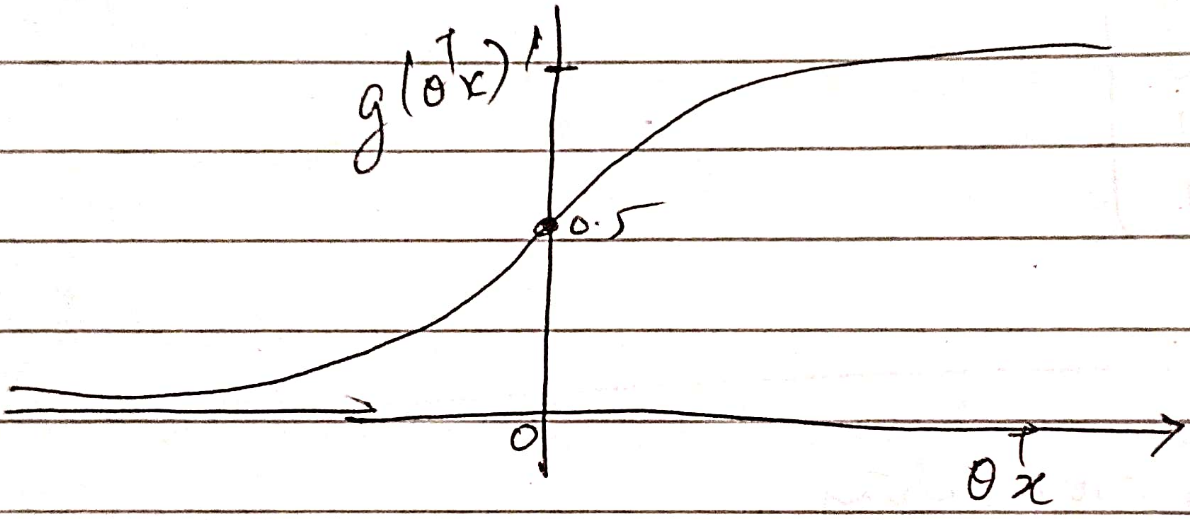
Logistic Regression

$$g(\theta^T x) = h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

$h_\theta(x) \geq 0.5 \quad y = 1$

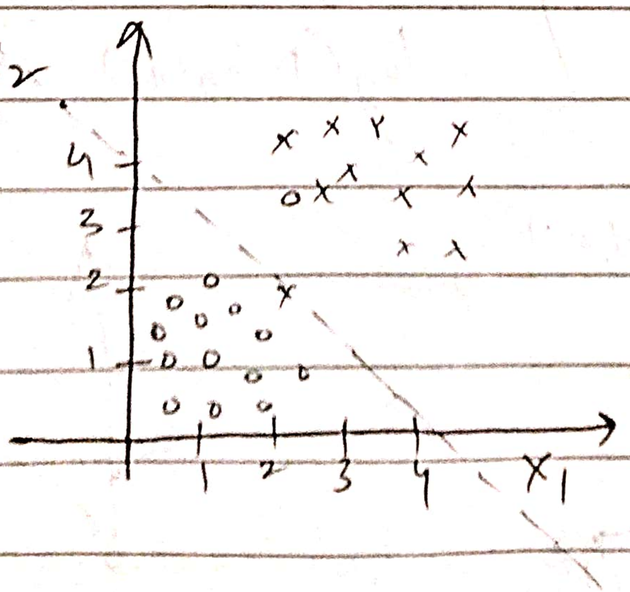
$h_\theta(x) < 0.5 \quad y = 0$



Decision boundary

$$h_\theta(x) = \frac{1}{1+e^{-\theta_0 + \theta_1 x_1 + \theta_2 x_2}}$$

X_1	X_2	Y
-	-	Δ
		Δ
-	-	0
		0
		0



$y = 1, \ h_\theta(x) \geq 0.5$

$\theta^T x \geq 0$

$y = 0, \ \theta^T x < 0$

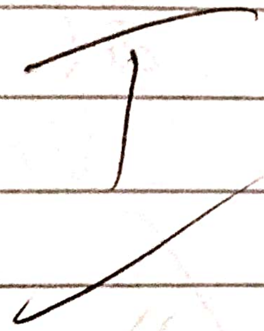
$$\theta = \begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix}$$

$-4 + x_1 + x_2 \geq 0 \qquad x_1 + x_2 \geq 4$

logistic regression

$P(Y=1|x)$

$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}} = \frac{1}{1+e^{-(\theta_0 x_0+\theta_1 x_1+\dots+\theta_n x_n)}}$$

