Binary Classification:

In this our dependent variable (y) is a discrete variable where it takes two values i.e. $y \in \{0,1\}$ where 0 represents negative class and 1 represents a positive class.

<u>Examples:</u>

Email : Spam or not ?

Online Transactions : Fraud or not ?

Tumor : malignant or Benign ?

<u>**Threshold classifier output**</u> : We use linear regression and map all predictions greater than 0.5 as a 1 and all less than 0.5 as a 0. However, this method doesn't work well because classification is not actually a linear function.

What if we use linear reg for classification problem ?

If we use linear regression model then out predicted output might be <0 or sometimes it might be >1, both the cases are not suitable for classification problem.

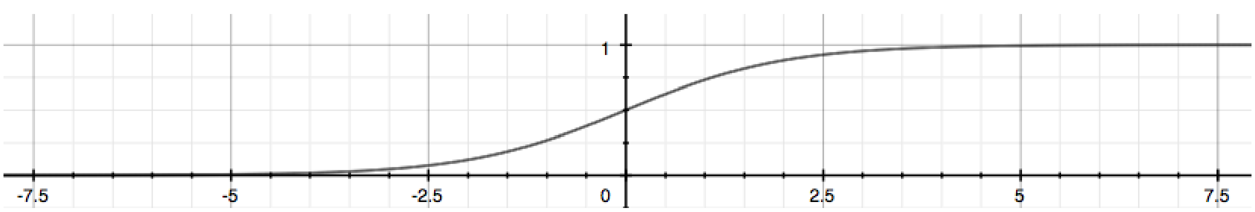
We can use classification algorithms such as **logistic Regression** for the binary classification.

<u>**Logistic Regression**</u> :

Our goal is to have our hypothesis function between 0 to 1, i.e. $0 <= h_\theta(x) <=1$.

θ

For Logistic regression we modify our hypothesis function as $h_\theta(x) = g(h_\theta(x))$, where g(z) is a logistic/sigmoid function, where $g(z) = \frac{1}{1+e^{-z}}$, the graph looks like below.



<u>Interpretation of Hypothesis Output</u>:

$h_\theta(x)$ = estimated probability that y=1 on input x.

Example of cancer prediction : $h_\theta(x)$ =0.7 → it tells that 70% chance of tumor being malignant tumor.

Another way of notation is $h_\theta(x) = P(y=1/x; \theta)$ → " *probability that y=1, given x, parameterized by* θ"

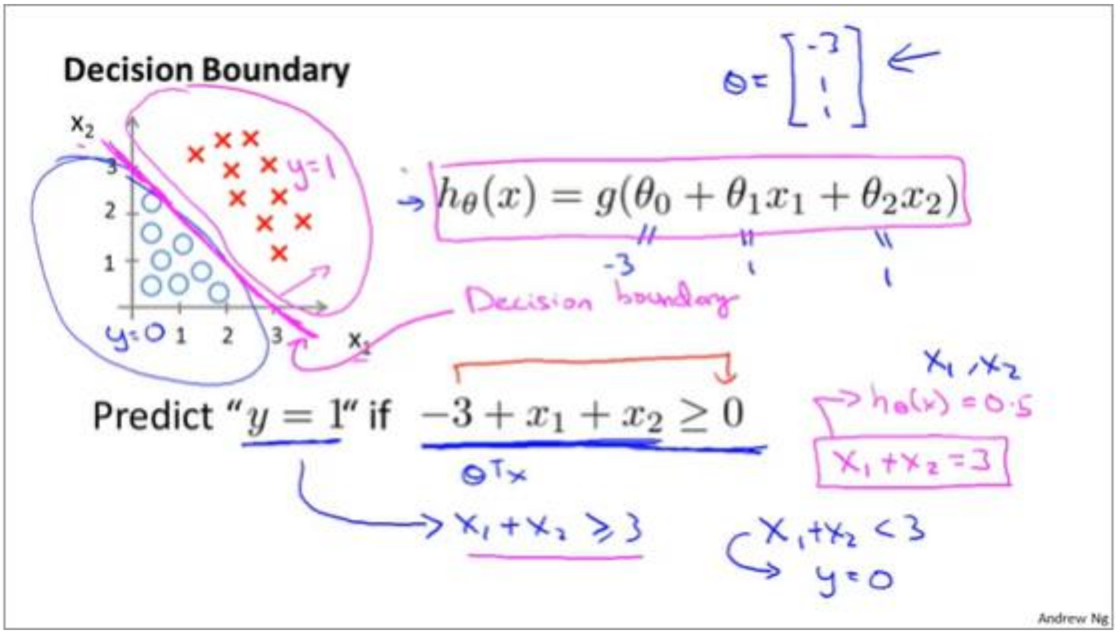
Also, $P(y=1/x; \theta) + P(y=0/x; \theta) = 1$.

