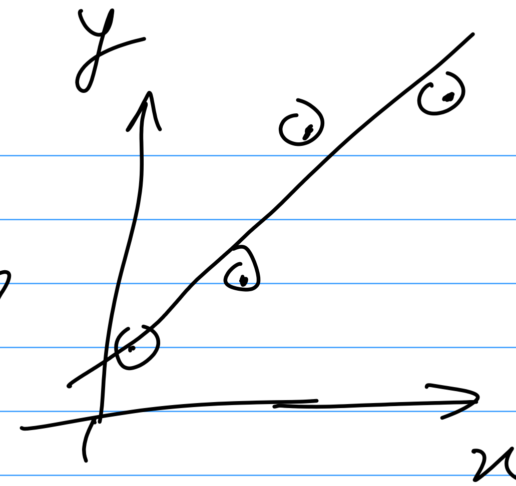
Linear Regression $y = \theta_0 + \theta_1 \cdot x$.

$x \to y$	
1.0	4.5
3.0	11.7
5.0	18.7
7.0 $\longrightarrow$	?

$$y = \underline{3.1} + \underline{1.8}x$$

$$y = f_\theta(x)\Big|_{x=7} = \ ?$$



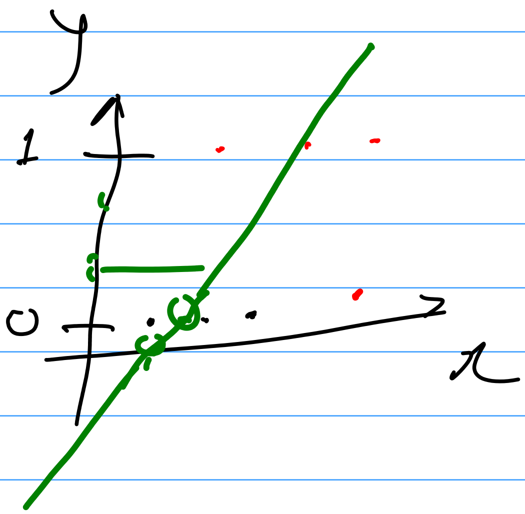
Classification $K = 2$.

$x \to y$	
1	0
3	0
5.0	1
7.0 $\to$	?

