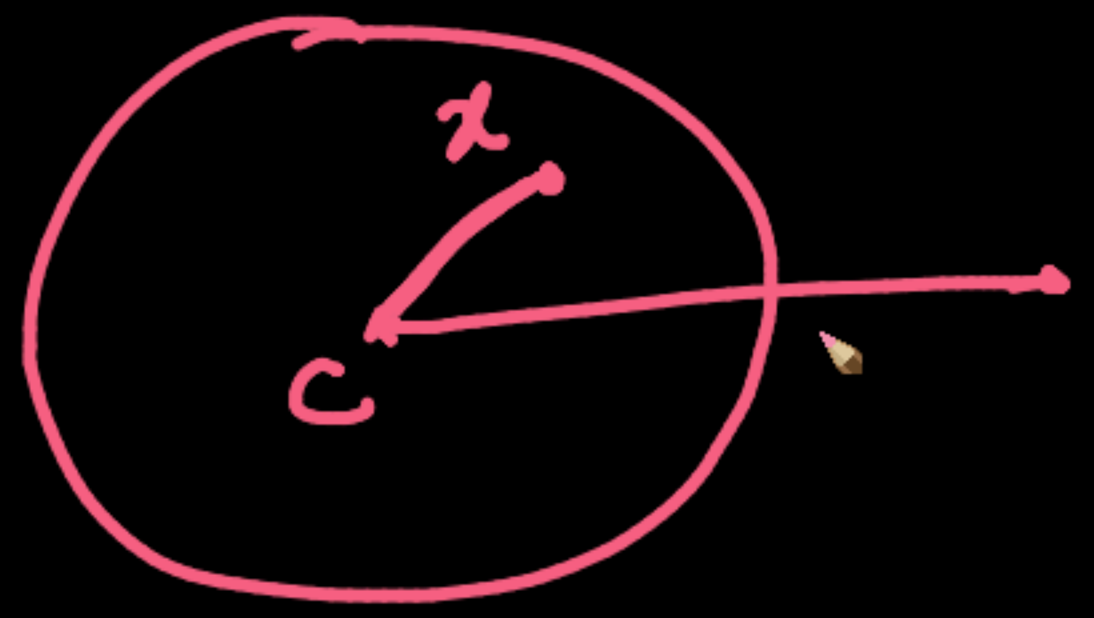
$$\sum_{i=1}^{d}(x_i - c_i)^2 - r^2 = 0$$

inside:

$$\sum_{i=1}^{d}(x_i - c_i)^2 - r^2 < 0$$

outside:

$$> 0$$



Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:15

Linear Algebra

basis →

→ Eigen values & eigen vec → PCA → calculus

→ Matrix operations → SVD ; MF ; rotation

SVD ; MF → RecSys

rotation → CV

covered

LA in the context

29 / 30

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:15

✓ Biz-problem (Fish-sorting / classfn)

intuition (geometry)

1st time

Learn

Math

(CG, LA, Calc)

✓

solve the problem & code

30 / 30

ML

numerical

categorical features (discrete r.v)

Fish: len, width ; shape → Type1 or Type2

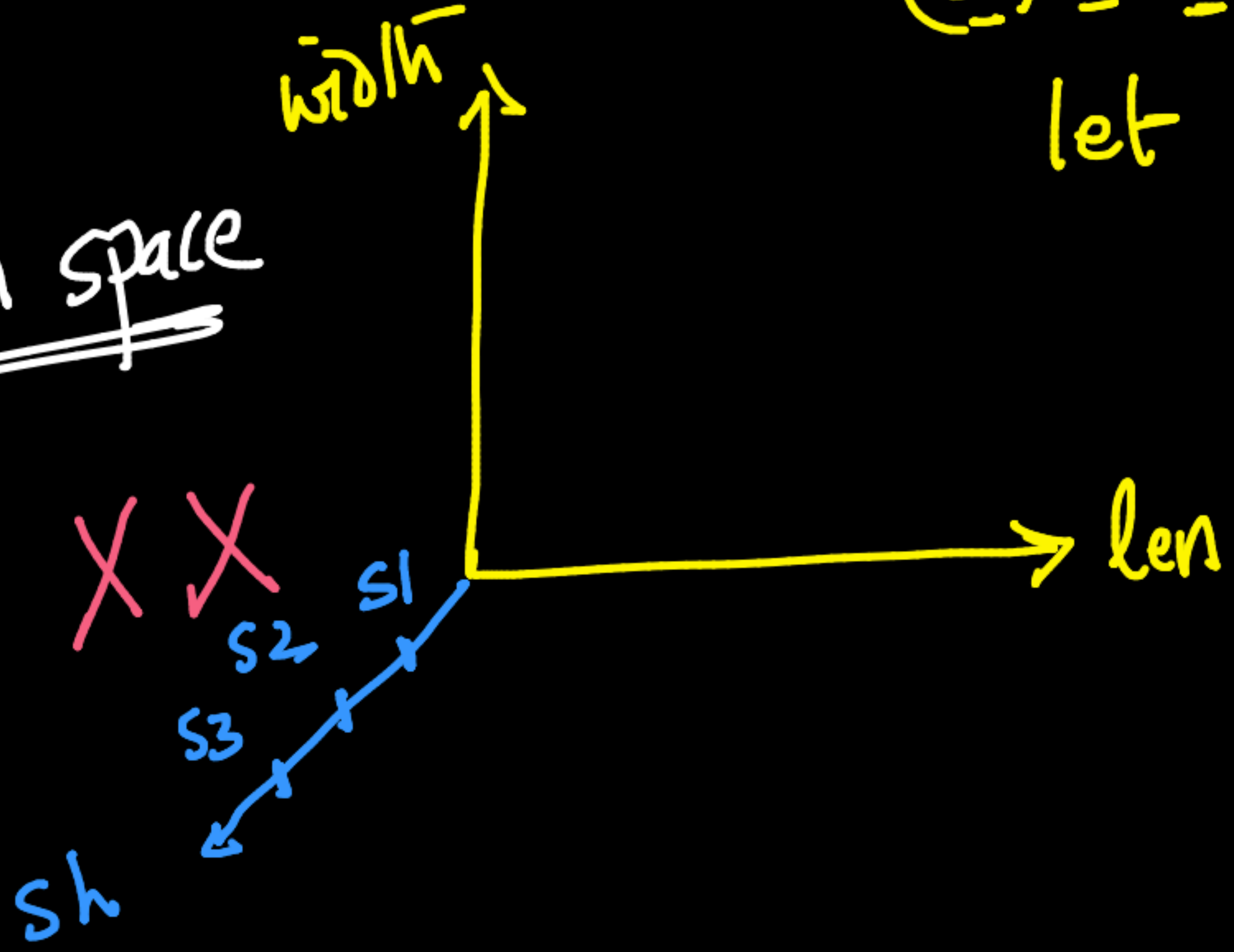
△ ▯ ○
(S1, S2, S3)

