

Sigmoid

$\text{sig}(\text{val}) \rightarrow [0, 1]$

$\text{sig}(x) = \frac{1}{1 + e^{-x}}$

$$\omega^* = \underset{\omega}{\text{argmax}} \sum_{i=1}^n \frac{1}{1 + \exp(-y_i \omega^T x_i)}$$
  
↑  
optimal

sklearn → Logistic

$$\omega^* = \underset{\omega}{\text{argmax}} \sum_{i=1}^n y_i \omega^T x_i$$

$$= \underset{\omega}{\text{argmax}} \sum_{i=1}^n \text{sig}(y_i \omega^T x_i)$$

$$= \sum_{i=1}^n \frac{1}{1 + \exp(-y_i \omega^T x_i)}$$

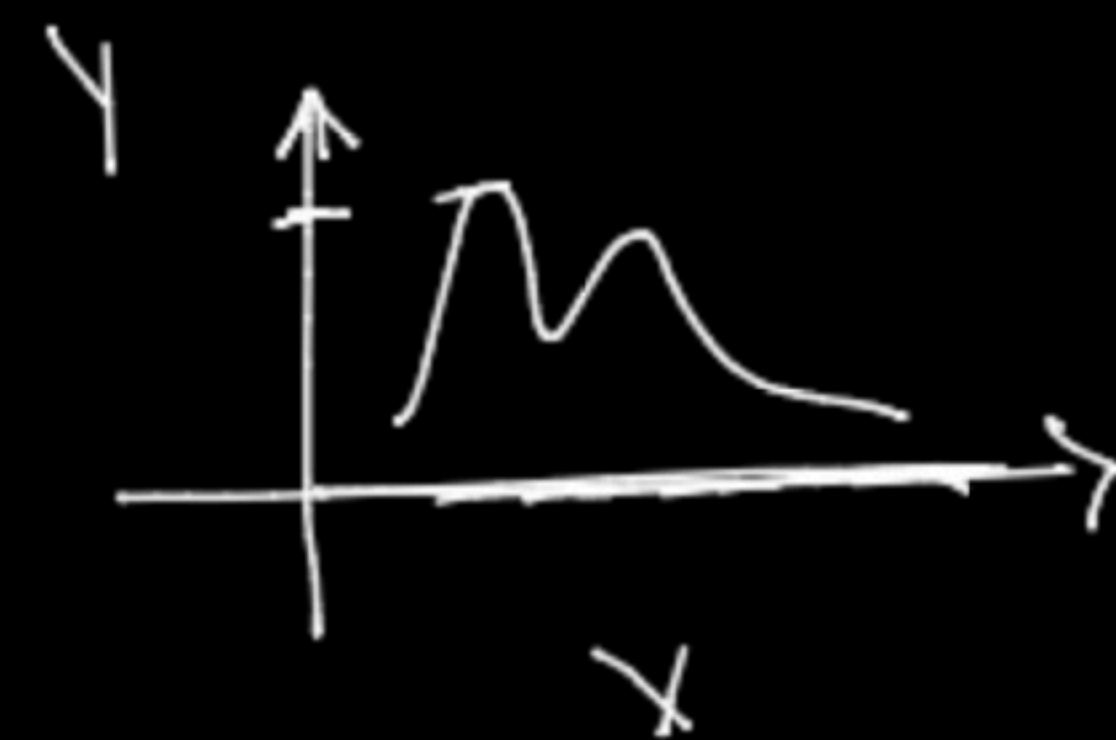
Model

Optimisation

$x \rightarrow \square \rightarrow f()$

$y = f(x)$

Max  $f(x)$  →  $x^*$



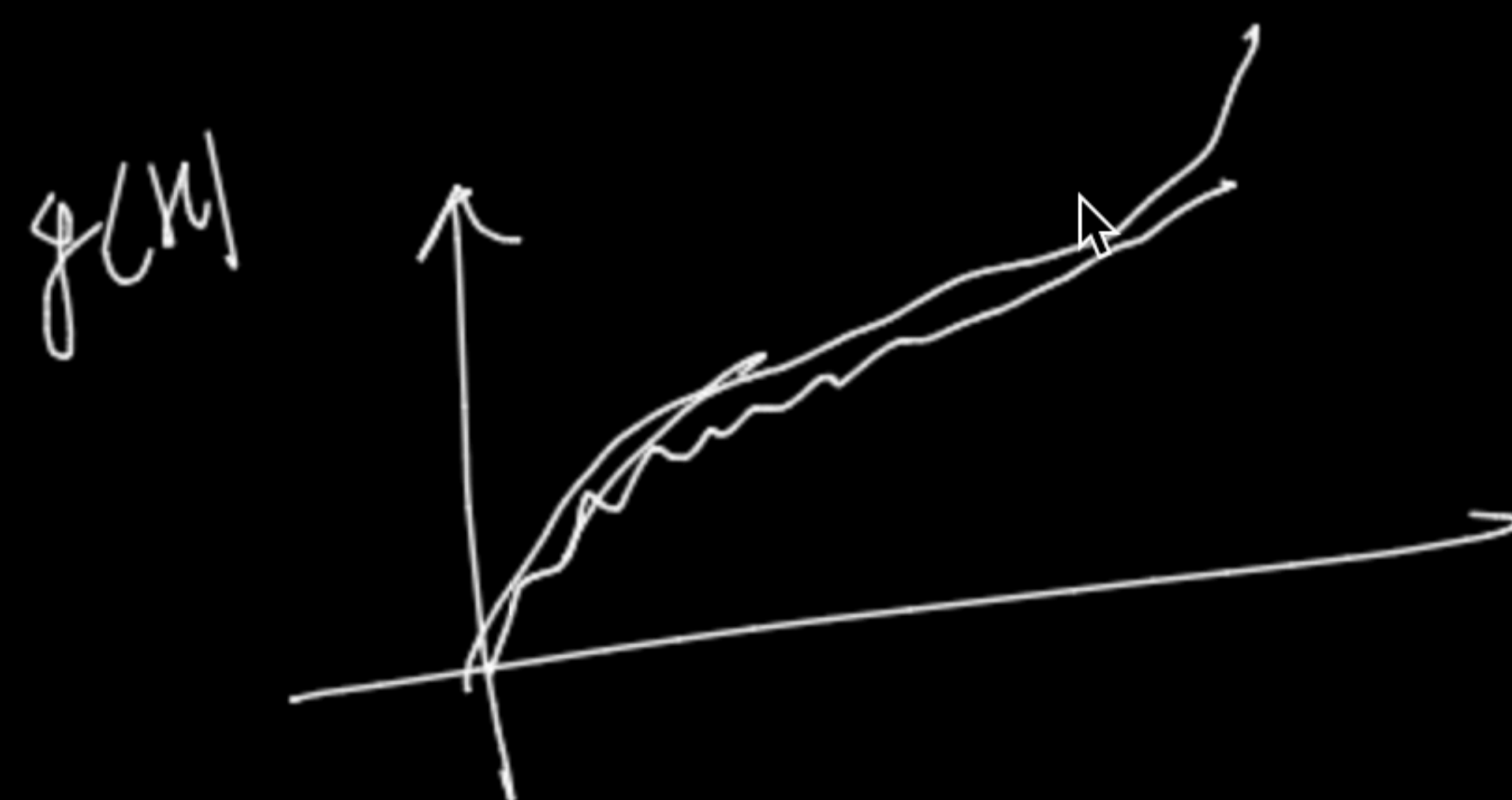
Monotonic fn  
↑

$g(x)$

$x \uparrow \quad g(x) \uparrow$

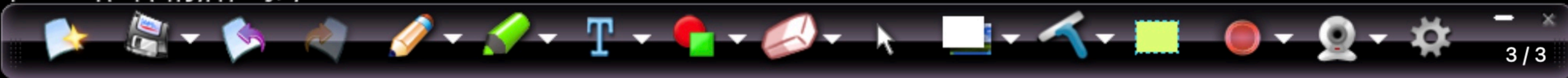
then  $x_1 > x_2$  then  $g(x_1) > g(x_2)$

