Gradients.

1> Vector , 2> Differentiation , 3> Partial derivatives.

4> Concept of maxima & minima

if you have $f(x, y)$

unit vector in y direction

Gradients $\nabla f(x,y) = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$

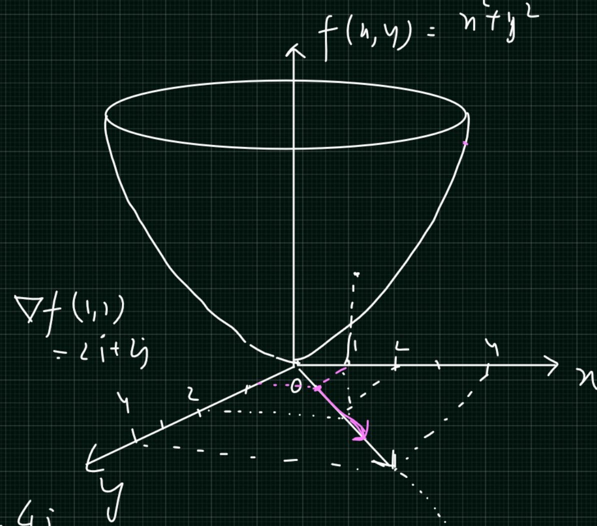
unit vector in x direction

$f(x,y) = x^2 + y^2$

$\nabla f(x,y) = \frac{\partial f}{\partial x} i + \frac{\partial f}{\partial y} j$

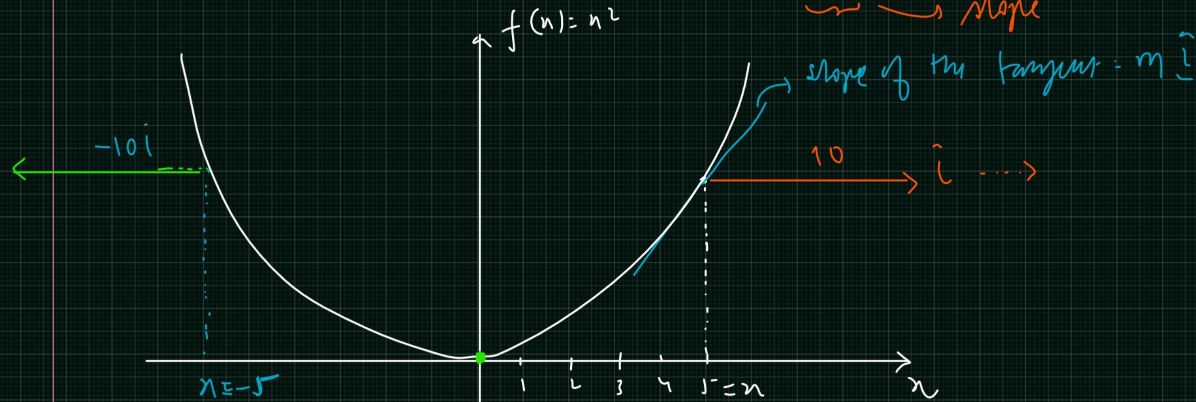
$\nabla f(x,y) = 2x\hat{i} + 2y\hat{j}$

$\nabla f(2,2) = 2\times 2\hat{i} + 2\times 2\hat{j} = 4i + 4j$

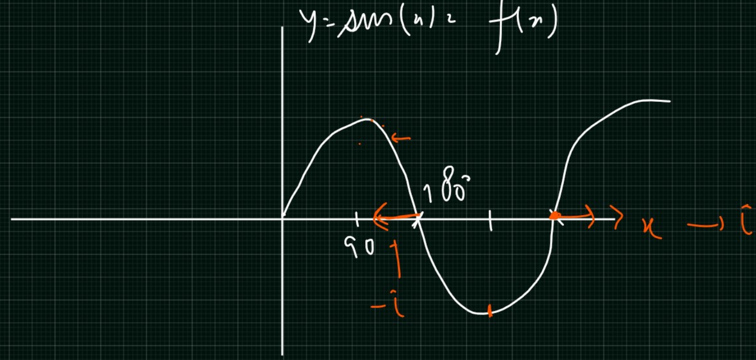


$f(x) = x^2$ $\nabla f(x) = \frac{\partial f}{\partial x}\hat{i} = 2x\hat{i}$ slope $\hat{i}$

