Logistic Regression

Mohd Ayyoob
(M.Tech CSA, IISc Banglore)

AGENDA

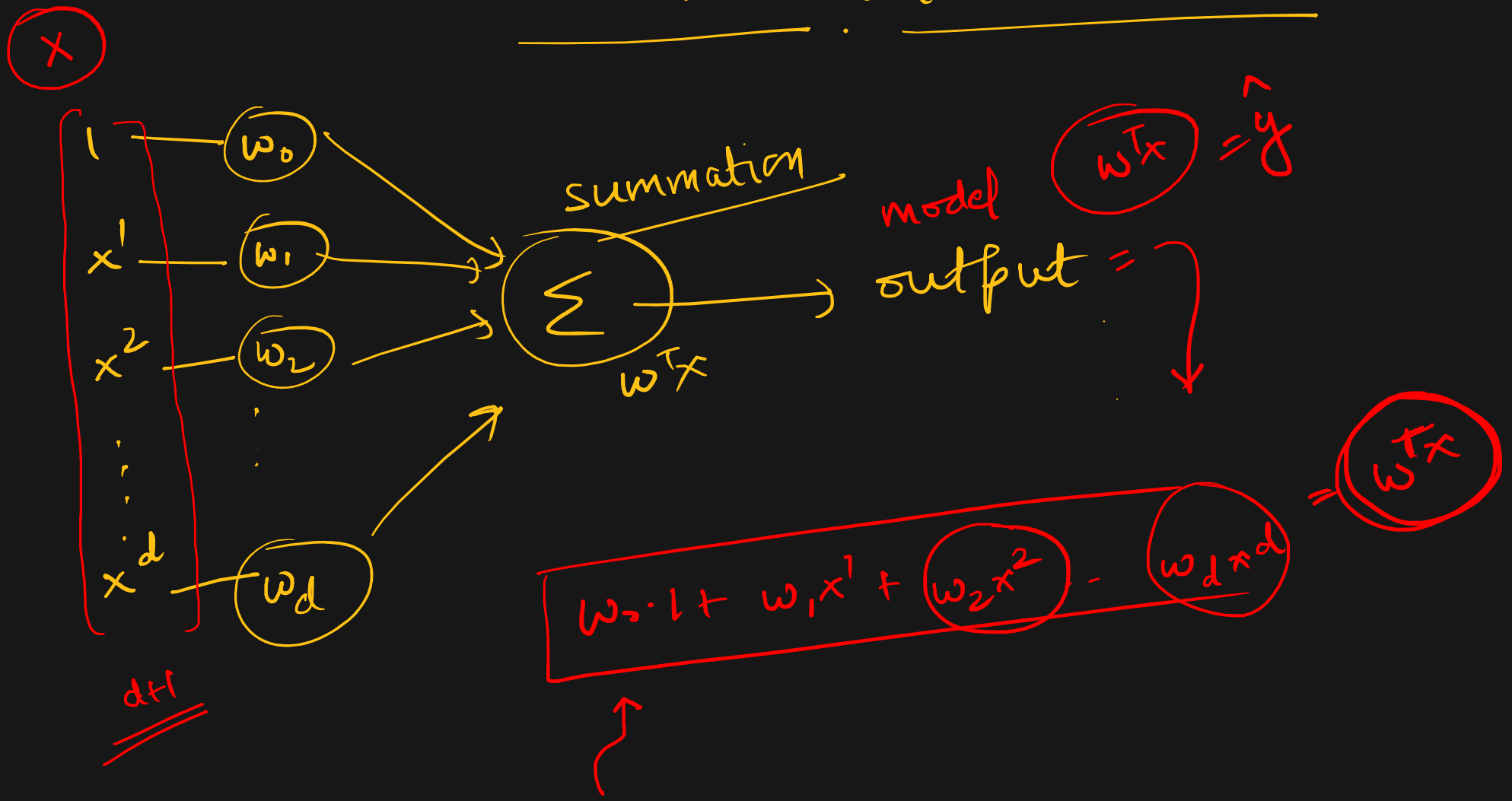
1. Why ‘regression’ in logistic regression when it is a classification algorithm?
2. Our requirement : We want to provide probability of data point belonging to each of the class.
3. How to do we convert output to prob ? Sigmoid function
4. How do we see goodness of our predictions? – Loss function
5. Loss function in logistic regression : cross-entropy