↳ then  $g(x)$  is said to be  
monotonically increasing



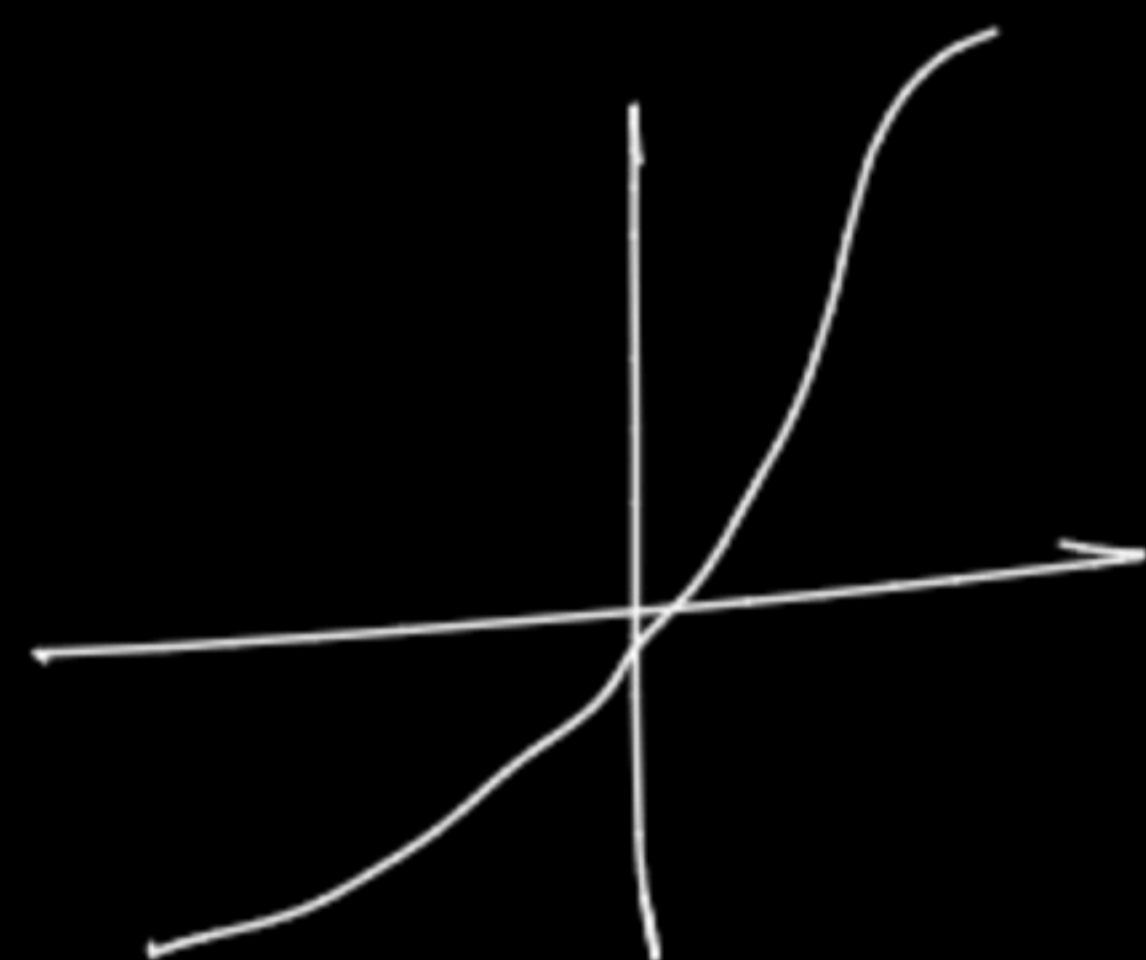
Optimal

problem



3/3

①  $x^*$   $\arg\min_x x^2$  is monotonically



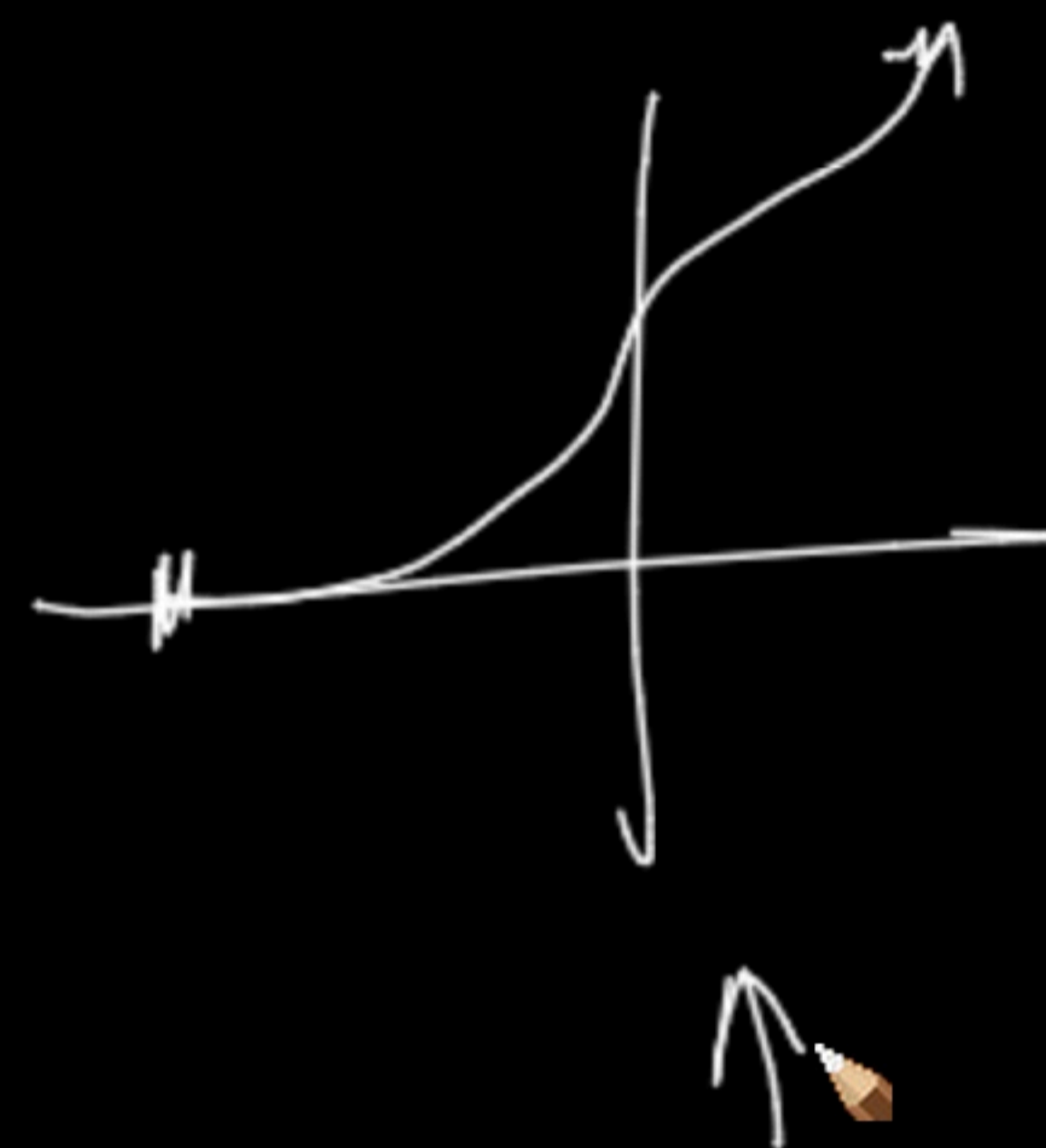
②  $g(x) = \log(x)$

$x^* = \arg\min_x f(x)$

$= \arg\min_x (g(f(x)))$

$= \arg\min_x (\log(x^2))$

$\therefore f(x) = x^2$



monotonically

$$w^* = \underset{w}{\operatorname{argmax}} \sum_{i=1}^n \frac{1}{1 + \exp(-y_i w^T x_i)}$$

with  $g(x) = \log(x)$

$$\log\left(\frac{w}{n}\right) \rightarrow \log w - \log n$$

$$w^* = \underset{w}{\operatorname{argmax}} \sum_{i=1}^n \log\left(\sigma(y_i w^T x_i)\right)$$

$$w^* = \underset{w}{\operatorname{argmax}} \sum_{i=1}^n \log\left(\frac{1}{1 + \exp(-y_i w^T x_i)}\right)$$

$$= \underset{w}{\operatorname{argmax}} \sum_{i=1}^n \log 1 - \log(1 + \exp(-y_i w^T x_i))$$

$$= \underset{w}{\operatorname{argmax}} \sum_{i=1}^n -\log(1 + \exp(-y_i w^T x_i))$$

ms  
arrow

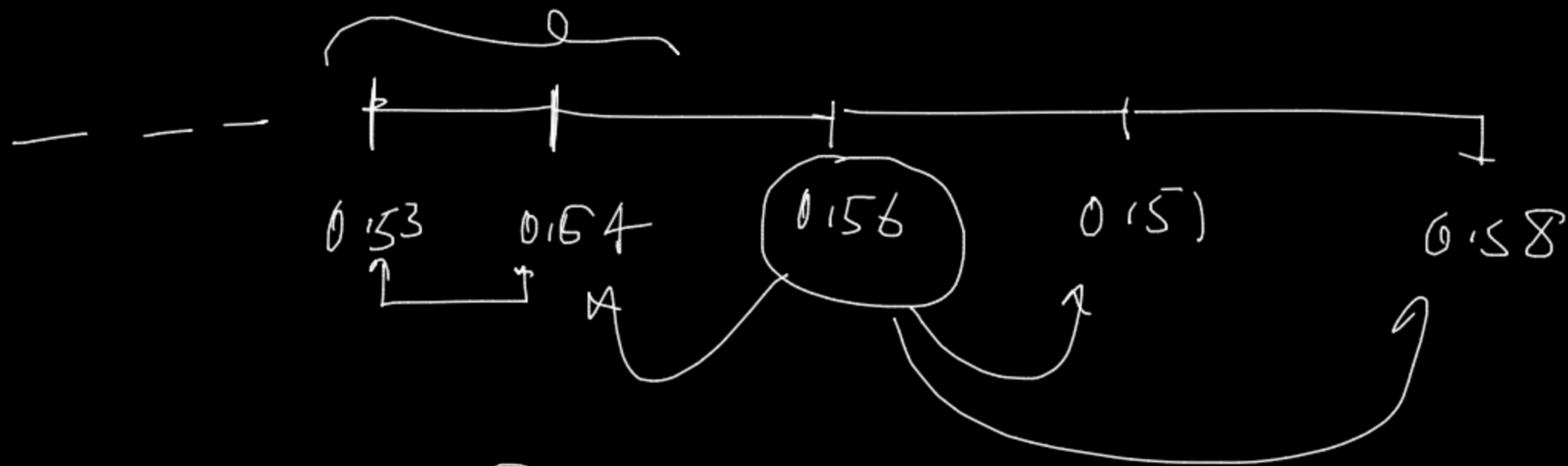
$$w^* = \underset{w}{\operatorname{argmax}} \sum_{i=1}^n -\log (1 + \exp(-y_i w^T x_i))$$

↑

$$\underset{w}{\operatorname{argmax}} -f(x) = \underset{w}{\operatorname{argmin}} f(x)$$