$\longrightarrow$ slope



$\nabla f(5) = 10\hat{i}$



$$\nabla f(x) = \cos(x)\,\hat{i}$$

$$\cos(180) = -1$$

$$\nabla f(x = 180°) = -2\,\hat{i}$$

$$\nabla f(x = 360) = 1\,\hat{i}$$

$$\nabla f(x,y) = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j}$$

$$\nabla f(x,y,z) = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}$$

$f \to$ loss

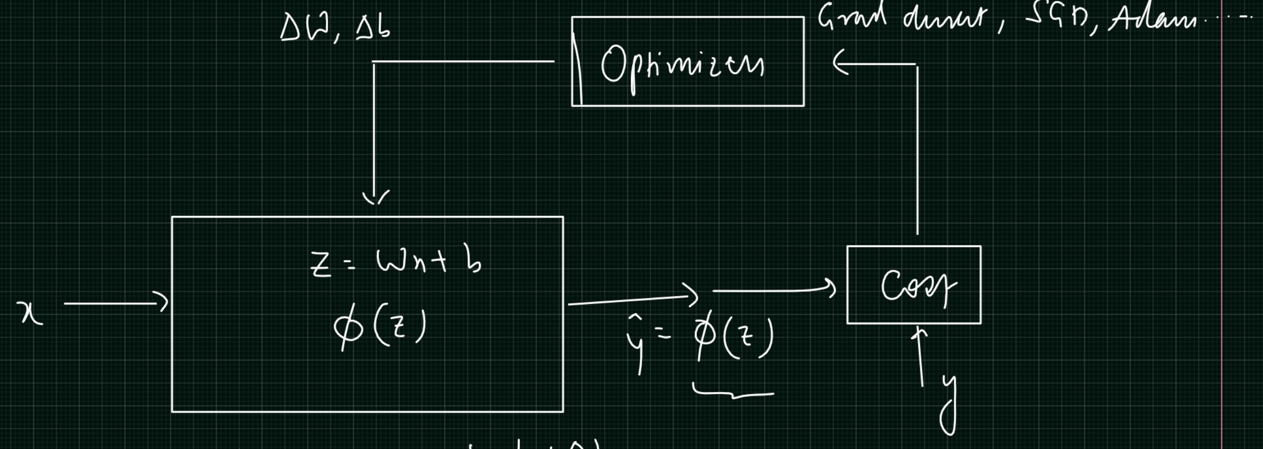
$f(w,b)$

$$\nabla f(w,b) = \frac{\partial f}{\partial w}\hat{i} + \frac{\partial f}{\partial b}\hat{j}$$

$$w = w - \eta\frac{\partial f}{\partial w} \qquad \Rightarrow \qquad \begin{bmatrix} w \\ b \end{bmatrix} = \begin{bmatrix} w \\ b \end{bmatrix} - \eta \underbrace{\begin{bmatrix} \frac{\partial f}{\partial w} \\ \frac{\partial f}{\partial b} \end{bmatrix}}_{\nabla f}$$

$$b = b - \eta\frac{\partial f}{\partial b}$$

$$\underbrace{\begin{bmatrix} w \\ b \end{bmatrix} = \begin{bmatrix} w \\ b \end{bmatrix} - \eta\nabla f}$$



$W = W + \Delta W$ $\qquad$ $b = b + \Delta b$

$\Delta W =$ $\qquad$ $\Delta b =$

$x_1, x_2, x_3 \ldots x_n \quad y \quad \hat{y}$

$$C = (y - \hat{y})^2$$

$$C = \left(y - \phi(z)\right)^2$$

$$C = \left(\underbrace{y - \phi\left(Wx + b\right)}\right)^2$$

$$C \longrightarrow f(W, b)$$

Assumptions :-

$$C = \left(y - \phi\left(Wx + b\right)\right)^2$$

i) treat y as constant $\quad k = 1$

ii) $b = 0$ $\qquad\qquad$ $\sigma(0.5 \times 2)$

iii) $\phi\left(Wx + \cancel{b}^{\,0}\right) \approx \underbrace{\phi(W \cdot x)}_{}^{x_1} \approx$ float

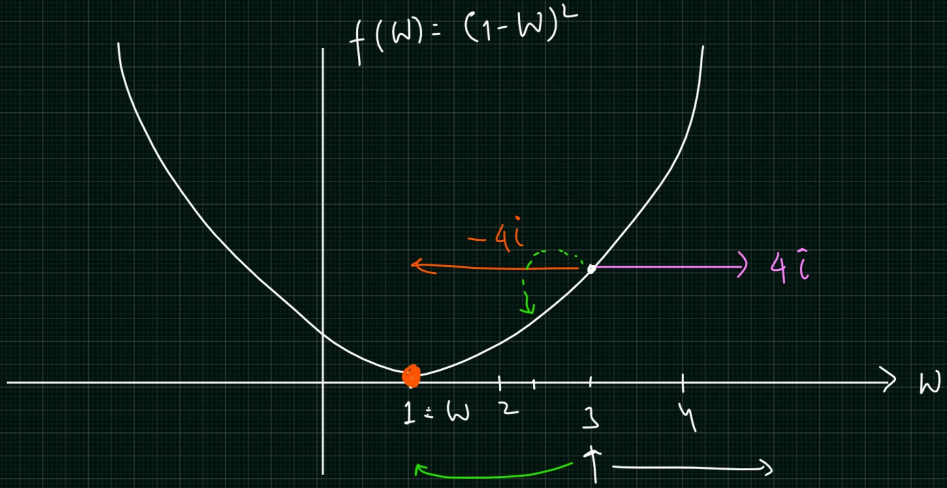
$\quad \hookrightarrow W \rightarrow$ float

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

$$C = \left(y - \phi\left(Wx + b\right)\right)^2$$

$$|||$$

$$\underbrace{C = (1 - W)^2} = f(W)$$



initilize random weight

Case 1. $W = 3$

$$c(3) = (1-3)^2 = 4$$

$$\nabla C(W=3) = \frac{\partial C}{\partial W} i = \frac{\partial (1-W)^2}{\partial W} i = \frac{\partial (1 + W^2 - 2W)}{\partial W} i$$

$$= (0 + 2W - 2) i$$

$$\nabla C(W) = (2W - 2) i$$

$$\nabla C(W=3) = (2 \times 3 - 2) i = (6-2) i = \underbrace{4i}$$

lets,

$$\Delta W = \frac{\partial C}{\partial W} \qquad W = W - \eta \frac{\partial C}{\partial W}$$

$$\Delta W = 4 \qquad W = W - \eta \nabla C$$

trial (i) = $W = W + \Delta W = 3 + 4 = \underbrace{7}$ ⇒ Gradient ascent

Target = 1

(ii) = $\Delta W = -\frac{\partial C}{\partial W}$

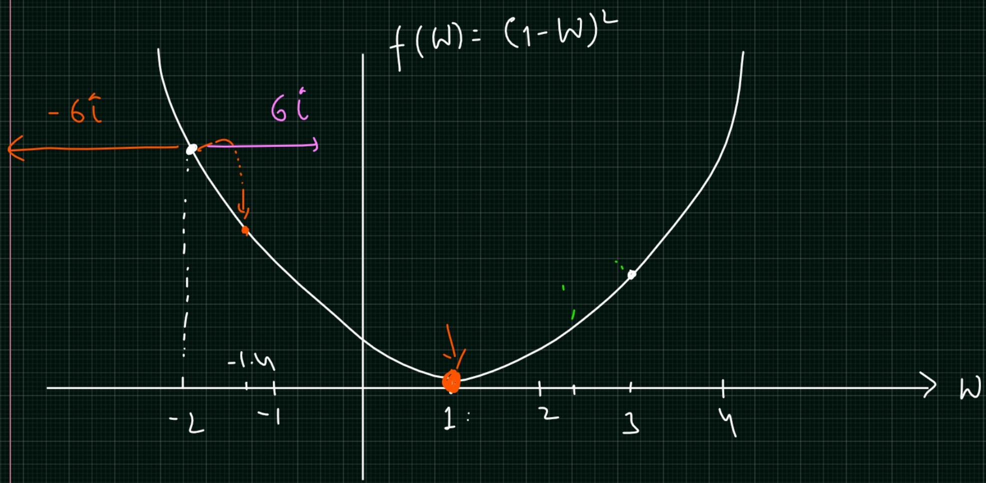
$$W = \underbrace{W + \Delta W} = 3 - 4 = -1 \Rightarrow$$ Gradient descent

Overshooting.

(iii) $\Delta W = -\eta \frac{\partial C}{\partial W} \Rightarrow \eta = 0.1$

$$W = W + \Delta W = 3 - 0.1 \times 4 = 3 - 0.4 = \underline{\underline{2.6}}$$

$$W = W - \eta \frac{\partial C}{\partial W}$$



Case 2

$W = -2$

$\nabla f(W) = (2W-2)\hat{i}$

$\nabla f(W=-2) =$

$= 2(-2) - 2\hat{i}$

$= -6\hat{i}$

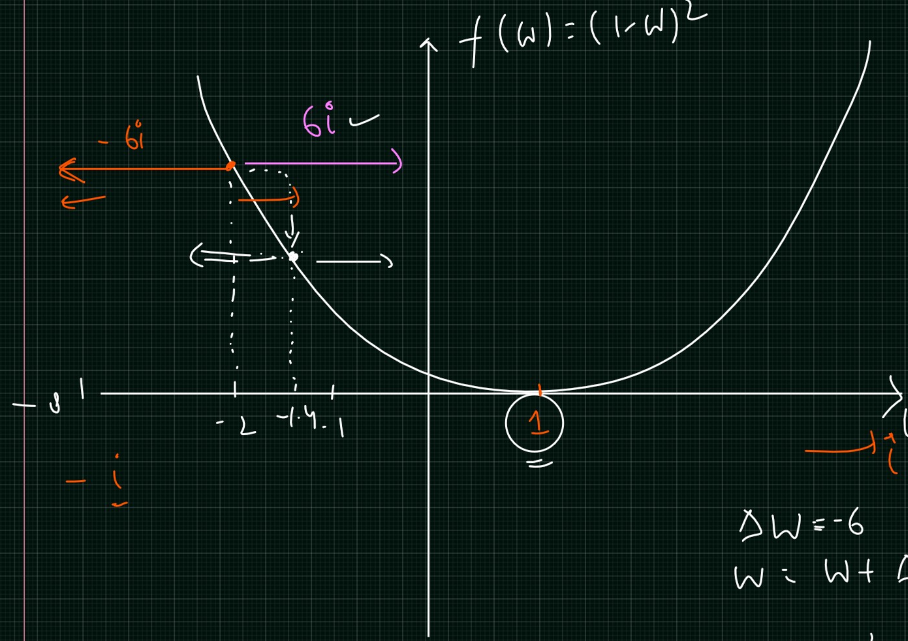
$$W = W \underbrace{- \eta \nabla f}$$

$$- 0.1 \cdot \underbrace{(-6)}$$

$$W = W + \underbrace{0.1 \times 6}$$

$$= -2 + 0.6$$

$$= \underline{-1.4}$$



$$\nabla f(w) = (2w-2)\,\hat{i}$$
$$= 2(-2)-2\,\hat{i}$$
$$= -6\,\hat{i}$$

$$\Delta W = -6$$
$$W = W + \Delta W$$
$$= -2 - 6$$
$$= -8$$

$$\Delta W = -(-6) = 6$$
$$\Delta W = 6$$
$$W = W + \Delta W = -2 + 6 = 4$$

$$\Delta W = -\eta \frac{\partial C}{\partial w}$$
$$= -0.1(-6)$$
$$\Delta W = 0.6$$

$$W = W + \Delta W$$
$$= -2 + 0.6 = -1.4$$