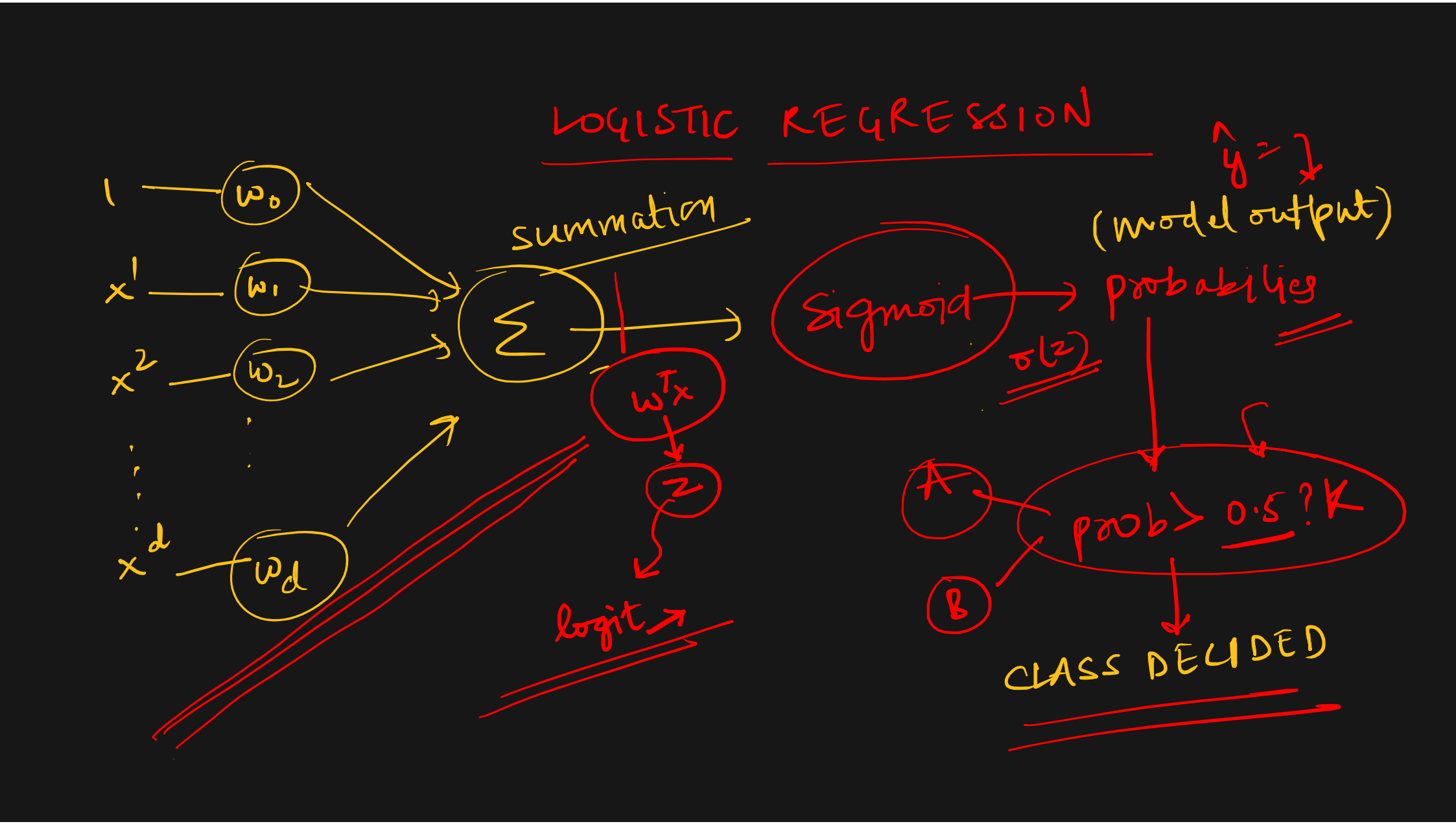
Assumption of logistic regression: ?