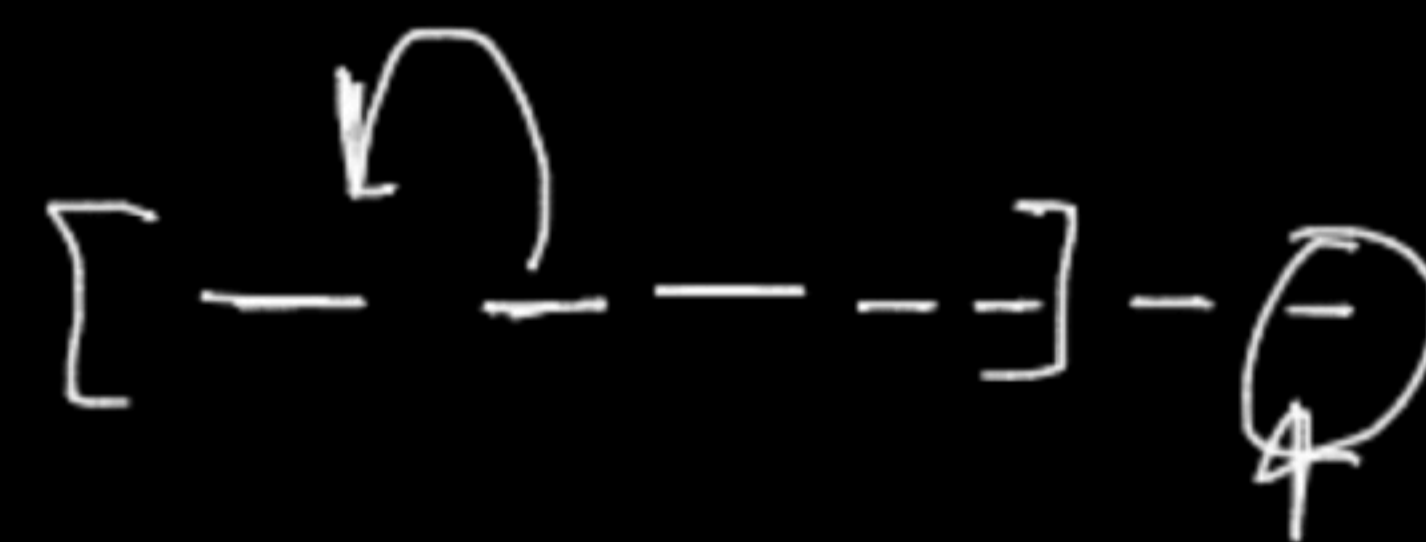
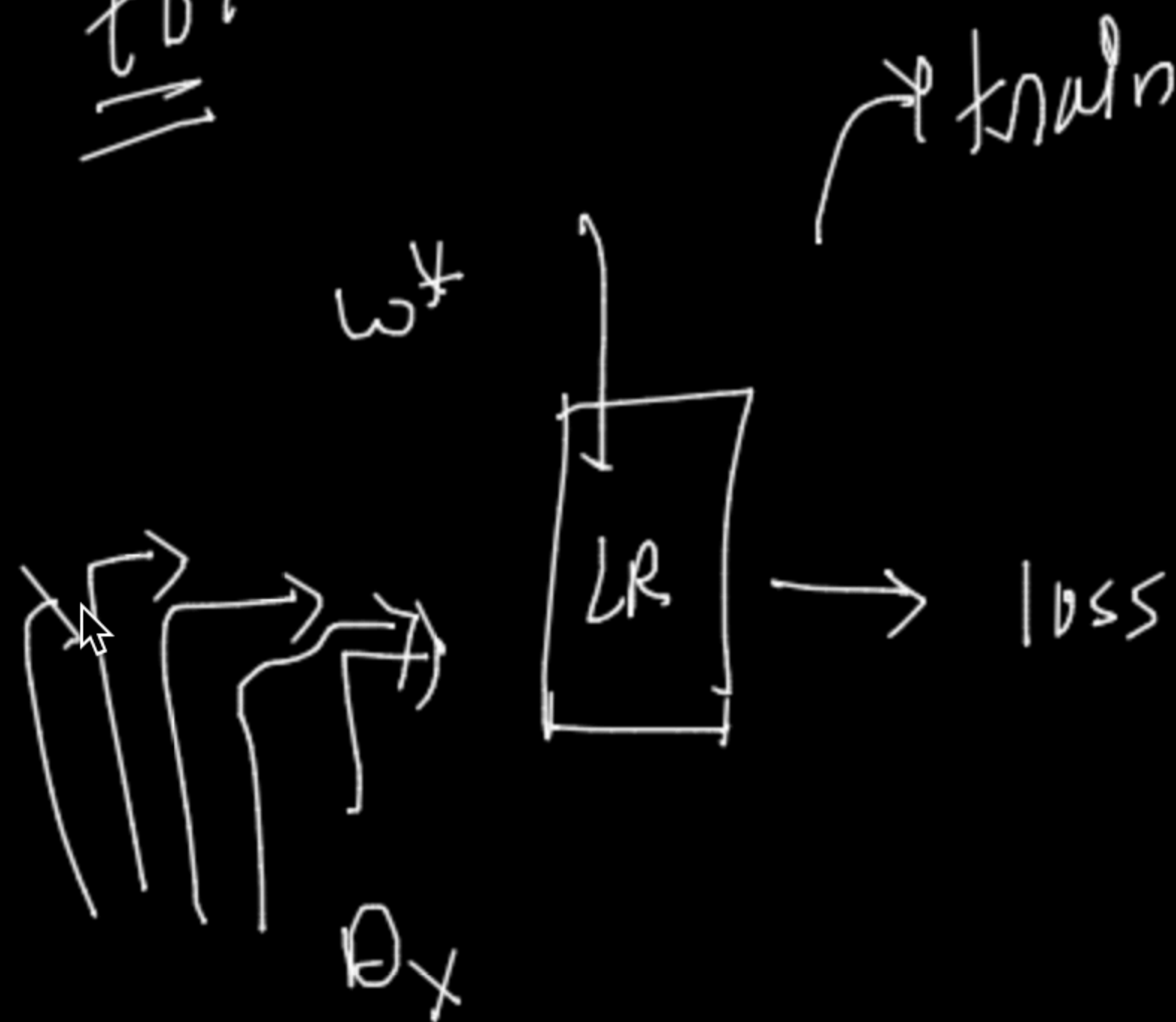
Optimisation Problem  
for Logistic Regression

