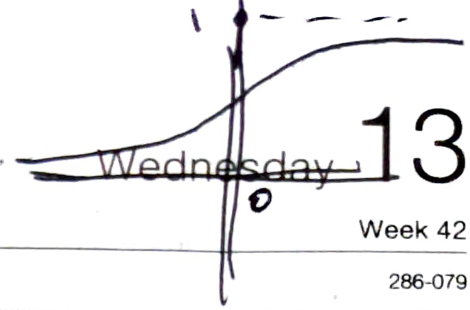

# Logistic regression:

where $g(z) = \dfrac{1}{1+e^{-z}}$

** sigmoid or logistic $f^n$ **

Want $h_\theta(x) \in [0,1]$

So we take ; $\boxed{h_\theta(x) = g(\theta^T x) = \dfrac{1}{1+e^{-\theta^T x}}}$

** $P(y=1 \mid x;\theta) = h_\theta(x)$

So; $P(y=0 \mid x;\theta) = 1 - h_\theta(x)$

and $y$ can take only two values $\{0,1\}$

So; $\boxed{P(y \mid x;\theta) = (h_\theta(x))^y (1-h_\theta(x))^{1-y}}$

$$\mathcal{L}(\theta) = P(\vec{y} \mid x;\theta) = \prod_{i=1}^{m} P(y^{(i)} \mid x^{(i)};\theta)$$

$$= \prod_{i=1}^{m} (h_\theta(x^{(i)}))^{y^{(i)}} (1-h_\theta(x^{(i)}))^{1-y^{(i)}}$$

Reason to take logistic function as $h_\theta(x)$



"log likelihood"

$$\ell(\theta) = \log \mathcal{L}(\theta)$$

$$= \sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1-y^{(i)}) \log (1-h_\theta(x^{(i)}))$$

$\mathcal{L}(\theta) \rightarrow$ Should be maximised $\Rightarrow$ $\ell(\theta) \rightarrow$ Should be maximised

To do that we shall use "Batch gradient ascent"

Can be used because $\ell(\theta)$ won't have local maximum; it will be concave and has single Maximum

i.e

$$\boxed{\theta_j := \theta_j + \alpha \frac{\partial}{\partial \theta_j} \ell(\theta)}$$

On substituting $\ell(\theta)$ + doing the derivatives we get

$$\boxed{\theta_j := \theta_j + \alpha \sum_{i=1}^{m} (y^{(i)} - h_\theta(x^{(i)})) \cdot x_j^{(i)}}$$

For logistic regression:

$$\theta_j := \theta_j - \alpha \sum_{i=1}^{m} \left( h_\theta(x^{(i)}) - y^{(i)} \right) \cdot x_j^{(i)}$$

where $h_\theta(x^{(i)}) = \dfrac{1}{1 + e^{-\vec{x}^{(i)}\vec{\theta}}}$