let

$$\begin{cases} S1 = 1 \\ S2 = 2 \\ S3 = 3 \end{cases}$$

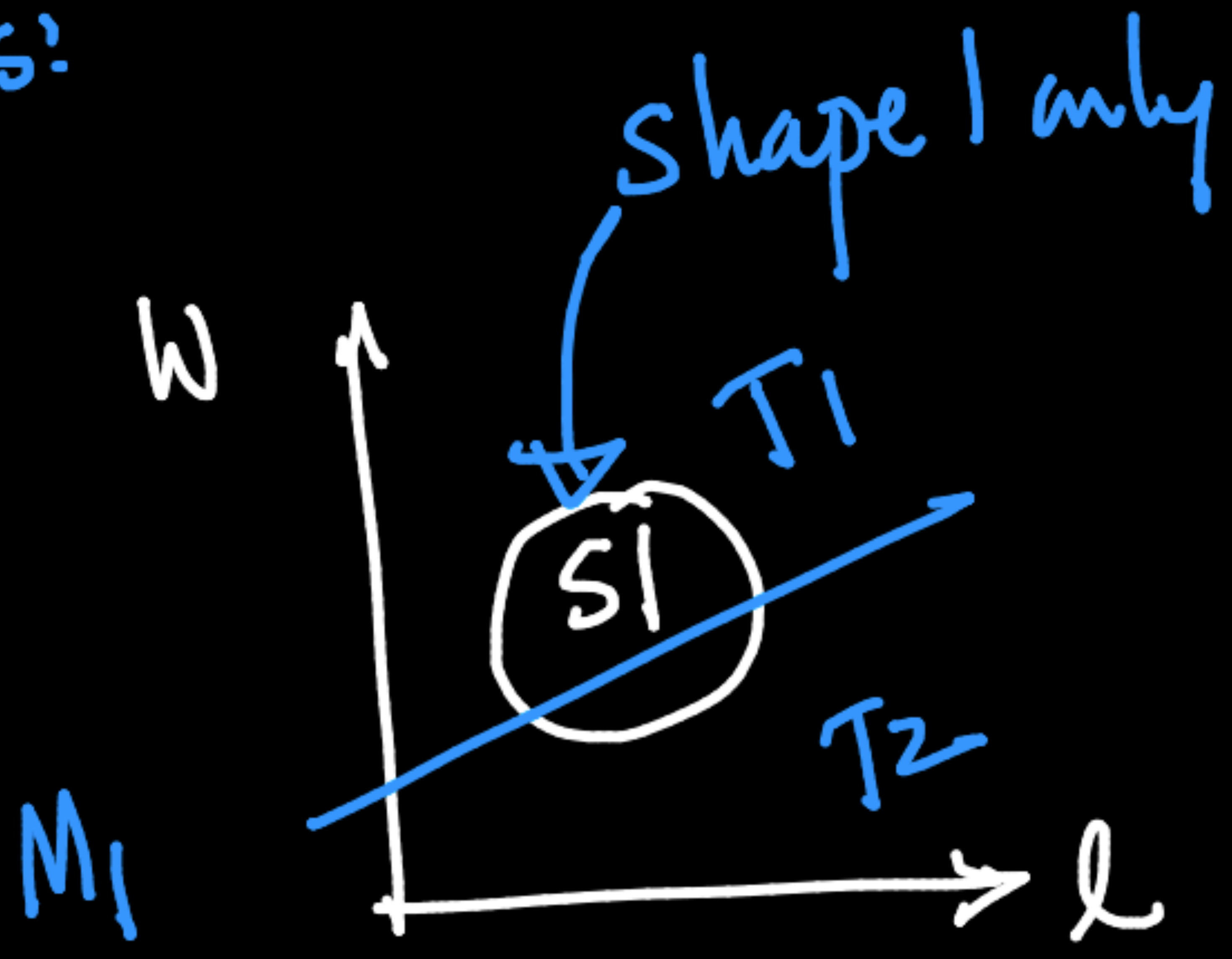
problem: artificially ordering

d-dim space

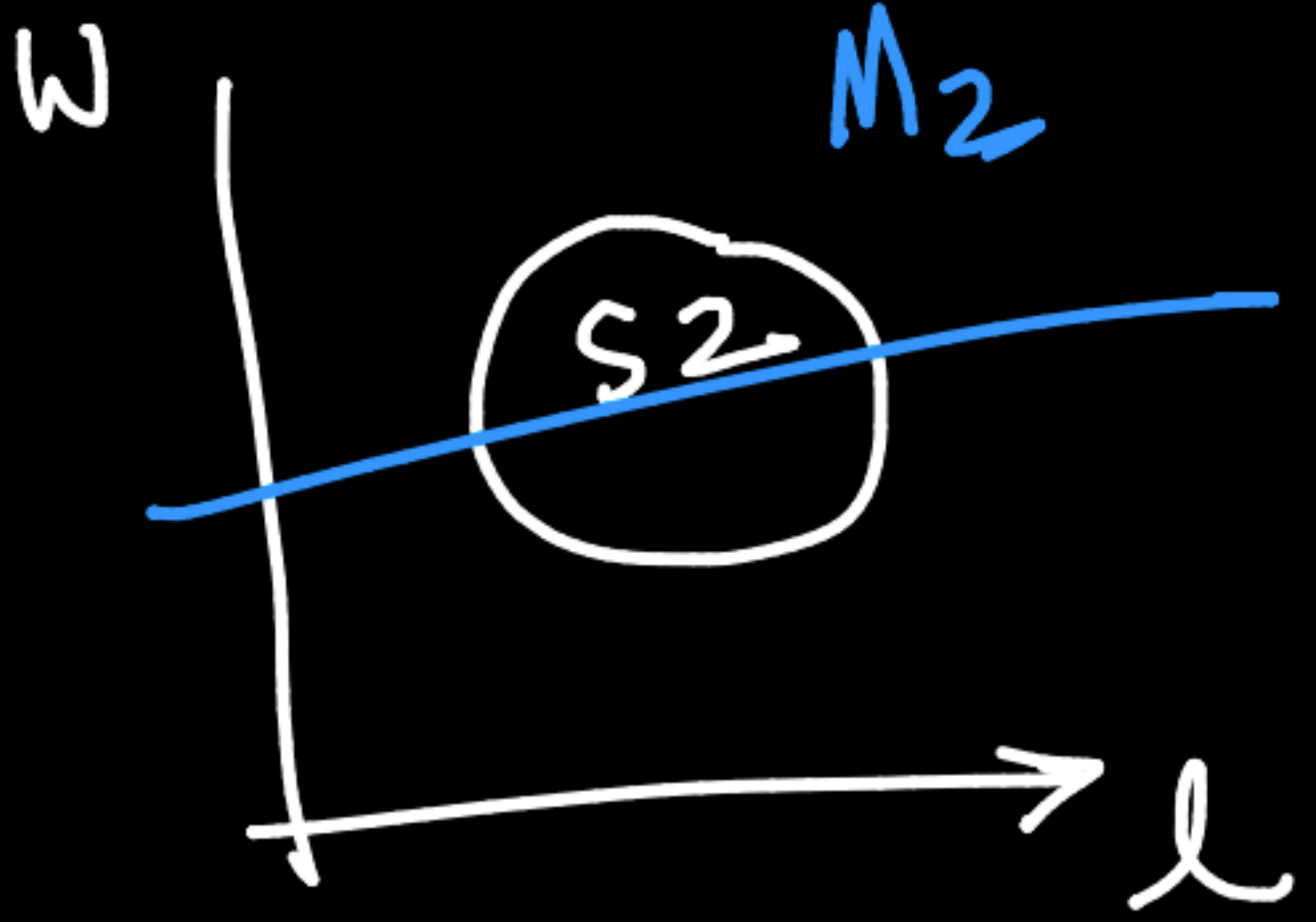


$$\begin{cases} S1 = 2 \\ S3 = 100 \\ S2 = 1000 \end{cases}$$

3-models:



2-dim



2-dim

2-dim

dim- of my space

len ; wid ;

