01 线性回归和逻辑回归

by Bosco Xue 2020.04.21

线性回归和逻辑回归都叫回归，但是前者用于回归问题，后者用于分类问题。

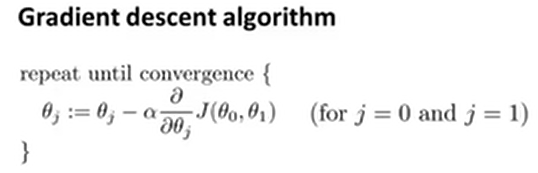
1. **线性回归**

（1）线性回归的Cost Function通常采用Square Error：



（2）线性回归的优化算法有梯度下降法和正规方程法：

a. 梯度下降法：沿梯度的反方向同步更新变量。



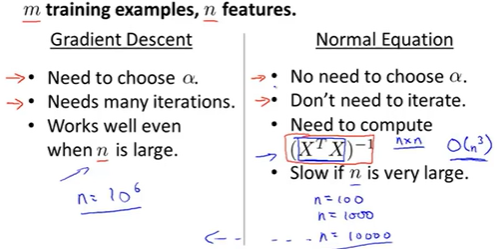
采用梯度下降法时若不同特征的scale差距较大，则可以采用特征缩放（归一化、标准化等）以加速收敛；学习率的选择可画出J-iteration times图像辅助判断，3倍递增。

b. 正规方程法：一种解析法，可由线代投影法或者矩阵求导得到。



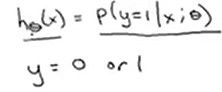
对于正规方程法，在特征冗余和m>>n的时候，可能矩阵求逆不存在。此外，由于解析法，不需要迭代，故也无需进行特征缩放。

c. 两者比较：

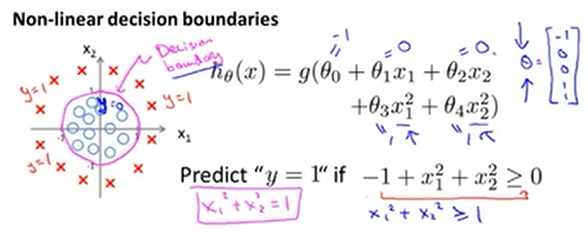


1. **逻辑回归**

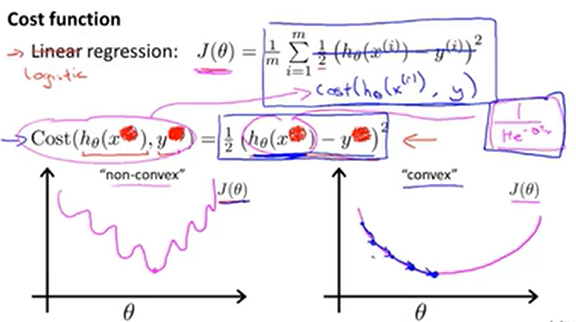
（1）模型假设：x的输出h(x)为某一分类的概率。



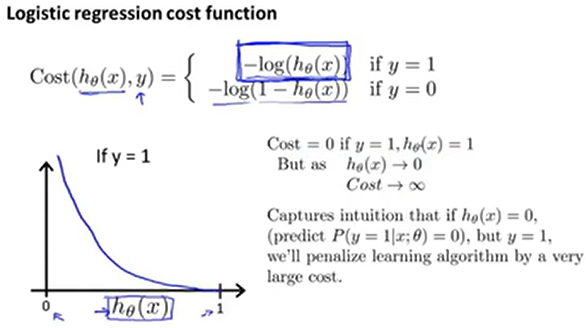
（2）为什么叫回归：因为LR实际上就是在线性回归模型外套上一个sigmoid函数，从而将函数值映射到0-1上，即要求的核心是hθ(x)，要拟合的dicision boundary（决策边界）即*hθ(x)*就是一个线性方程。



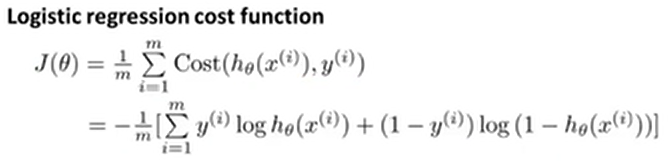
（3）为什么代价函数不用线性回归中的平方误差：因为此时*hθ(x)*代入的话，得到的代价函数为一非凸函数。



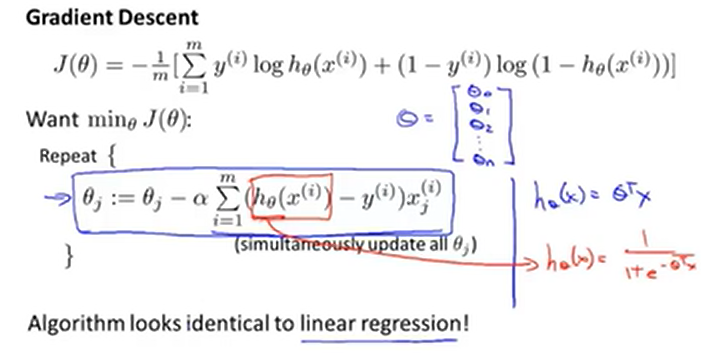
（4）如何理解逻辑回归的代价函数要取对数：按照代价函数设计的思想，在y=1时，hθ(x)尽量大，在y=0时，hθ(x)尽量小，即代价函数取为y=1时，-hθ(x)，y=0时，-(1-hθ(x))，似乎也是可以的，但是取对数后，在y=1时，-log(hθ(x))在hθ(x)的值很小时很敏感，惩罚大，代价大，在y=0时同理。



另一种理解是，根据样本可得到极大似然函数，两边取对数后再取负值以求最小值作为代价函数。



（5）采用梯度下降法优化的话，得到的式子和线性回归是一样的。



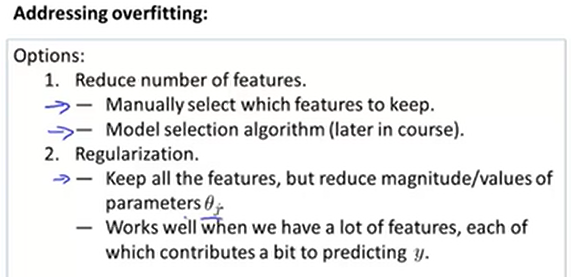
（6）通过one vs rest 可以将LR算法应用于多分类问题，此时有几个种类就要训练出几个LR分类器，最后取预测值最大的作为预测结果。

1. **正则化Regularization**

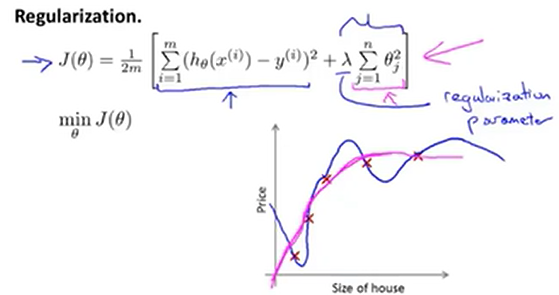
（1）overfitting

课上给出的过拟合处理方法有：

* 减少特征数量
* 正则化：去掉某些特征可能丢失一些信息，所以考虑保留所有特征，但是让每一个θ都很小，使得y的预测不那么容易聚焦于某些特征，增加鲁棒。

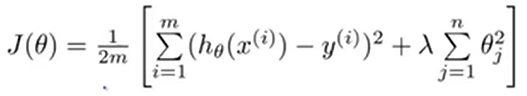


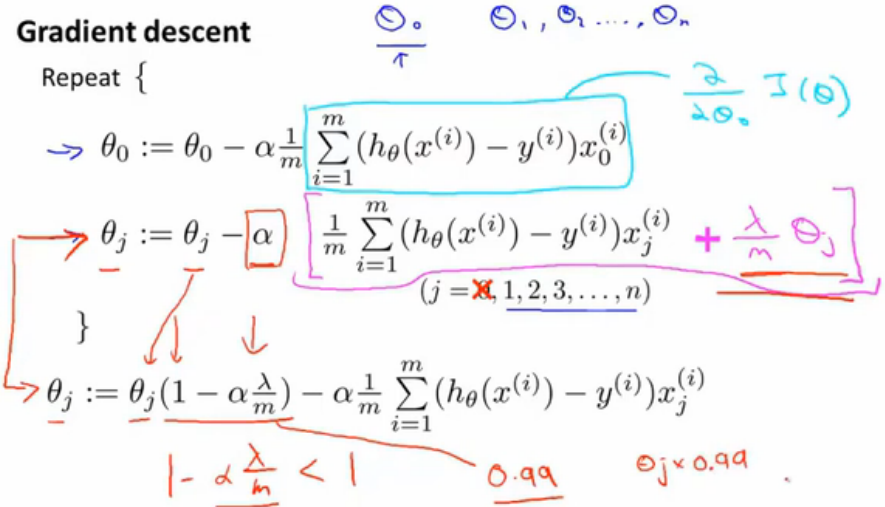
* 正则化参数**λ**对代价函数中加号左边的拟合训练数据与加号右边的更小的θ这两者进行了权衡。比如，当**λ**很大时，θ都接近0，此时的hθ(x)= θ0，这样又变成欠拟合。



（2）线性回归的正则化

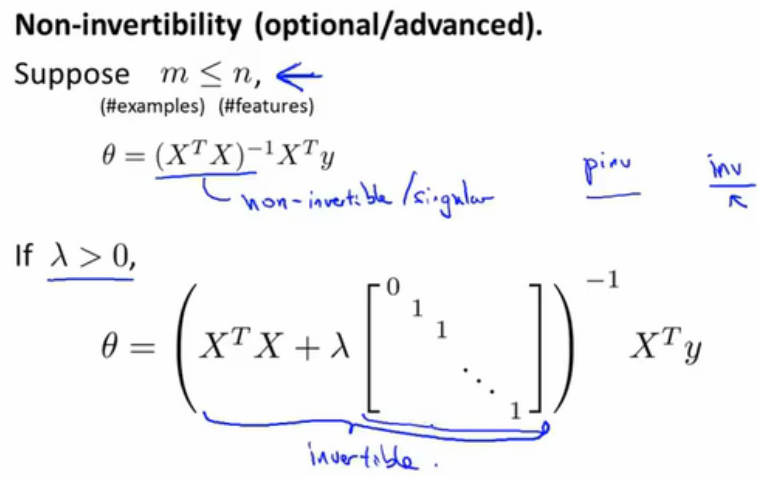
* 梯度下降法的正则化





* 正则方程的正则化

正规方程在遇到XX不可逆的情况时，正则化项可以使括号内的矩阵可逆。



（3）逻辑回归的正则化

代价函数和线性回归的一样，带正则化项。

