

## Polynomial fitting

- Rewrite in matrix-vector form

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^M \\ 1 & x_2 & x_2^2 & \cdots & x_2^M \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^M \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_M \end{bmatrix}$$

- This can still be written as

$$\hat{Y} = Xw$$

- Loss  $J(w) = \frac{1}{N} \|Y - Xw\|^2$

- The  $i$ -th row of the design matrix  $X$  is simply a transformed feature  $\phi(x_i) = (1, x_i, x_i^2, \dots, x_i^M)$

$$i = 1, 2, 3, \dots, N$$



①  $x + 1 = 2 \Rightarrow x = 1$

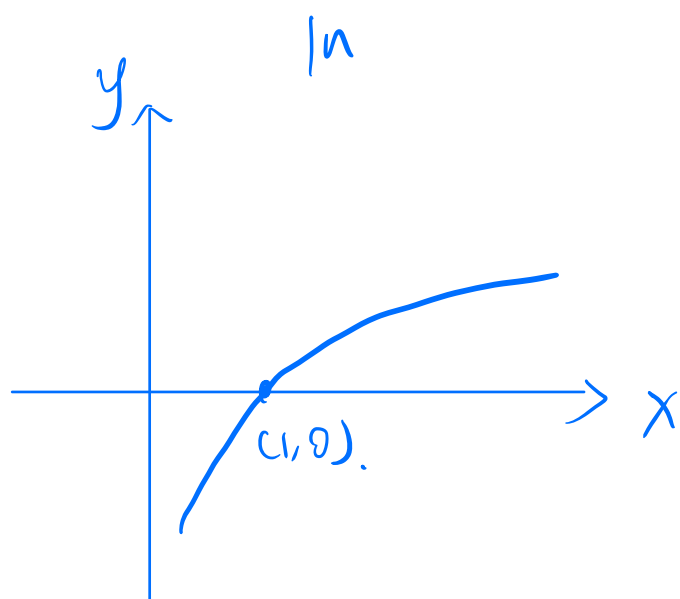
one function  $\rightarrow$  one element  
 $x$ .

②  $\begin{cases} x + y = 2 \\ x + 2y = 3 \end{cases} \Rightarrow \begin{cases} x = 1 \\ y = 1 \end{cases}$

two functions  $\rightarrow$  two  $\dots$   $\begin{matrix} x \\ y \end{matrix}$

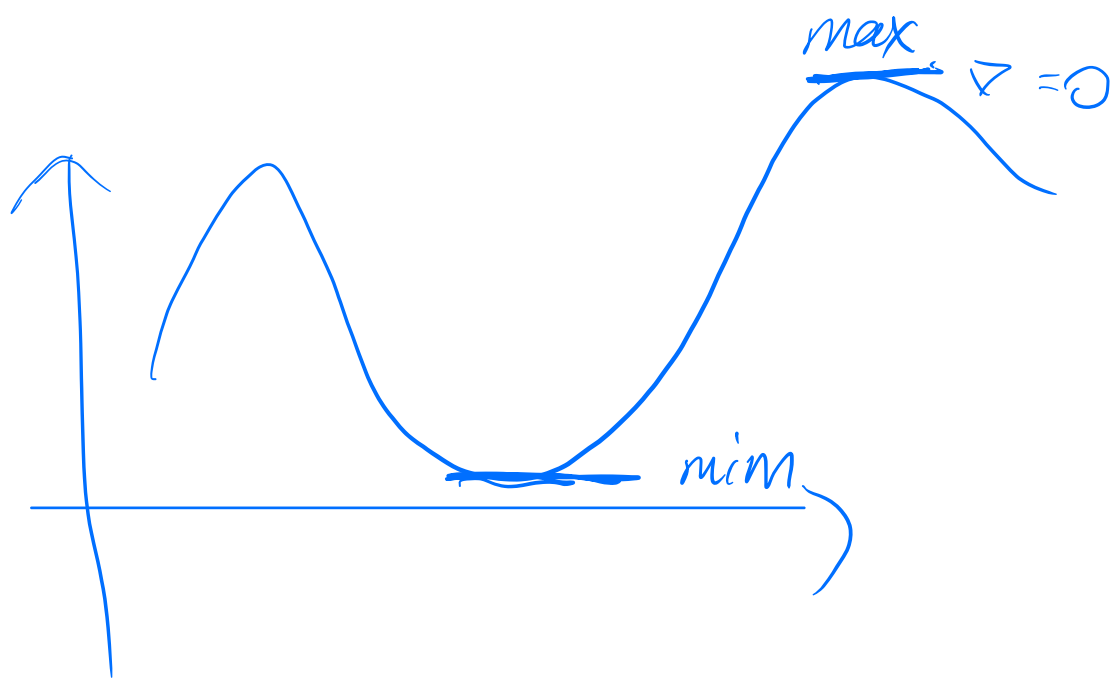
③  $N$  . function.

$N$  element

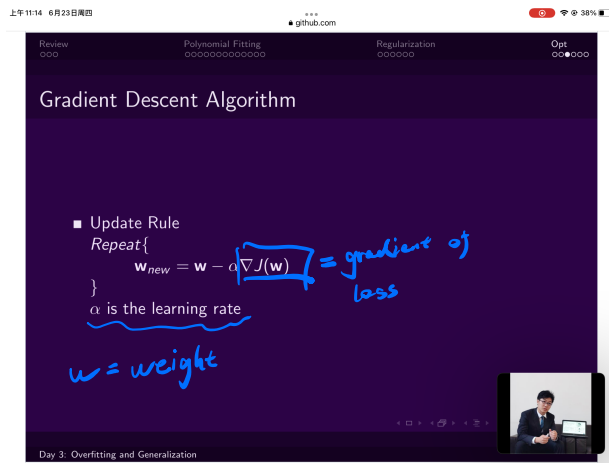


$$y = \ln(x).$$

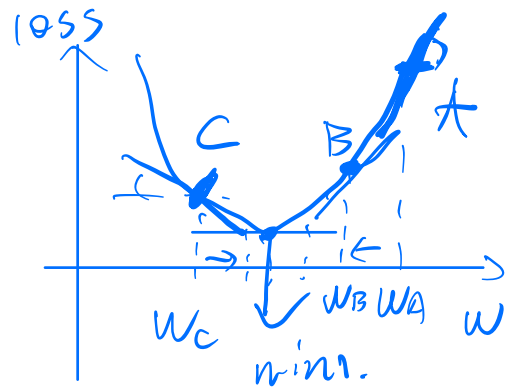
$\sim 0.9, 0.99$



$$w_{new} = w - \alpha \nabla J(w)$$



$$w_{new} = w - \alpha \cdot \nabla J(w)$$



$$A: \nabla J(w) > 0$$

$$\alpha \cdot \nabla J(w) > 0$$

$$w_{new} = w - \alpha \cdot \nabla J(w) \downarrow$$

$$C: \nabla J(w) < 0$$

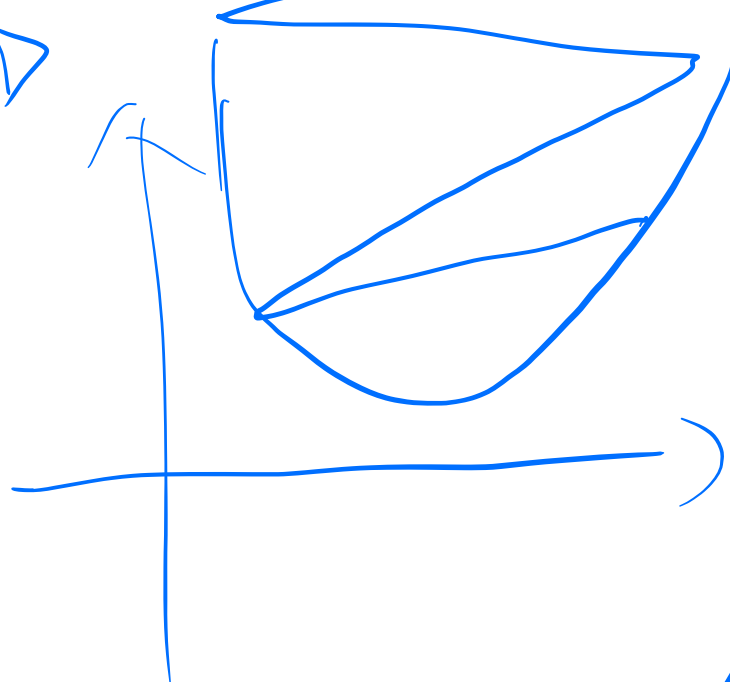
$$\alpha \cdot \nabla J(w) < 0$$

$$w_{new} = w - \alpha \nabla J(w) \uparrow$$

$$\alpha = 0.1, 0.01$$

Gradient Descent.  $\Rightarrow$  GD.