numeric | categorical

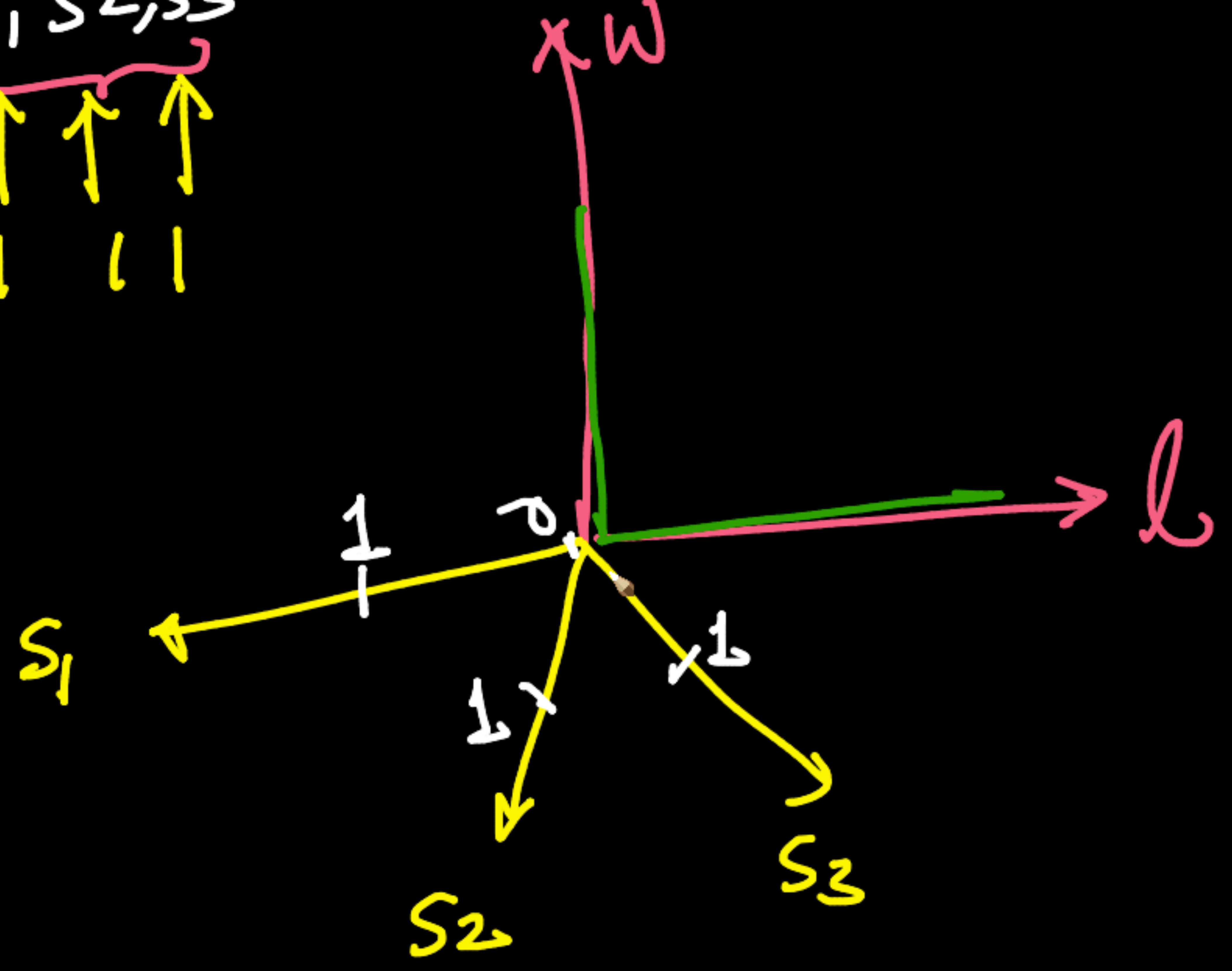
len, wd ; shape

s_1, s_2, s_3

1 1 | 1 1 1

5-dim

d-dim space



5-dim space

5-dim vec

$$x_1 = \langle \overset{l}{10}, \overset{w}{12}, \overset{s}{\boxed{S1}} \rangle \longrightarrow \begin{array}{ccccc} l, & w, & \overset{is}{S1}, & \overset{is}{S2}, & \overset{is}{S3} \\ \lbrack 10, & 12, & 1, & 0, & 0 \rbrack \end{array}$$

shape

one-hot encoding

Q

$$x_2 = \langle \overset{l}{11}, \overset{w}{2}, \overset{s}{S_2} \rangle$$

$$\downarrow$$

$$[11, 2, 0, 1, 0]^T$$

Q

fish → len, wid ; color , shapes , bone-density

catg (above color) · catg (above shapes) · ordinal (above bone-density)

len, wid → numerical

color → $C_1 \dots C_{10}$

shapes → $S_1, \dots, S_4$

bone-density → low / medium / high

dim → 2 | 10 | 4 | 3 or 1

10	0	1
5	1	2
0	2	3

l : -1

m : 0

h : 1

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:55

Best Sep π_d

Given:

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{n=1000} \; ; \; \underbrace{x_i \in \mathbb{R}^d}_{} \; ; \; y_i \in \{-1, +1\}$$

d-dim

0 & 1

-ve

Type 1

+ve

Type 2

d-dim feature vector

Class-label

37 / 44

One strategy! Brute force (Extremely Terrible)

$\Pi_d : w^T x + w_0$

w → d-dim

w_0 → 1

(d+1) → variables

↓

∞-many combinations

$\{w_0, w_1, w_2, \ldots w_d\}$

w_0	w_1	w_2	
1	0	2	- - -
1	2	2	- - -

→ [minimal] misclassification rate (- - -)