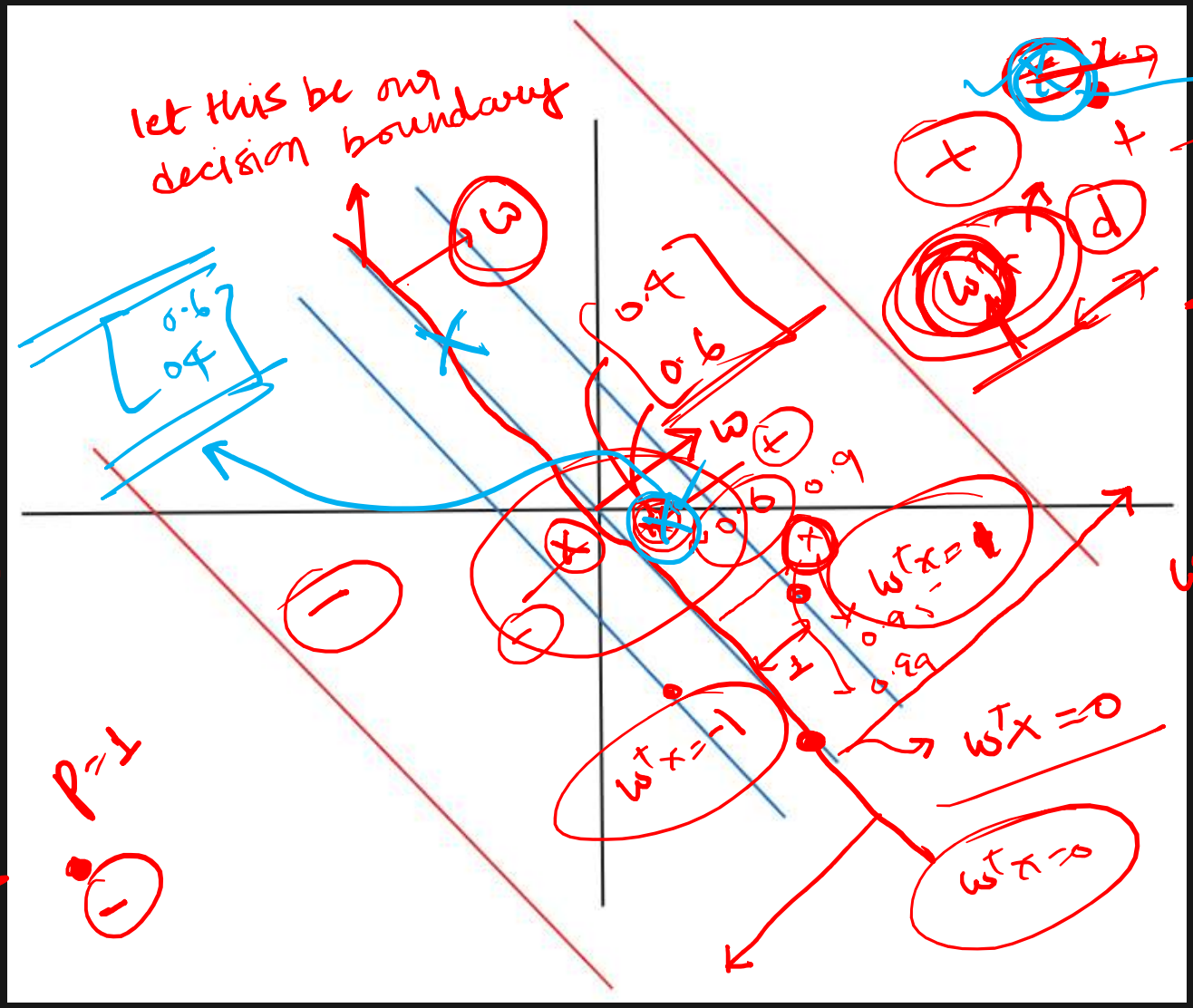
→ data is linearly seperable

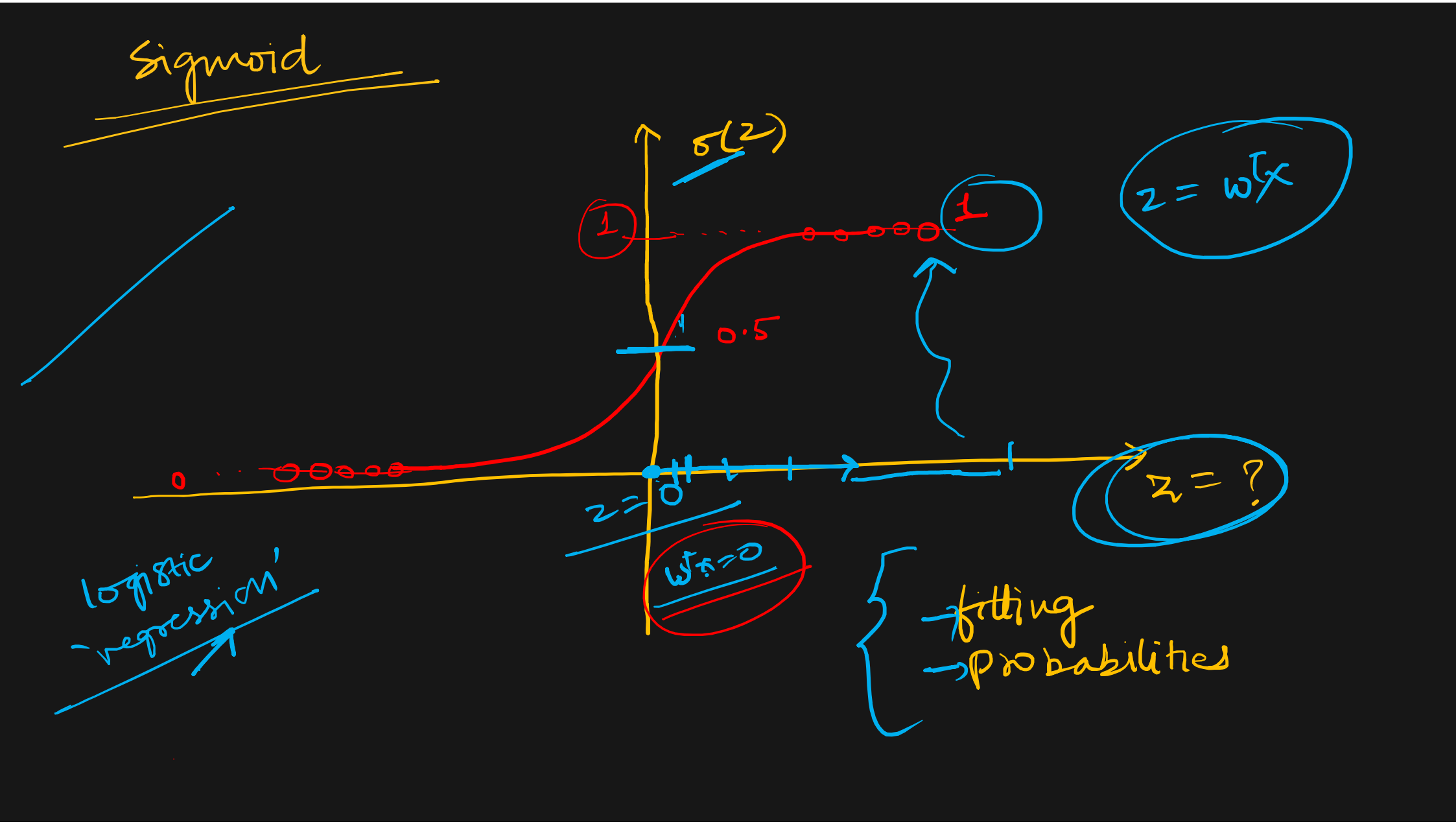
given D, if some says logistic regression give train-acc = 100%. → D is linearly seperable

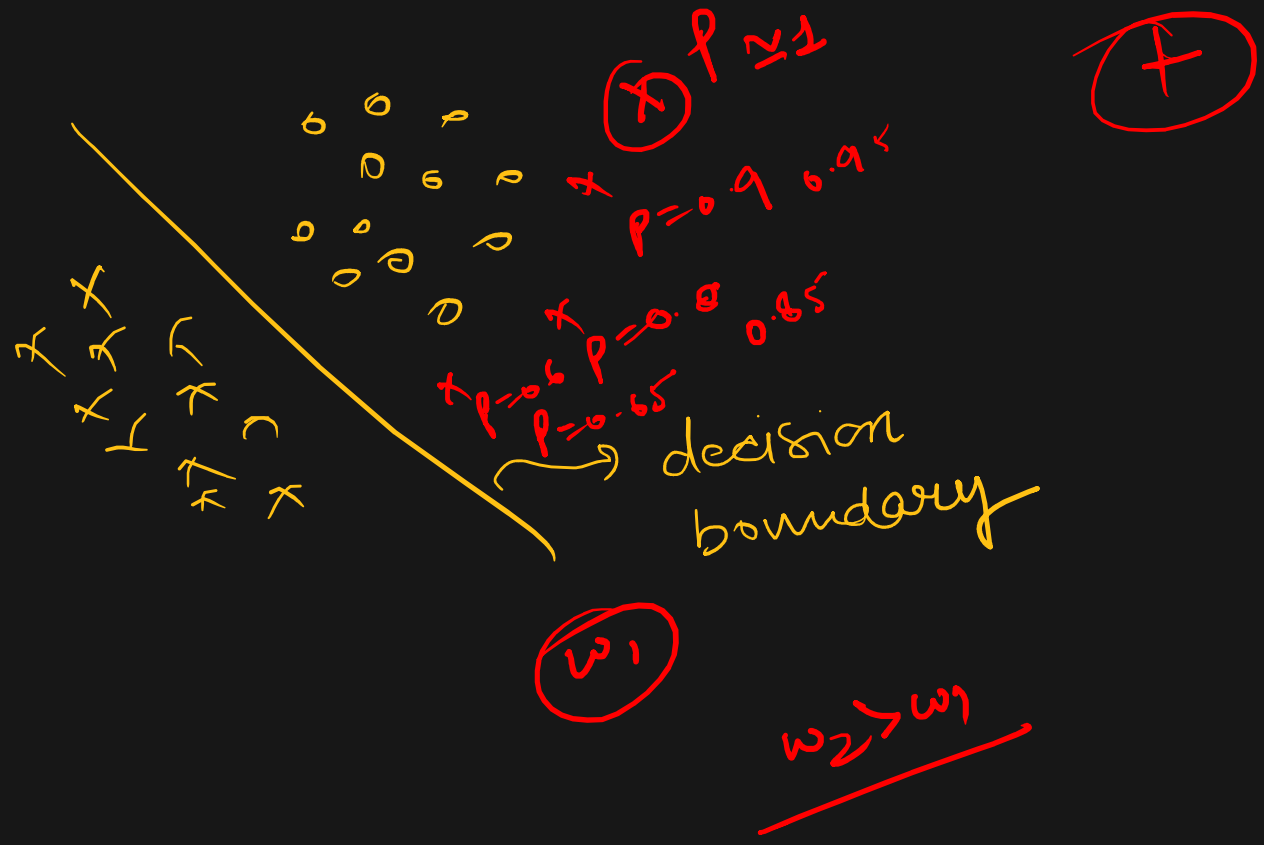
LINEAR REGRESSION




LOGISTIC REGRESSION
1
x^1
x^2
x^d
w_0
w_1
w_2
w_d
summation
Σ
$w^T x$
z
logit
sigmoid
$\sigma(z)$
probability
$\hat{y} =$?
(model output)
A
B
prob > 0.5 ? K
CLASS DECIDED