SGD



1 GB

$$y = 3x^2 + 2x + 1$$

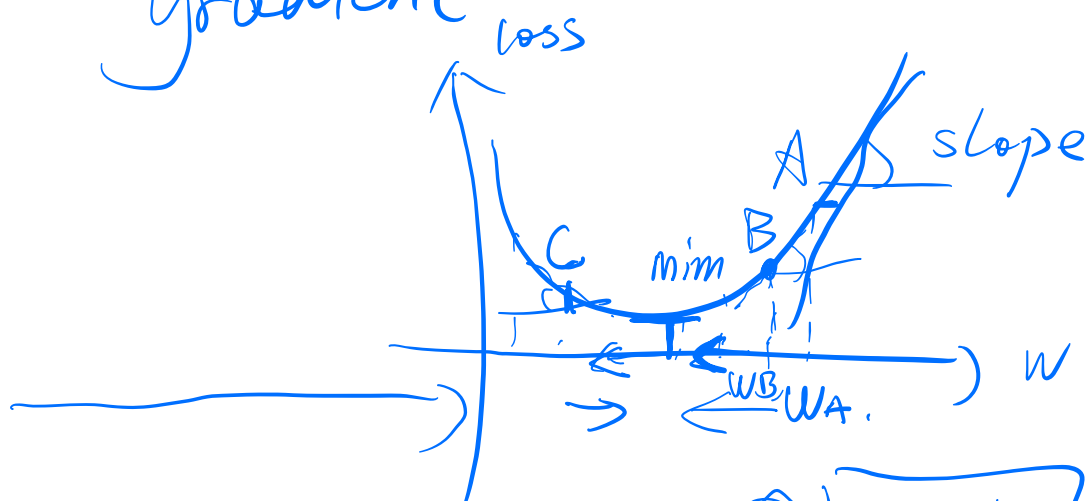
$$\nabla \rightarrow 0$$

+  $\infty$

$$\|W\|_2 =$$

$$w_1^2 + w_2^2 + \dots$$

gradient



$$W_{\text{new}} = W - \underbrace{(\alpha \nabla J(W))}_{> 0}$$

≡

> 0

3

2

$$V \Rightarrow 0$$

$$y = 3x^2 + 2x + 1$$

$$\nabla y = 6x + 2 \Rightarrow 0$$

$$x = -\frac{1}{3}$$

1 d'ime.

[1,

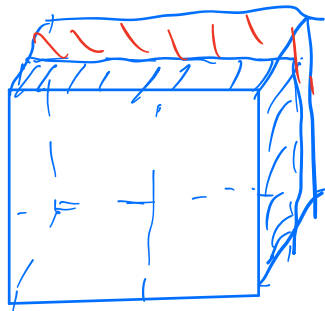
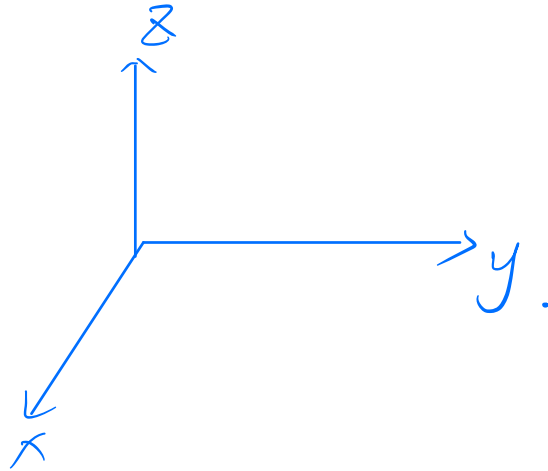
2,

3.]

2 d'ime.

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

3 dim



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

$$\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$