& max dist of points from Π_d

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:43

3D-space

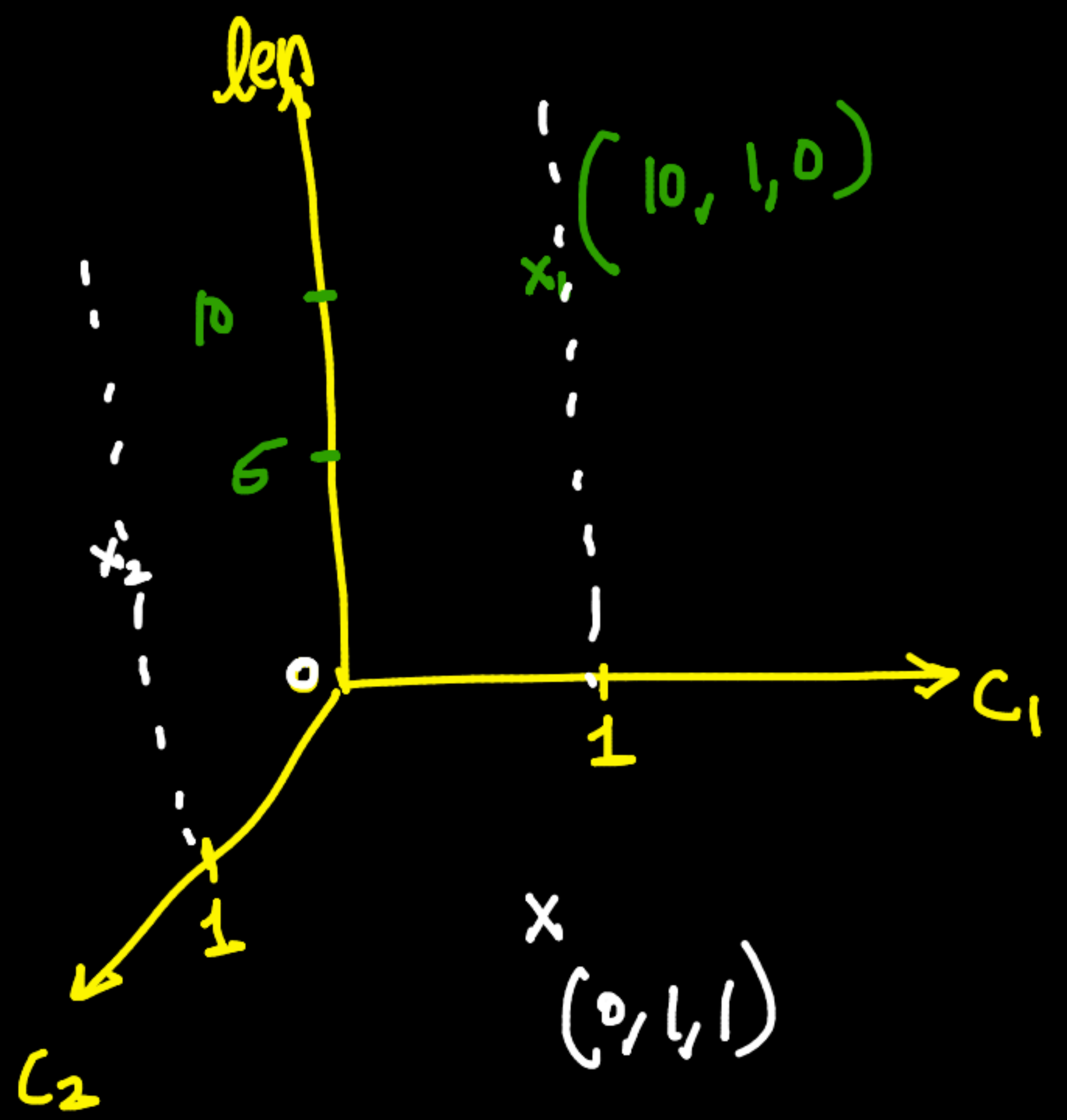
l, color

c_1, c_2

↓ 1 dim ↓ 2-dim

$x_1 = 10, C_1$

$x_2 = 5, C_2 \rightarrow (5, 0, 1)$