Decision Boundary :

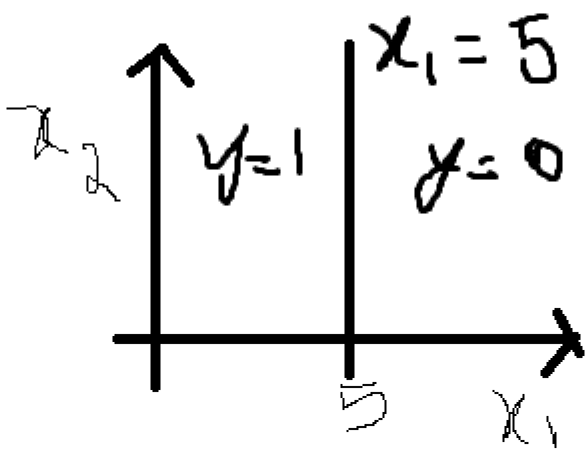
From sigmoid graph, we can say that g(z) >= 0.5 when z>=0, suppose if z=0 →g(z) =1/2

That is

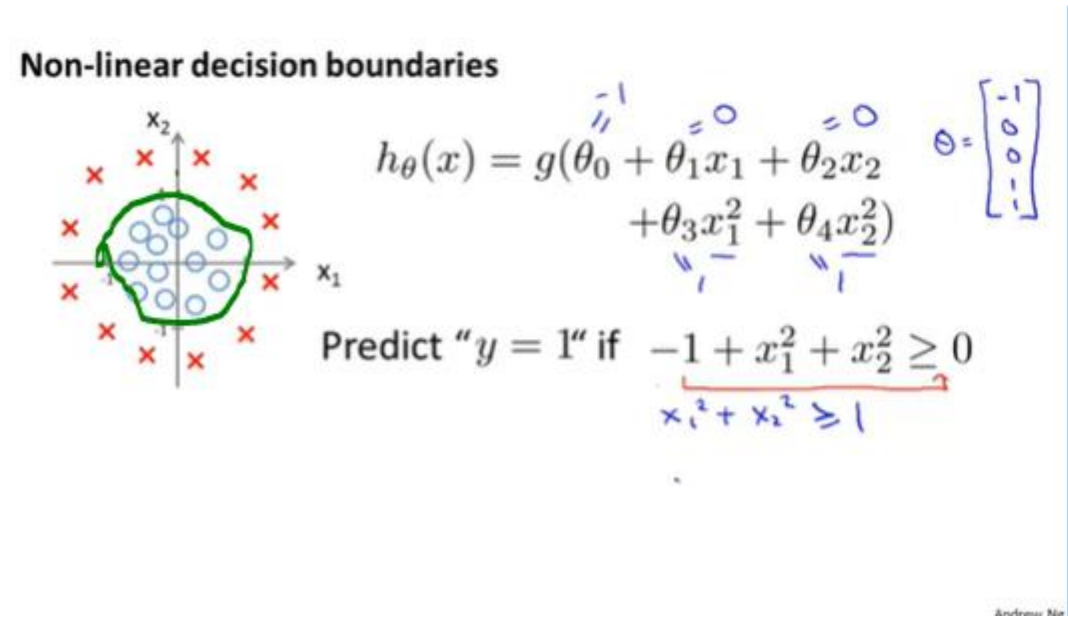
- for y=1, → $g(\theta^T X) >= 0.5$ i.e. $(\theta^T X) >= 0$. Similarly
- for y = 0 → $g(\theta^T X) < 0.5$ i.e. $(\theta^T X) < 0$.



- Consider logistic regression with two features $x1$ and $x2$. Suppose theta_0 = 5 , theta_1 = -1 and theta_2=0, so that $h\vartheta(x)=g(5-x1)$. Which of these shows the decision boundary of $h\vartheta(x)$?
- For y=1, $\theta^T X$ i.e. 5-x1 >= 0 → x1 <= 5 is the required decision Boundary equation.



- The decision boundaries might not always be linear, it can also be non-linear as well.



<u>Cost Function for Logistic Regression</u>: