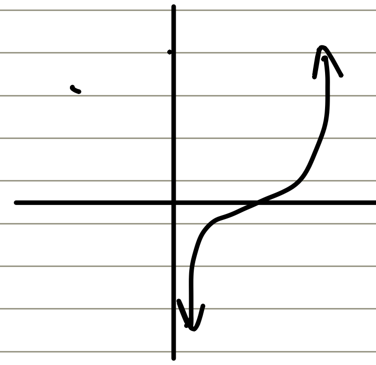
Logistic Regression:

Output is between $0 \to 1$ (binary classifier)

Logit Function: $f(p) = \log\left(\frac{p}{1-p}\right)$

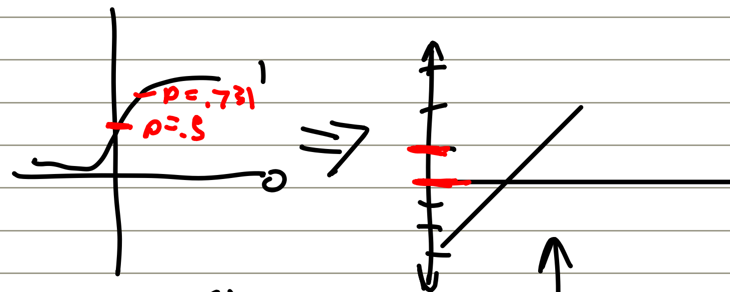
↳ log odds

↳ Takes in $(0,1)$ and makes it $(-\infty, \infty)$



$f(.5) = \log 1 = 0$

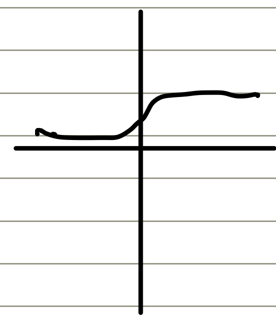
$f(.731) = \log(2.712) = 1$



Stretches out into line

Logit inverse = sigmoid $\sigma = \frac{1}{1+e^{-x}}$

↳ Takes $(-\infty, \infty)$ and makes it $(0,1)$



Linear Regression = $h(x) = \theta^T x + b$

⇓

Logistic $h(x) = \sigma(h(x)) = \frac{1}{1+e^{-\theta^T x}} \Rightarrow h(x) = \begin{cases} > 0.5, & \text{if } \theta^T x > 0 \\ < 0.5, & \text{if } \theta^T x < 0 \end{cases}$

Cost Function

if $y=1$: $\text{cost} = -\log(h(x))$ ⇒ penalize for close to 0

if $y=0$: $\text{cost} = -\log(1-h(x))$ ⇒ penalize for close to 1

$\text{cost}(h(x), y) = -y\log(h(x)) - (1-y)\log(1-h(x))$

Putting all errors in sample together:

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

Gradient Descent:

$$J' = \frac{1}{m}\sum_{i=1}^{m} (h(x^i) - y^i) \cdot x^i$$