Why do we need sigmoid?
What if we don't have sigmoid?
let this be our decision boundary
0.6
0.4
0.4
0.6
0.9
0.6
0.1
0.99
$w^T x = 1$
$w^T x = 0$
$w^T x = -1$
$w^T x = 0$
$w^T x$
$\sigma(w^T x) = prob$
$z = w^T x$
$\sigma(z)$
prob
$w^T x \geq 0$
$w^T x_2 = 20$
$w^T x_2 \geq 0$
$w^T x_3 = 0$


Sigmoid
σ(z)
1
1
0.5
0
z = 0
z = ?
z = w^T x
w^T x = 0
logistic regression
fitting
probabilities


p ≈ 1
p = 0.9
0.9
p = 0.8
0.85
p = 0.65
decision boundary
w1
w2 > w1

Q) for a point x on decision boundary of logistic regression, is $\sigma(w^T x) = w^T x$?

a) Yes

b) No

c) No idea

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$(z = w^T x)$

$\sigma(w^T x) =$

$\sigma(0) = 0.5 \neq ?$ $w^T x \neq 0$

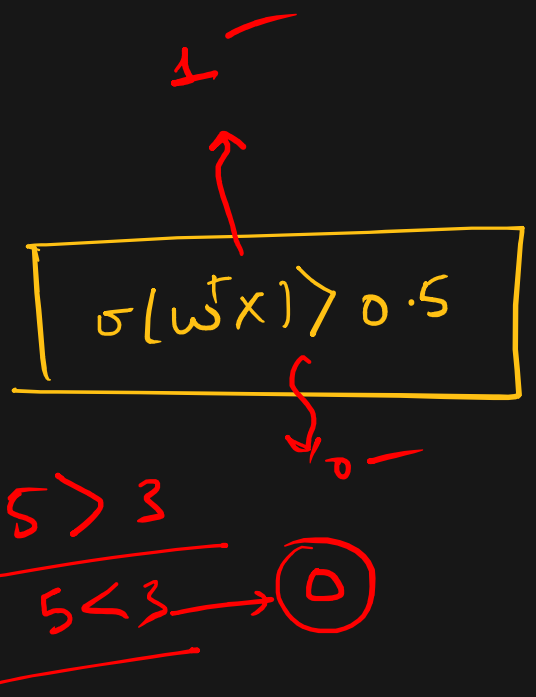
Q) For case of logistic regression, given a random x and w. For this, what would be the model's output

a) $w^T x$ → logit (z)

b) $\sigma(w^T x)$ → prob $(\hat{y})$

c) logical value of this operation

d) NONE

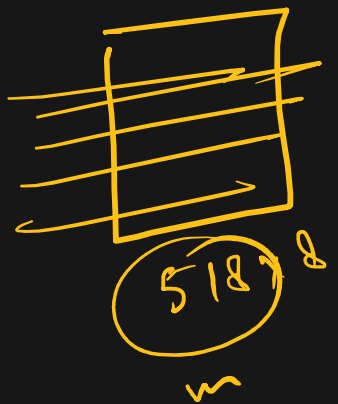


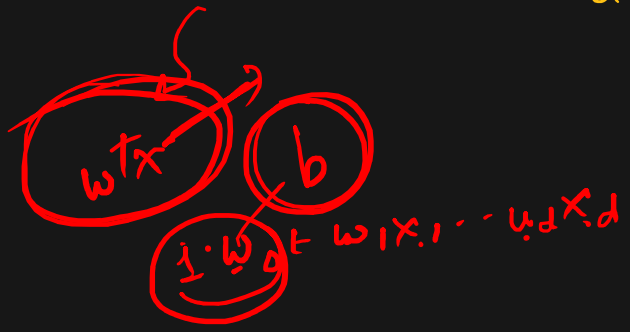
$$\frac{d(\sigma(z))}{dz} = ? \quad \sigma(z)(1-\sigma(z)) = \sigma'(z)$$

