逻辑回归：

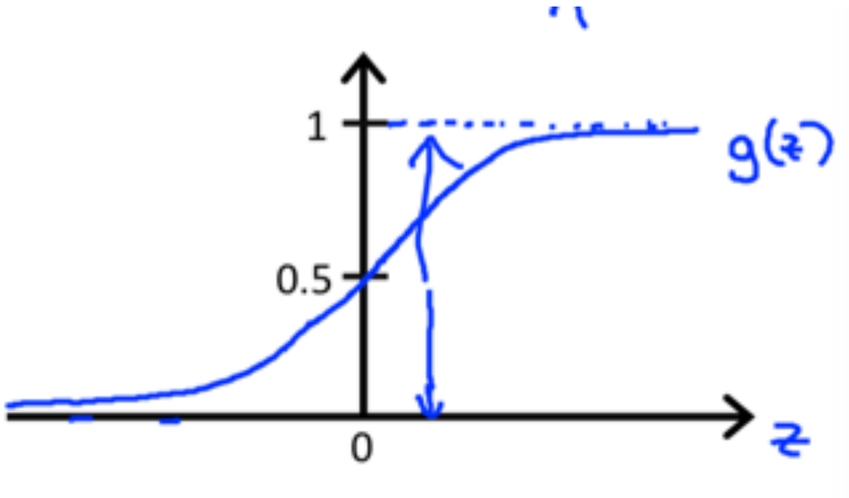
逻辑回归所解决的问题：二元的分类问题。例如：一封邮件是不是垃圾邮件？一张图片中的肿瘤是良性的还是恶性的？

结果的表示方法：是(1) / 否(0)

判断结果的依据：$h_\theta(x)$ >= 0.5

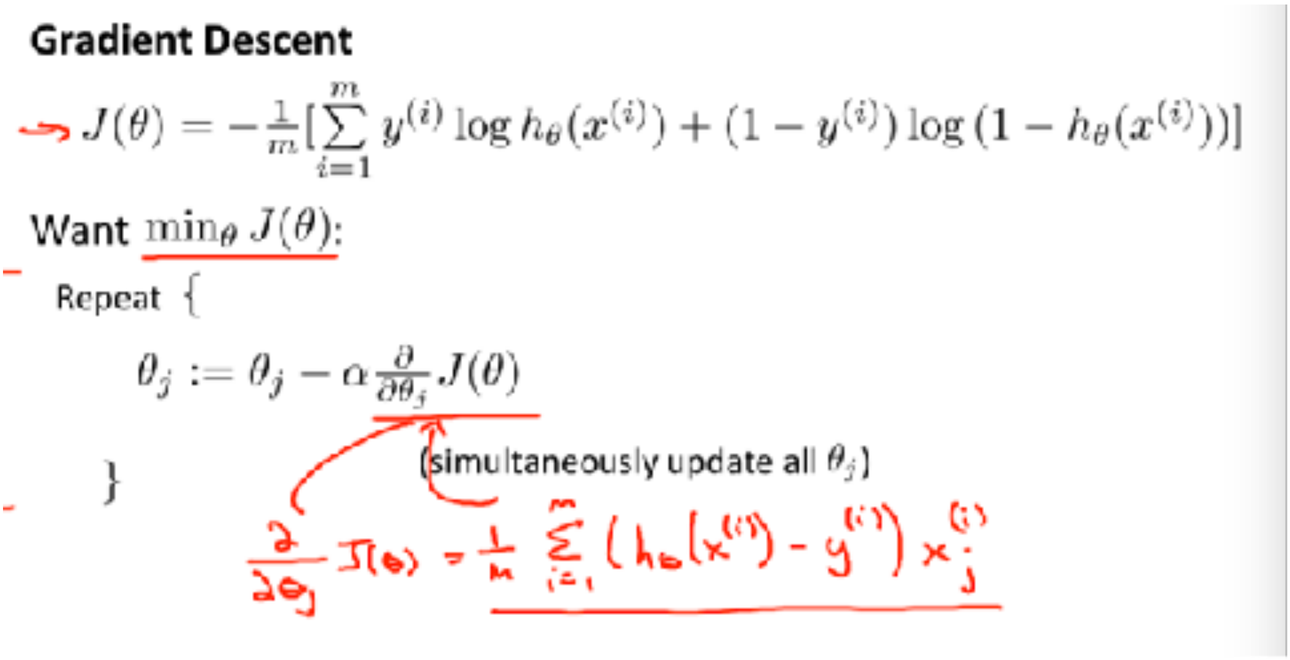
判断的函数：Sigmoid function. $h_\theta(x) = \dfrac{1}{1+e^{-\theta^T x}}$

该函数的图像如下：



其中z等于$\theta^T x$，可以看出利用z等于0作为判断的分界线，对于一组参数 θ 给定一个x，z是这个x的y的预期，对于直线上方的点，其y会大于0；而对于直线下方的点，其y会小于0；据此对整个数据空间一分为二。（实际上强行将问题分为直线上和直线下）

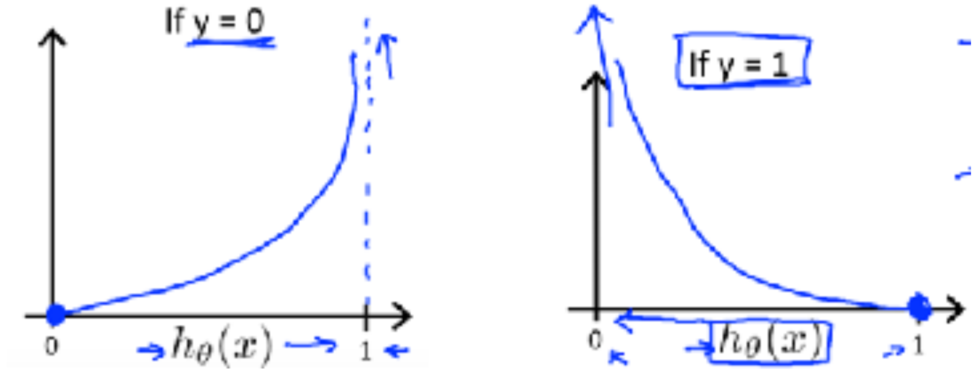
损失函数和参数的更新：



损失函数的设计颇为巧妙，按照分类的结果其分为两个部分：

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

当y为0的时候，预测的结果越靠近1受到的惩罚越大；当y为1的时候，预测结果越靠近0受到的惩罚越大。



优化器的选择：

除了梯度下降法外，还有一些其他的优化器可供选择：

Optimization algorithm

Given θ, we have code that can compute

- $J(\theta)$
- $\frac{\partial}{\partial\theta_j}J(\theta)$ (for $j=0,1,\ldots,n$)

Optimization algorithms:
- Gradient descent
- Conjugate gradient
- BFGS
- L-BFGS

Advantages:
- No need to manually pick α
- Often faster than gradient descent.

Disadvantages:
- More complex

衍生出来的问题：

多元分类问题，使用逻辑回归为 One-vs-all 方法，在对第i组分类的时候，其他组都是异类，反复利用这一方法可以画出k条线，最终将图形两两区分开。