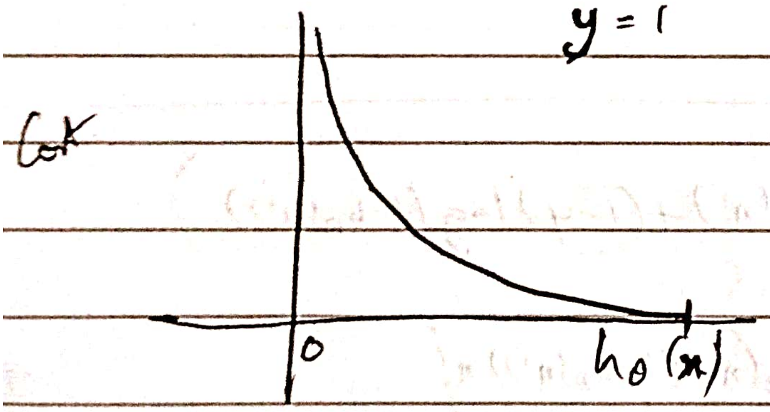


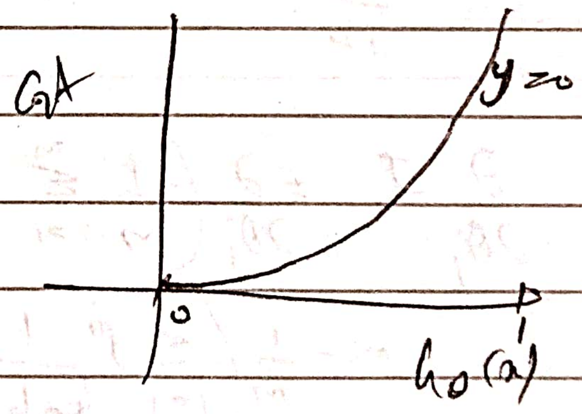
$h_\theta(x) \geq 0.5, \; y=1$

$h_\theta(x) < 0.5, \; y=0$

$$\text{error} = J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \text{Cost}(h_\theta(x^i), y^{(i)})$$

$$\text{Cost}(h_\theta(x), y) = \begin{cases} -\log h_\theta(x) & y=1 \\ -\log(1-h_\theta(x)) & y=0 \end{cases}$$





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

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

$$\theta_j = \theta_j - \alpha\frac{\partial J}{\partial \theta_j}$$

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

$$g(t) = \frac{1}{1+e^{-t}}$$

$$\frac{d}{dt}\left(\frac{1}{1+e^{-t}}\right) = \frac{(1+e^{-t})\cdot 0 - 1\cdot e^{-t}\cdot(-1)}{(1+e^{-t})^2} = \frac{e^{-t}}{(1+e^{-t})^2}$$

$$= \left(\frac{1}{1+e^{-t}}\right)\cdot\left(\frac{e^{-t}}{1+e^{-t}}\right) = g(t).(1-g(t))$$

$$\frac{d}{dt}\left(1-\frac{1}{1+e^{-t}}\right) = \frac{d}{dt}\left(\frac{e^{-t}}{1+e^{-t}}\right)$$

$$= \frac{(1+e^{-t})(e^{-t})(-1) - e^{-t}e^{-t}(-1)}{(1+e^{-t})^2}$$

$$= \frac{-e^{-t} - e^{-2t} + e^{-2t}}{(1+e^{-t})^2} = \frac{-e^{-t}}{(1+e^{-t})^2}$$

$$= -g(t)(1-g(t)).$$

$$\frac{\partial J}{\partial \theta_j} = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)}\frac{1}{h_\theta(x^{i})}\cdot\frac{\partial}{\partial\theta_j}(h_\theta(x^{i})) + (1-y^{(i)})\frac{1}{1-h_\theta(x^{i})}\cdot\frac{\partial(1-h_\theta(x^{i}))}{\partial\theta_j}$$

$$= -\frac{1}{m}\sum_{i=1}^{m} y^{(i)}\frac{1}{h_\theta(x^{i})}\cdot h_\theta(x^{i}).1-h_\theta(x^{i})\,x_j^{i} + (1-y^{(i)})\frac{1}{1-h_\theta(x^{i})}\,-h_\theta(x^{i})\,1-h_\theta(x^{i})\cdot x_j^{i}$$

$$= -\frac{1}{m}\sum_{i=1}^{m} y^{(i)}(1-h_\theta(x^{i}))\,x_j^{i} - (1-y^{(i)})\,h_\theta(x^{i})\cdot x_j^{i})$$

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

$$= -\frac{1}{m}\sum_{i=1}^{m}\left(y^{(i)} - h_\theta(x^{i})\right)x_j^{i} \quad \text{or} \quad \underline{\underline{\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{i}) - y^{(i)}\right)x_j^{i}}}$$