max-value of $\sigma(z)$? : max 1 min. 0

max-value of $\sigma'(z)$? : 0.25 ↗ $z=0$

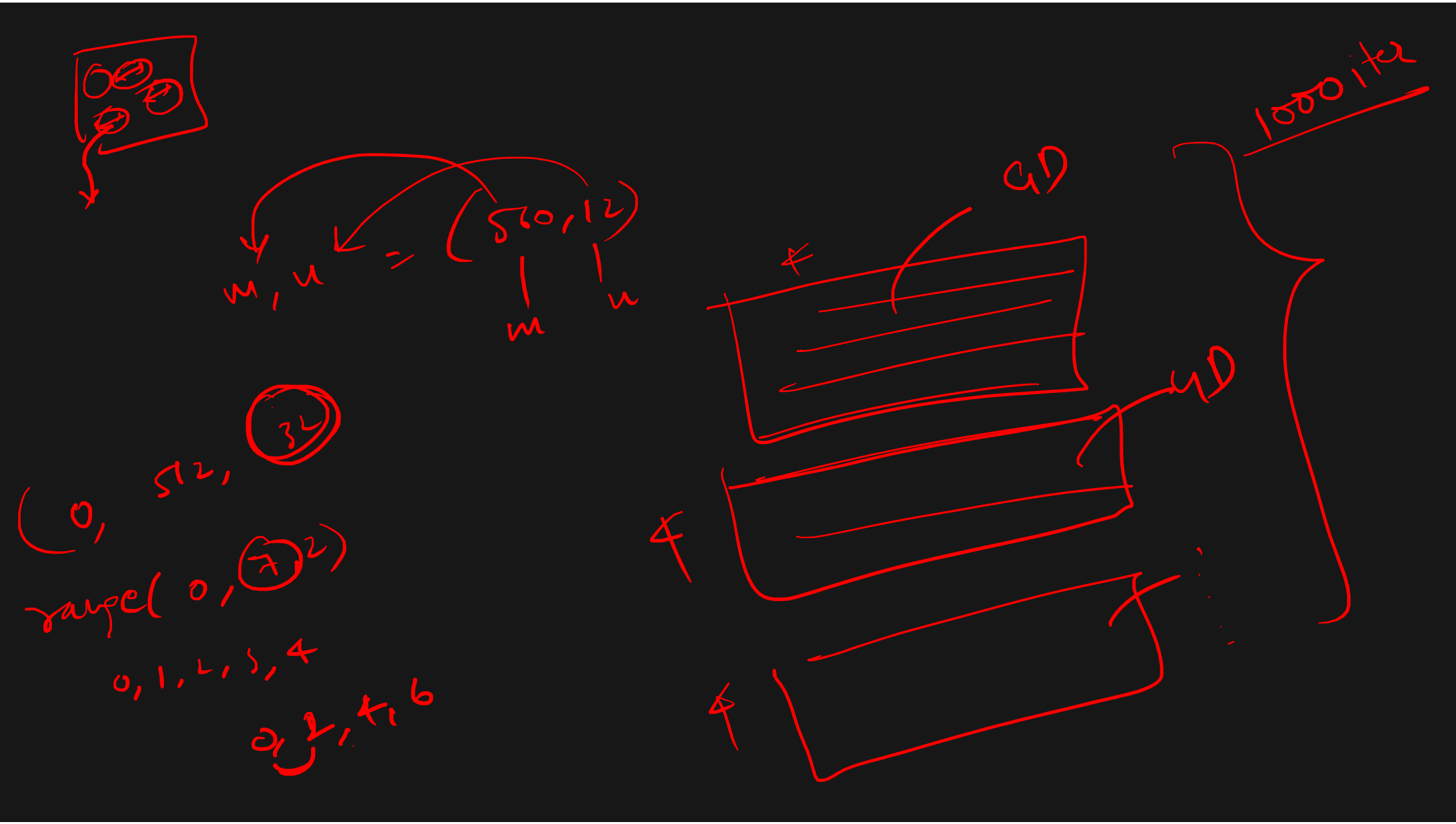
derivative

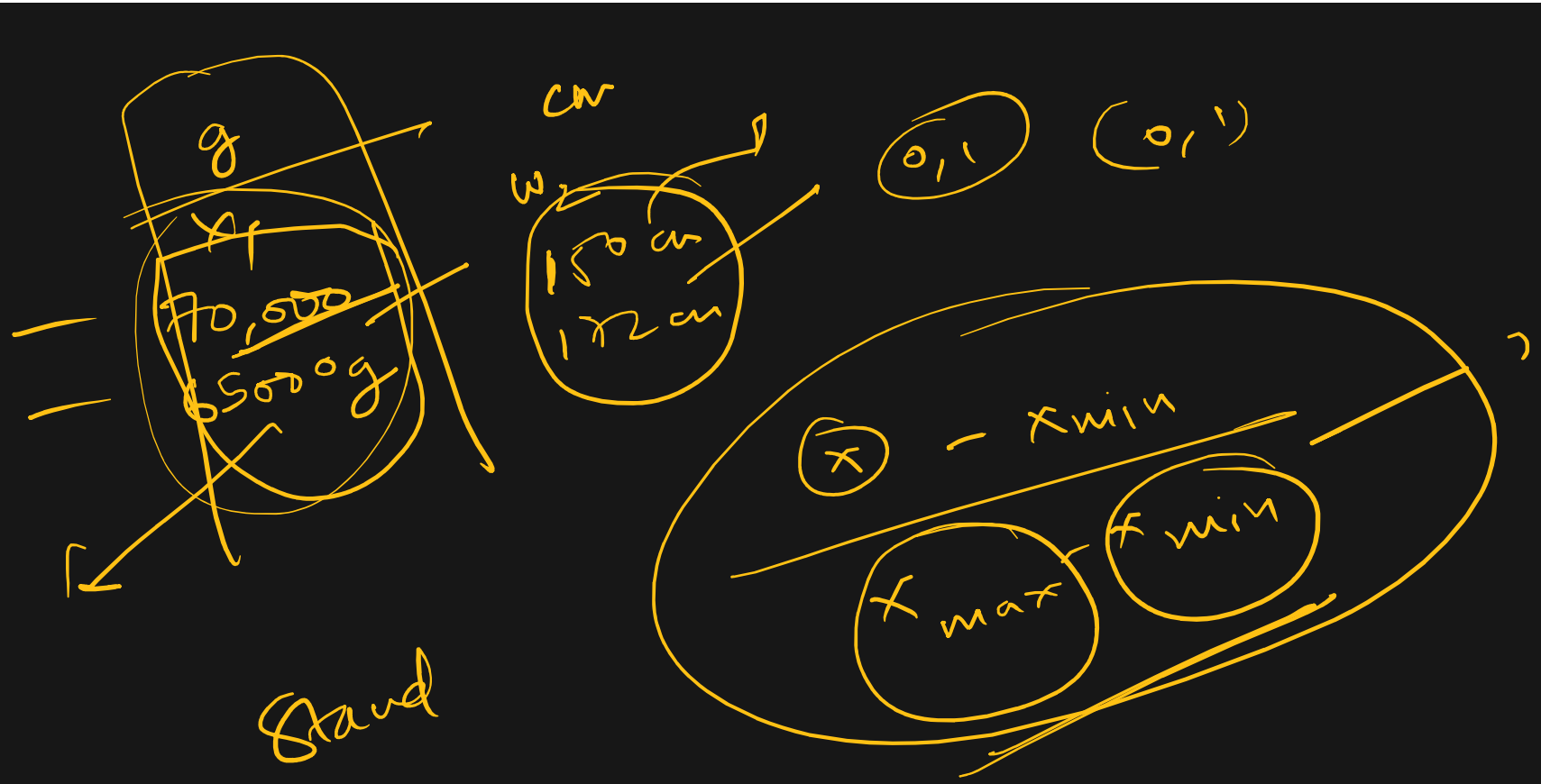




$x(1-x)$

$\sigma(w^T x + b)$


m, n = (50, 12)
m
n
(0, 512,
32)
range(0, 7, 2)
0, 1, 2, 3, 4
0, 2, 4, 6
GD
GD
1000 iter


g
70,000
5500 g
cm
150 cm
182 cm
0,1
(0,1)
x - xmin
xmax
xmin