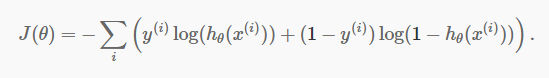
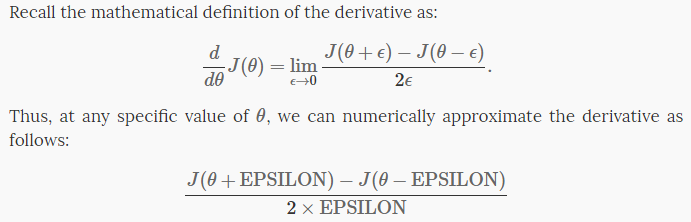
二分类算法的目标函数

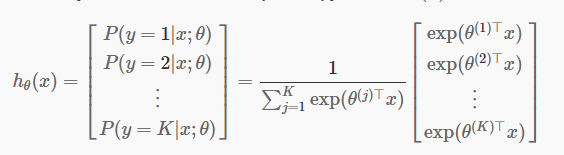


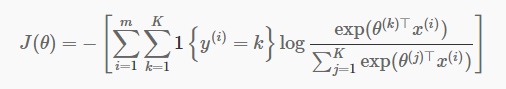
Note that only one of the two terms in the summation is non-zero for each training example (depending on whether the label y(i) is 0 or 1). Wheny(i)=1 minimizing the cost function means we need to make  hθ(x(i)) large, and when  y(i)=0 we want to make  1−hθ large as explained above.

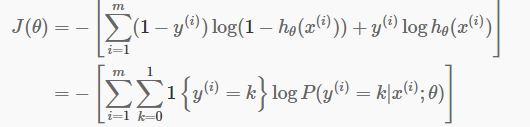


判断你的梯度计算方法是否正确，（通常取e=10-4）

Softmax regression (or multinomial logistic regression)









NMF

Find two non-negative matrices (W, H) whose product approximates the non-  
negative matrix X. This factorization can be used for example for  
dimensionality reduction, source separation or topic extraction.

components:H transformation: W  
The objective function is::  
  
 0.5 \* ||X - WH||\_Fro^2  
 + alpha \* l1\_ratio \* ||vec(W)||\_1  
 + alpha \* l1\_ratio \* ||vec(H)||\_1  
 + 0.5 \* alpha \* (1 - l1\_ratio) \* ||W||\_Fro^2  
 + 0.5 \* alpha \* (1 - l1\_ratio) \* ||H||\_Fro^2

Where::  
  
 ||A||\_Fro^2 = \sum\_{i,j} A\_{ij}^2 (Frobenius norm)  
 ||vec(A)||\_1 = \sum\_{i,j} abs(A\_{ij}) (Elementwise L1 norm)

解法：

ALS（迭代最小二乘法）

cd(坐标下降法)

pg(Project Gradient solver)

流程：(cd)

一.初始化W与H

二.判断是否对W和H进行正则化

令l1\_W=l1\_H=alpha\*l1\_ratio l2\_W=l2\_H=alpha\*(1-l1\_ration)

三.循环迭代更新W与H

(1).更新W

1.HHt=H\*H'+eye()\*l2\_W ,XHt=X\*H'-l1\_W\*ones() (X=WH XH'=WHH')

2.grad=W\*HHt-XHt

3.W=W-grad./diag(HHt)

4.W[W<0]=0

(2). 更新H

1.HHt=W'\*W+eye()\*l2\_H XHt=X'\*W-l1\_H\*ones() （X=WH X'=H'W' X'W=H'W'W）

2.grad = H'\*HHt-XHt

3.H'=H'-grad./diag(HHt)

4.H'[H'<0]=0

(3).判断是否达到退出条件，不满足则回到(1)