$$w^* = \underset{w}{\operatorname{argmin}} \sum_{i=1}^n \log (1 + \exp(-y_i w^T x_i))$$



loss

tol



$w^*$   
 $\uparrow$   
 weight vector



$$\sigma(w^T x_i) = p(y_q = +ve)$$

(case -)

if  $w_i = +ve$

$$\hookrightarrow x_{q,i} \uparrow \Rightarrow (w_i x_{q,i}) \uparrow$$

$\downarrow$  implies

$$\sum_{i=1}^d w_i x_{q,i} \uparrow$$

$\downarrow$

$$\sigma(w^T x_{q,i}) \uparrow$$

$$w = \langle w_1, w_2, w_3, w_4, \dots, w_d \rangle$$

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

$$f_1, f_2, f_3, \dots, f_d$$

$$(x_q) \xrightarrow{?} y_q$$

[ ? ]

$$\text{if } w^T x_{q,i} > 0 \rightarrow +ve$$

$$\text{if } w^T x_{q,i} < 0 \rightarrow -ve$$

$$\Rightarrow P(Y_q = +w) \uparrow$$

(use  $i_1^0$ )

$$i_1 \quad \boxed{\omega_1 = -w}$$

$$x_{q,i} \uparrow \Rightarrow \omega_1 x_{q,i} \downarrow$$

$$\Rightarrow \sum_{i=1}^d (\omega_i x_{q,i}) \downarrow$$

$$\Rightarrow \sigma(\omega_1 x_{q,i}) \downarrow$$

$$P(Y_q = +w) \downarrow$$

$$P(Y_q = -w) \uparrow$$