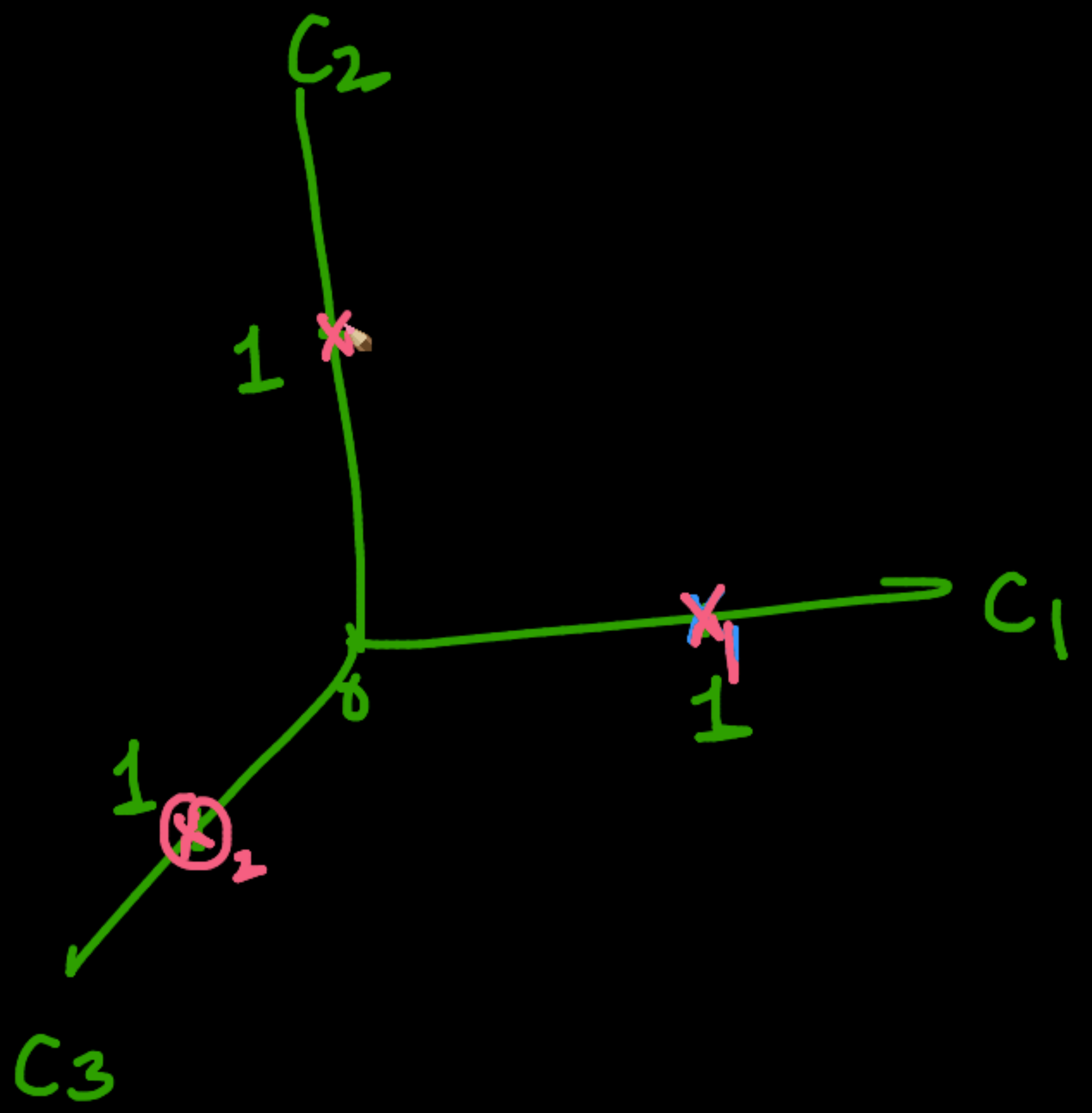
39 / 39

Fish $\rightarrow$ color

c_1, c_2, c_3

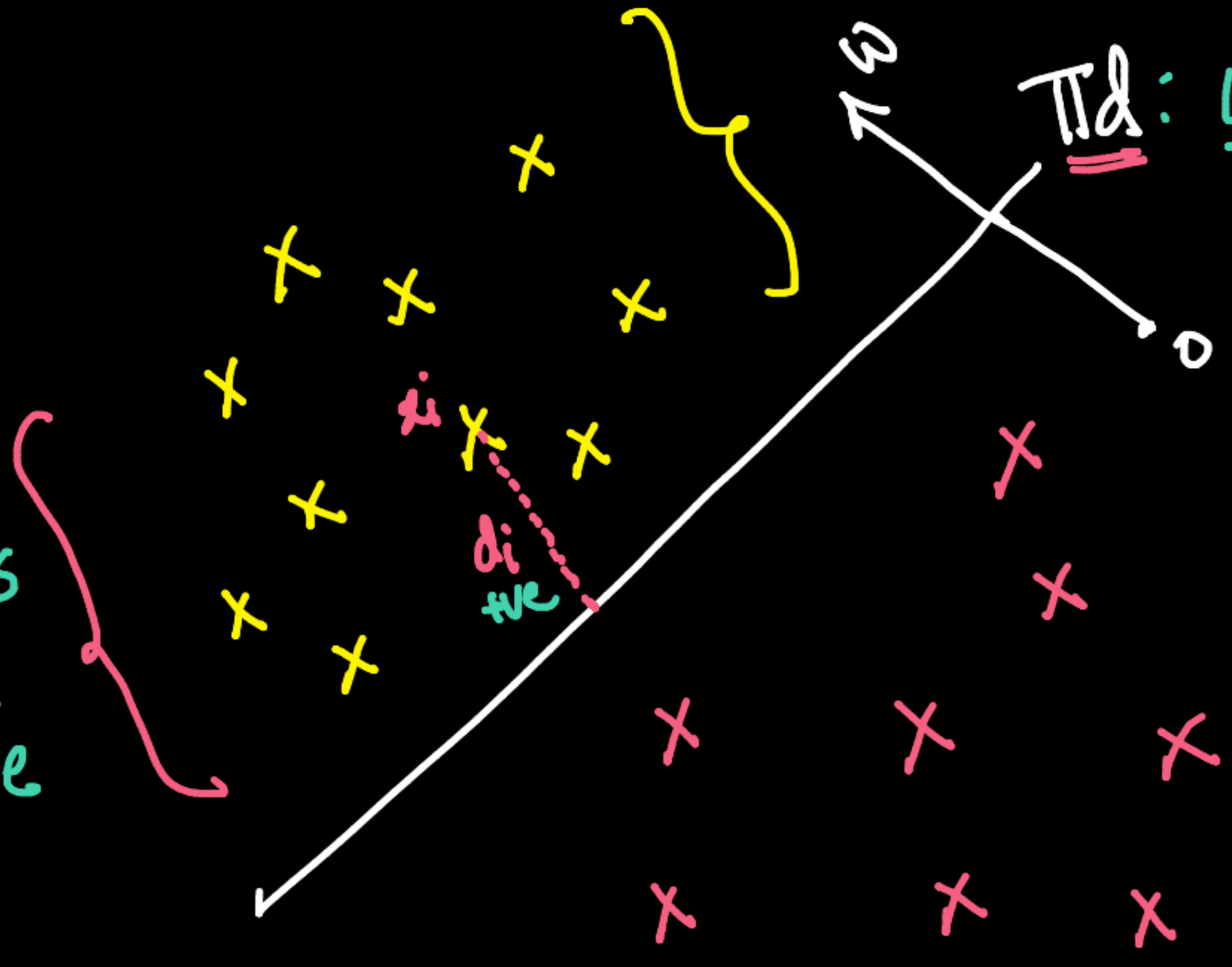
$X_1 \rightarrow c_1 \rightarrow (1,0,0)$

$X_2 \rightarrow c_3 \rightarrow (0,0,1)$



