1.PCA降维的公式及原理分析:

假设对称矩阵A的所有特征值都不一样，那么: -> 对角矩阵

对于矩阵Y，它的协方差

假设Y=QX（此时不考虑降维），则 = =

PCA的本质:让协方差最小，方差最大 -> 协方差矩阵的对角元最大，非对角元为0->对角矩阵满足

我们已经得到两个公式: 和 ，假设Q =

可以得到: = ，因为协方差矩阵是对称半正定矩阵(特征值>=0)

所以可以得到:当Q = 时，是对角矩阵，因此如果Q少取几行就实现了降维

2.极大似然估计公式及原理分析:

极大似然估计起源于贝叶斯

假设特征为D，标签为A，P为概率

贝叶斯公式为:P(A|D) =

当样本给定的时候P(D)为常量，同时假定先验概率P(A)近似相等，那么我们可以得到:

P(A|D) = P(D|A)

对于机器学习的意义：将已知特征求标签的概率转化为已知标签求特征的概率，从无监督问题变成了有监督问题

但是极大似然的最大问题是:不同参数的P(A)不一定是相等的，只能说在样本不够大的时候是近似相等的

极大似然的两个假设：

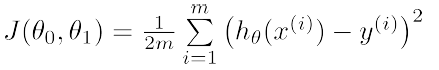
1. 已经发生的事件是独立重复事件，符合同一分布
2. 已经发生的事件是可能性（似然）最大的事件

3.线性回归

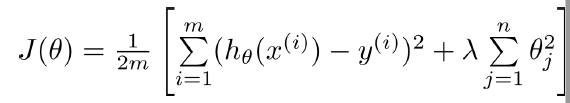
公式:

h(x) =

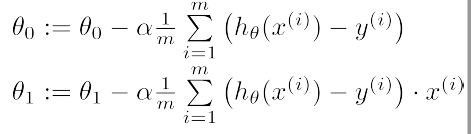
损失函数:



加上L2正则化后：



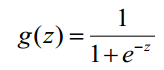
梯度下降:分别对求导，



4.逻辑回归

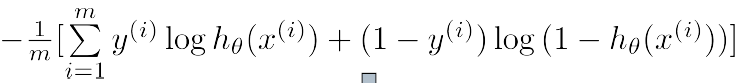
LR实际上是sigmoid+线性回归，用于判断边界，起到分类的作用

sigmoid公式：

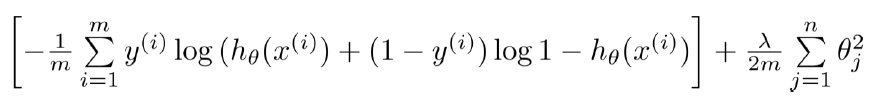


如果输入z = 线性回归，那么就是LR

损失函数:



损失函数+正则化项:



逻辑回归使用实例: