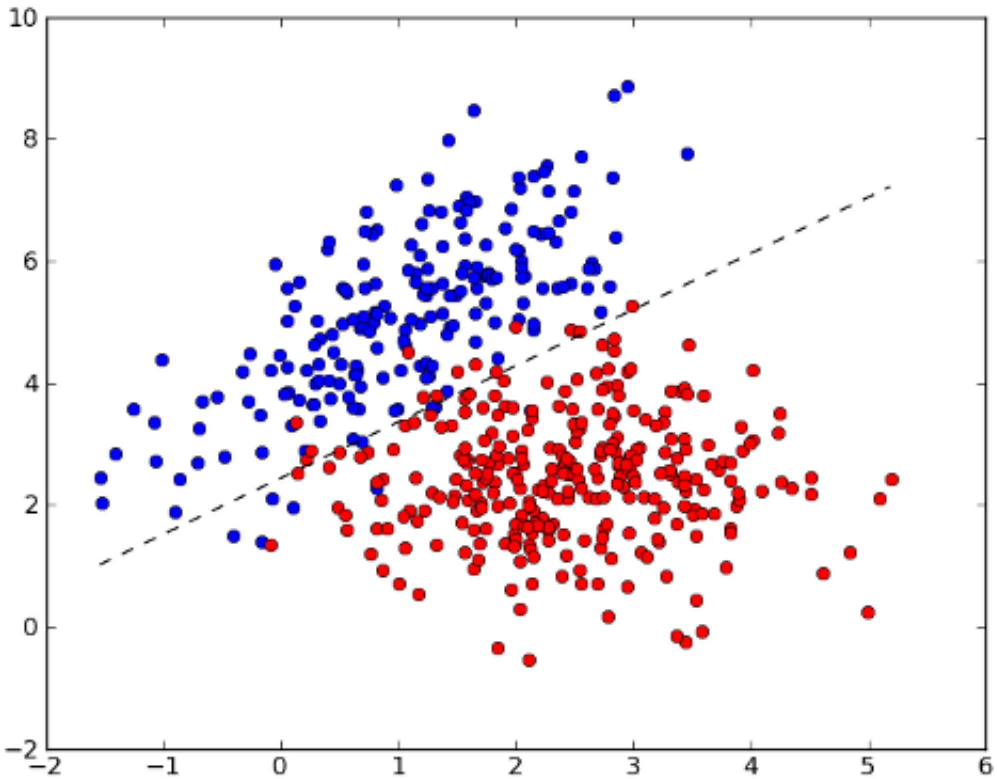
In [ ]:

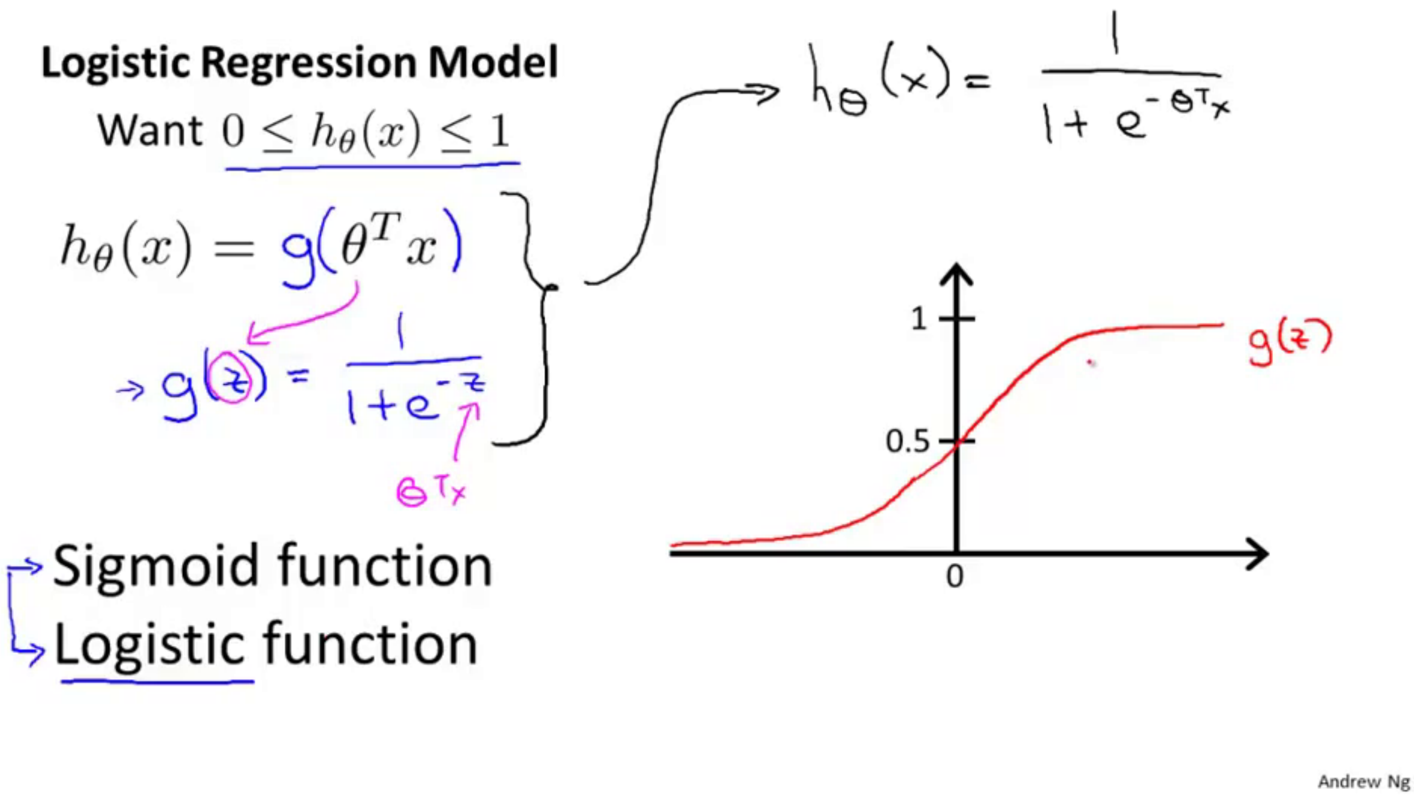
Logistic Regression



In logistic regression, our goal is to find a hyperplane, or a decision boundary, that separates the data. The hyperplane is always 1 dimension less than the dimension of the data. For example, if you have a 2 dimensional data, the hyperplane is 1 dimensional, or a line. And, if you have a 3 dimensional data, the hyperplane is a 2 dimensional, or a plane. Anyways, we'll focus on the 2 dimensional data case.

The decision boundary determines whether or not a sample is in one class or the other. In the image above, the decision boundary determines whether a sample is blue, $P(blue|sample)$, or red $P(red|sample)$.

Like in linear classification, the decision boundary in logistic regression is a line, which as we saw earlier, can be represented by $X\theta$. When a sample is on the decision line $X\theta$, we say that $P(blue|sample) = P(red|sample) = .5$. Normally, however, since the range of $X\theta$ is $[-\inf, \inf]$, we need to convert this to a probability and the way to do that is by using a sigmoid function $P(blue|sample) = \frac{1}{1+e^{-\theta^T x}}$.



Let's rewrite this another way:

$$P(blue|sample) = \frac{1}{1 + e^{-\theta^T x}}$$
$$1 + e^{-\theta^T x} = \frac{1}{P(blue|sample)}$$
$$e^{-\theta^T x} = \frac{1}{P(blue|sample)} - 1$$
$$e^{-\theta^T x} = \frac{1}{P(blue|sample)} - \frac{P(blue|sample)}{P(blue|sample)}$$
$$e^{-\theta^T x} = \frac{1 - P(blue|sample)}{P(blue|sample)}$$
$$log(e^{-\theta^T x}) = log(\frac{1 - P(blue|sample)}{P(blue|sample)})$$
$$-\theta^T x = log(\frac{1 - P(blue|sample)}{P(blue|sample)})$$
$$\theta^T x = log(\frac{P(blue|sample)}{1 - P(blue|sample)})$$

```
In [ ]:
```