$\Pi_d : w^T x + w_0 = 0$

binary classification



class = y_i's → +ve

+1

class = y_i's = -ve

-1

for each +ve labelled points

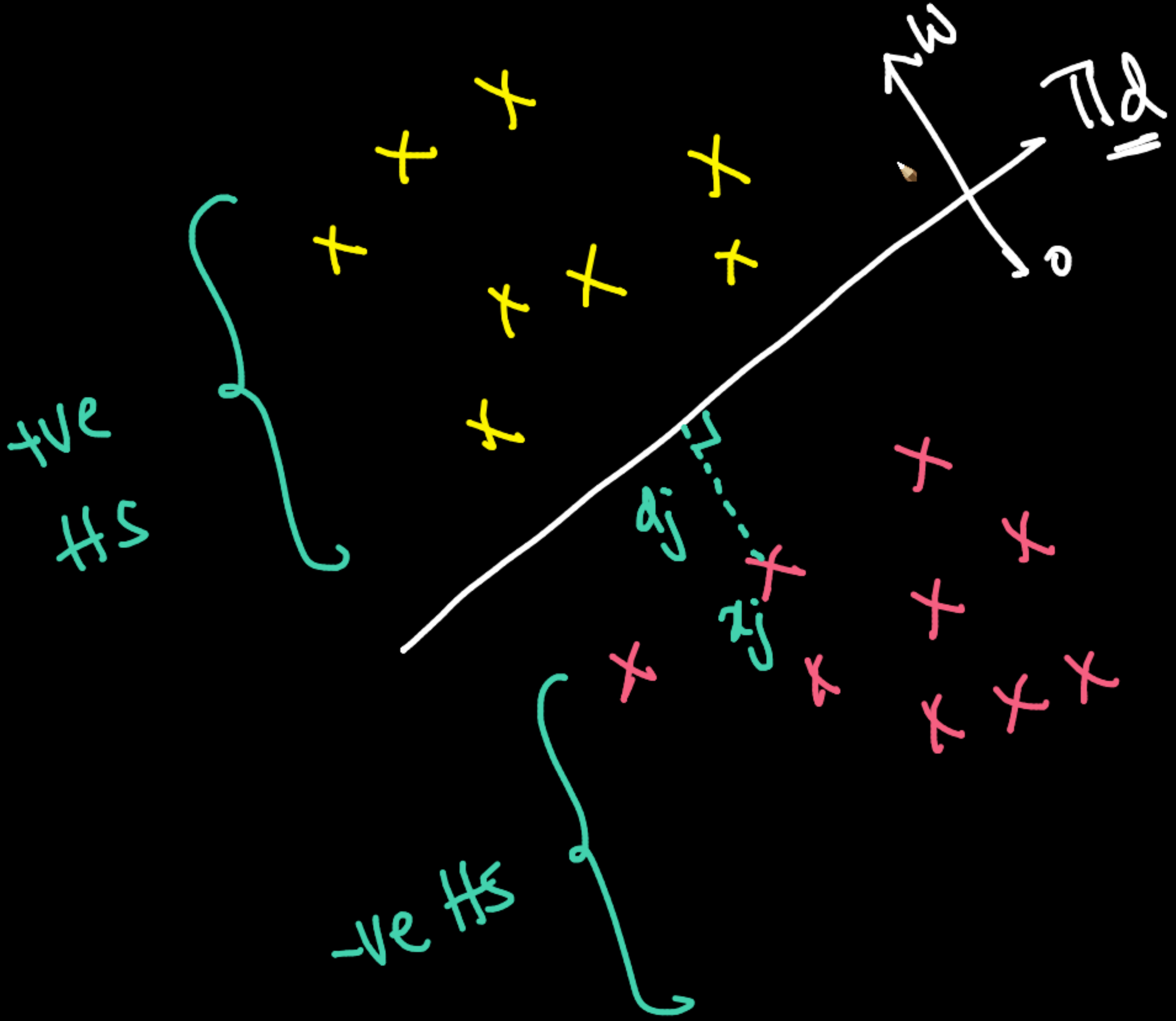
$y_i > 0$

↓

$w^T x_i + w_0 > 0$ — +ve half space (1)