$$y = \theta_0 + \theta_1 x$$

$$y = 3.1 + 1.8x$$

$$y = f_\theta(x)_{x=7} = \ ? \ 0$$



$e = 2\ldots$

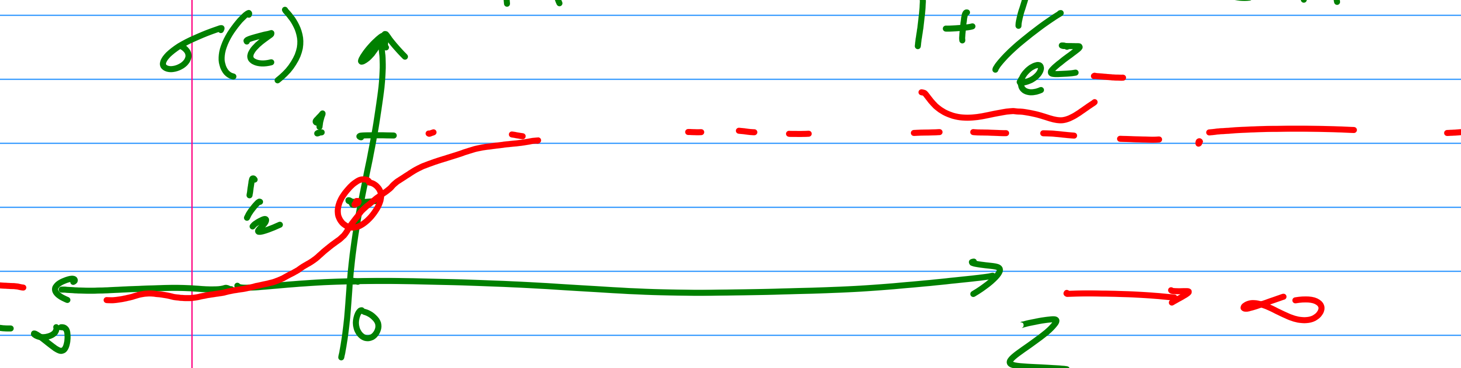
Logistic function

Specific Example: Sigmoid

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

Function Asymptote

$$= \frac{1}{1+\frac{1}{e^z}} = \frac{e^z}{e^z+1}$$



$z \to \infty,\ e^z \to \infty,\ \frac{1}{e^z} \to 0,\ \sigma(z) \to 1$

$z \to -\infty,\ 1/e^z \to \infty,\ \sigma(z) \to \frac{1}{1+\infty} \to 0$

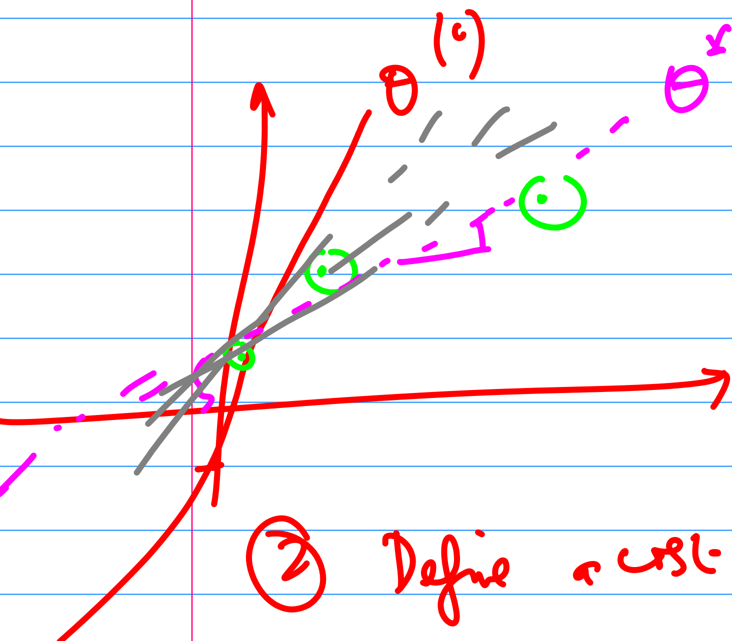
$\frac{d(\sigma(z))}{dz} =$

Gradient Descent

$y = \theta^T x = \theta_0 + \theta_1^T x$.

① Start with some random guess of $\theta^{(0)} = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$

$= \begin{bmatrix} -1.5 \\ 2 \end{bmatrix}$



② Define a cost function, MSE $\sum_{i=1}^{N} (y_i - \hat{y}_i)^2$

$\nabla_\theta L = \begin{bmatrix} \frac{\partial L}{\partial \theta_0} \\ \frac{\partial L}{\partial \theta_1} \end{bmatrix}$

$$\theta^{(i+1)} = \theta^{(i)} - \alpha \cdot \nabla_\theta L$$

hypothesis class

$$y = \sigma(\theta^T x) = \frac{1}{1+e^{(\theta_0 + \theta_1 x)}}$$

$x \to y \leftarrow \hat{y}$

$$\theta^{(0)} = \begin{bmatrix} \theta_0^{(0)} \\ \theta_1^{(0)} \end{bmatrix}$$

$\alpha = 0.01$

Cost function : $MSE((y-\hat{y})^2) \not\Leftrightarrow$ convex

Classification suitable cost function (log-likelihood, crossentropy)

$e^{?} = 0$

$e^{-\infty} = \frac{1}{e^{\infty}} = \frac{1}{\infty} = 0$

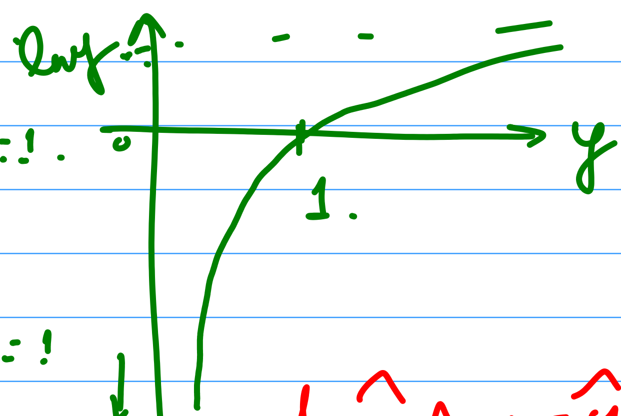
$$L(D, f_\theta) : I \to \mathbb{R}$$

$$L = -(y \log \hat{y} + (1-y)\log(1-\hat{y}))$$

a $L \to 0 \Rightarrow$ good

$L \to \infty \Rightarrow$ Bad

$0>$



$f(y) = \ln y$

$f(y)|_{y=1} = 0$

$f(y)|_{y=0} = -\infty$

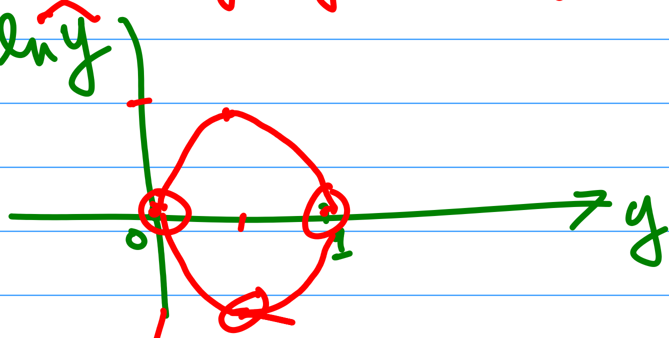
$\ln y = x$ $a^{?} = 1$.

$e^{(\ln y)} = e^{x}$

$y = e^{x}$ $e^{?} = 1$

$x = ?, y = 1.$ $\to 0$ min $-y \ln \hat{y}$ when $y = \hat{y}$

$(1-y)\ln(1-y)$. $-y \ln \hat{y}$



$y \in [0, 1]$

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

$f(y)|_{y=0} = 0$

$f(y)|_{y=1} =$