d_i as large as possible

$$\frac{w^T x_i + w_0}{\|w\|} \text{ maximal} \quad (2)$$



For all -ve labelled points ($y_i<0$)

① → $w^T x_j + w_0 < 0$

② $d_j = \frac{w^T x_j + w_0}{\|w\|}$

dj's are -ve

dj's → large -ve values

+ve labelled: $y_i = +1$

$$\frac{\overbrace{w^T x_i + w_0}^{+ve}}{\|w\|} \cdot \overset{+1}{y_i} \longrightarrow \text{maximal}$$

-ve labelled: $y_i = -1$

$$\frac{\overbrace{w^T x_i + w_0}^{-ve}}{\|w\|} \cdot \overset{-1}{y_i} \longrightarrow \text{maximal}$$

$\boxed{x_i \to y_i}$ $\mathcal{D}$

Find Π_d s.t.

for each point x_i;

$$z_i = \frac{w^T x_i + w_0}{\|w\|} \cdot y_i \longrightarrow \text{Maximal}$$

(as large as possible)

y_i: $+1/-1$

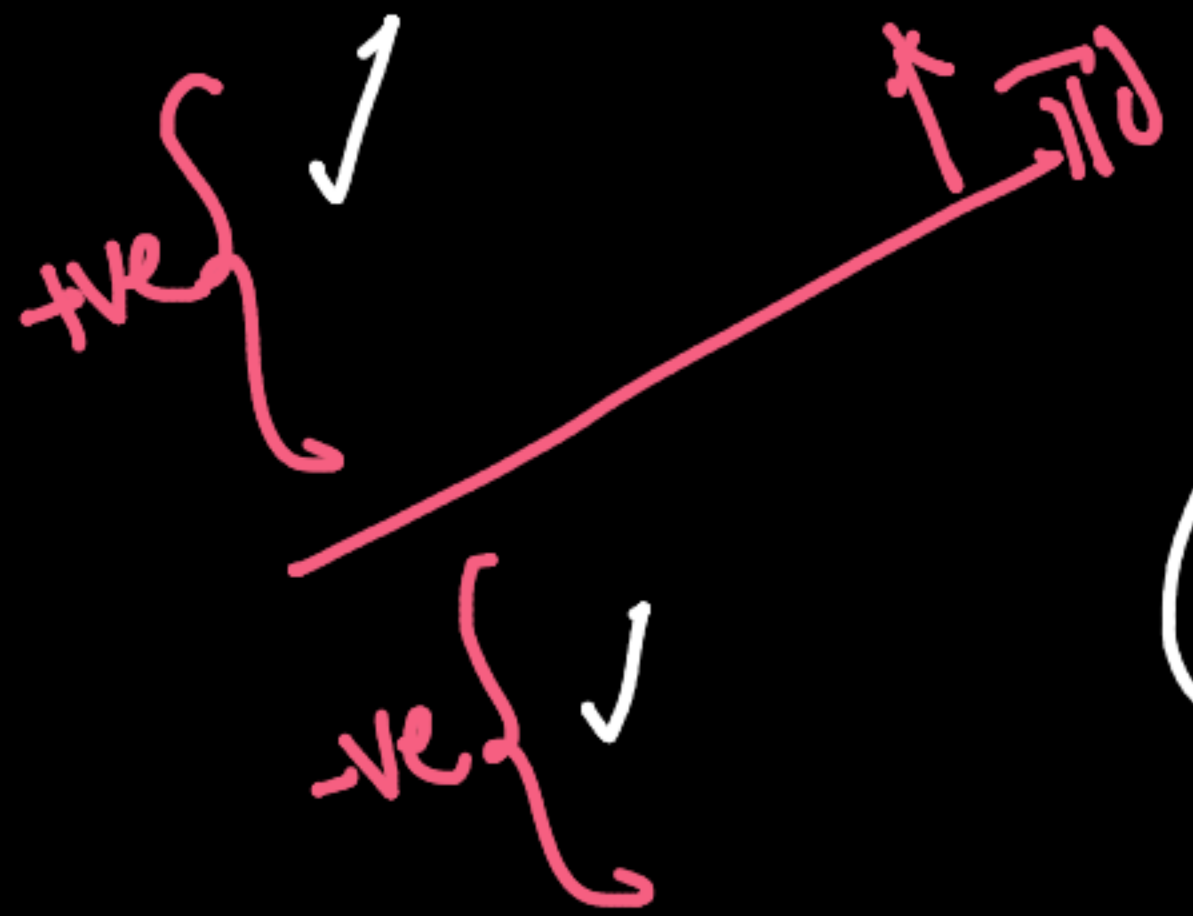
$$\mathcal{D} = \left\{ (x_i, y_i)_{i=1}^{n=1000} - - - - \right\}$$

params (args)

$$\max_{\underline{w}, \underline{w_0}} \overbrace{z_1, z_2, z_3 \dots z_n}$$

simultaneously

$$\max_{w, w_0} (z_1 + z_2 + \dots z_n) = \max_{w, w_0} \sum_{i=1}^{n} z_i$$



π_d

$$\boxed{\max_{\underset{d}{w}, \underset{1}{w_0}} \sum_{i=1}^{n} y_i \frac{(w^T x_i + w_0)}{\|w\|}}$$

optimization

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:16

Q

$$\overbrace{\max \quad \underline{z_1, z_2}, \ldots z_n}$$ Simultaneously

$$\downarrow$$

$$\underbrace{\max \sum_{i=1}^{n} z_i}$$ ✓

✓

Q why not

x_3 lies on Π_2

$$\boxed{0}$$

$$\max \quad z_1 \cdot z_2 \cdot \boxed{z_3} \cdots z_n$$

$$\downarrow$$

$$\underbrace{\max \prod_{i=1}^{n} z_i}$$

48 / 48

detour

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{n} \ \dots$$

$$\max_{w, w_0} \ \overbrace{\sum_{i=1}^{n} \frac{y_i (w^T x_i + w_0)}{\|w\|}}^{f(w, w_0)}$$

$$\max_{w, w_0} f_{\mathcal{D}}(\overbrace{w, w_0}^{\text{dtl var}}) = \min_{w, w_0} \overbrace{- f(w, w_0)}$$

dtl var

one-var

min x $f(x) = x^2 - 2x + 5$

Find x that minimizes $f(x)$

Calculus in one-var

multivariate calc

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:25

11^{th} $\&$ 12^{th}

$$\min_x \boxed{f(x)} = x^2 - 2x + 5$$

next class ← why ?

① $$\frac{d\,f(x)}{dx} = \boxed{2x-2} = 0 \Rightarrow x = 1$$

② $$\frac{d^2 f(x)}{dx^2} = 2 > 0$$

minimal-val of $f(x)$

51 / 52

Maxima & minima

→ Functions ; Limits ; Continuity } one-variable

→ Diff ; slope & geom

→ Maxima & minima

→ partial derivatives } d+1 var multivariate calc

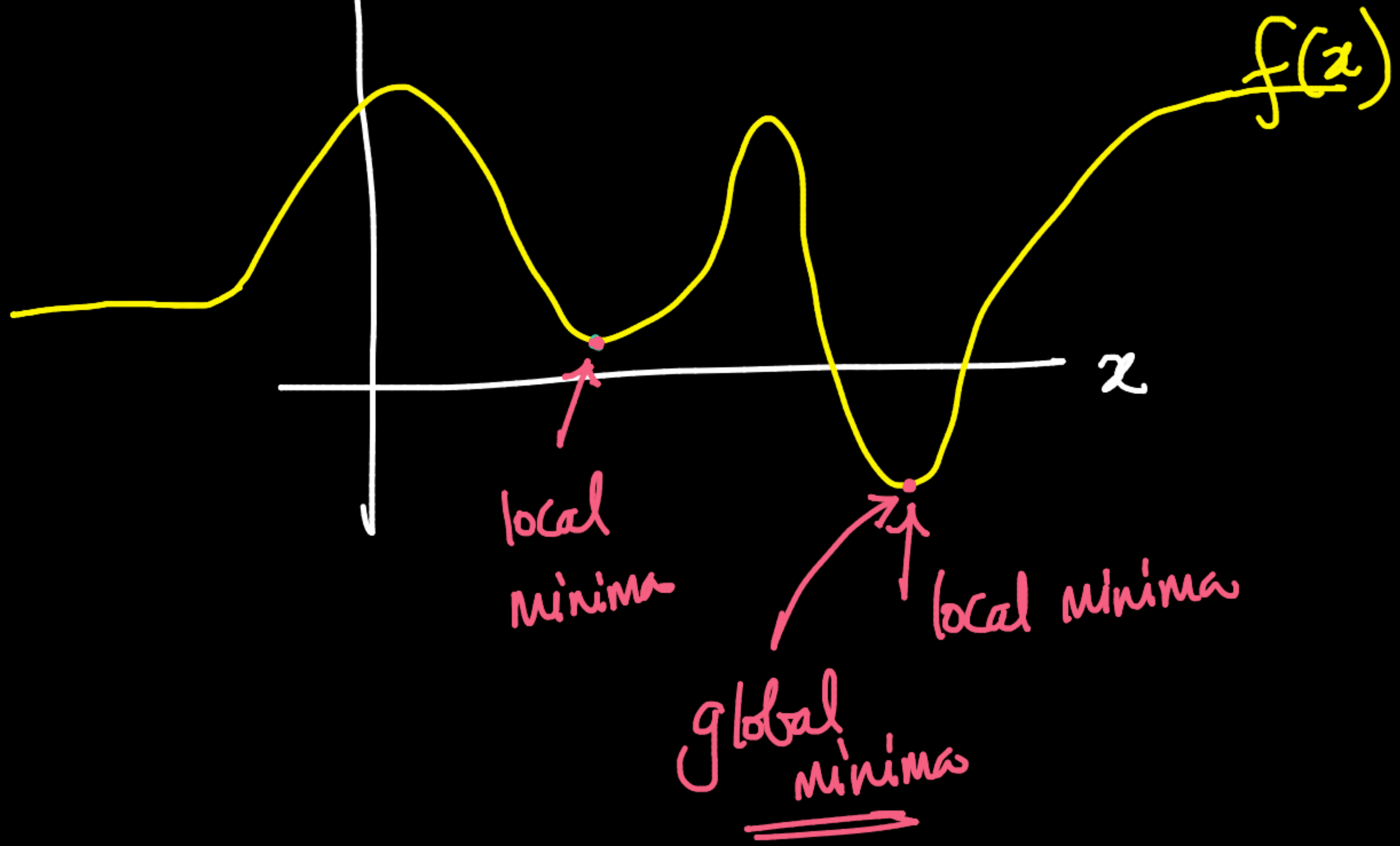
→ Gradient descent

2-3 classes

→ SGD

→ constrained optimization → PCA

$y = f(x)$



You have left the meeting

We get frequent requests for your notes!

Notes written by you helps in understanding the topic better. You can upload the notes in two simple steps mentioned below

1. **Scan the QR code with your iPad**
   Scanner should be present in the top menu on your iPad
2. **Upload Notes on the generated link**
   All notes uploaded will be visible in the saved version of this session

OR

Drag and drop files or click here to upload

Files Uploaded from your computer appear here

54 / 54