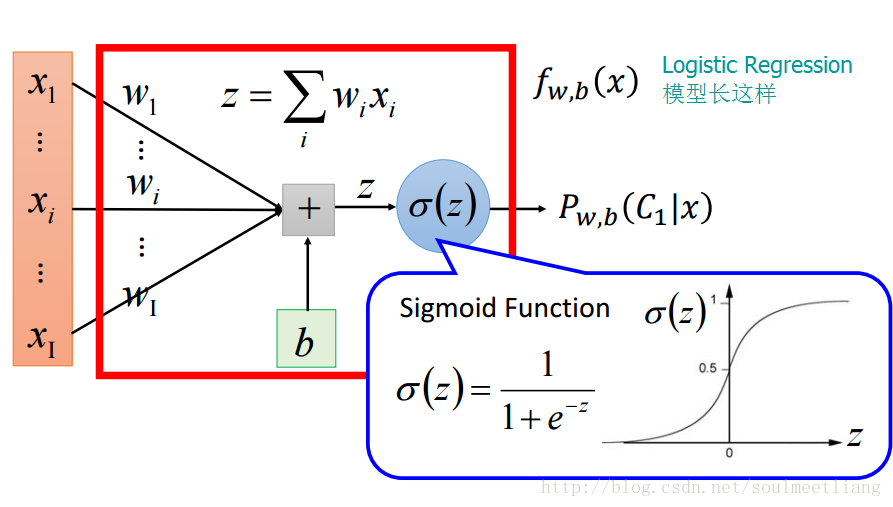
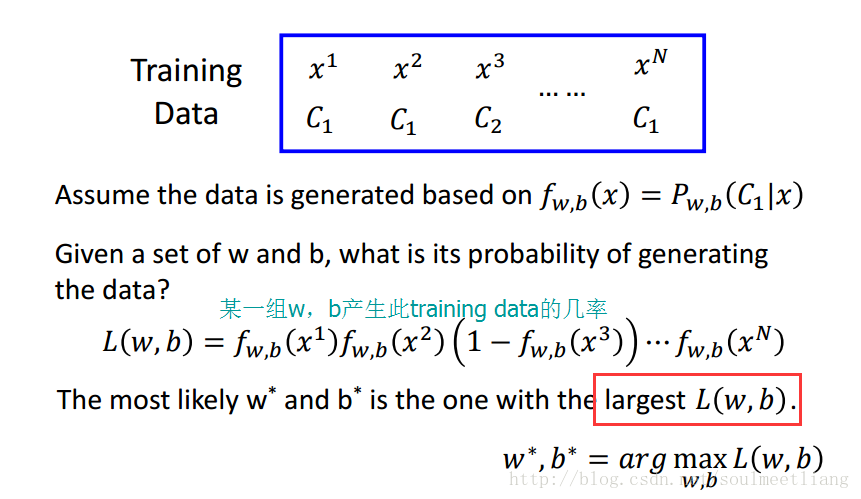
## Logistic regression

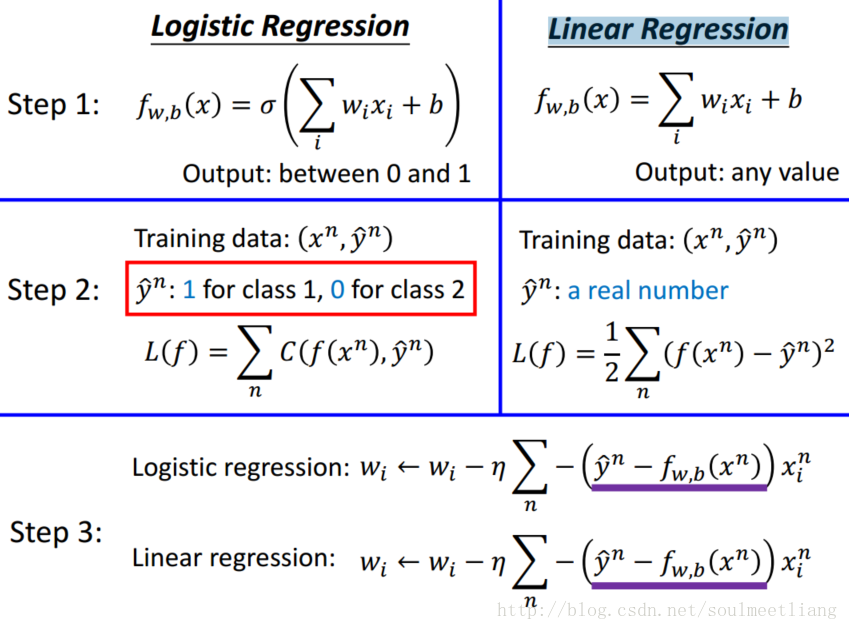
Logistic regression与Regression在函数上的差别，主要在于Logistic regression需要对结果使用sigmoid functon，从而logistic regression function的值在0~1之间，而regression的值可以是any value.

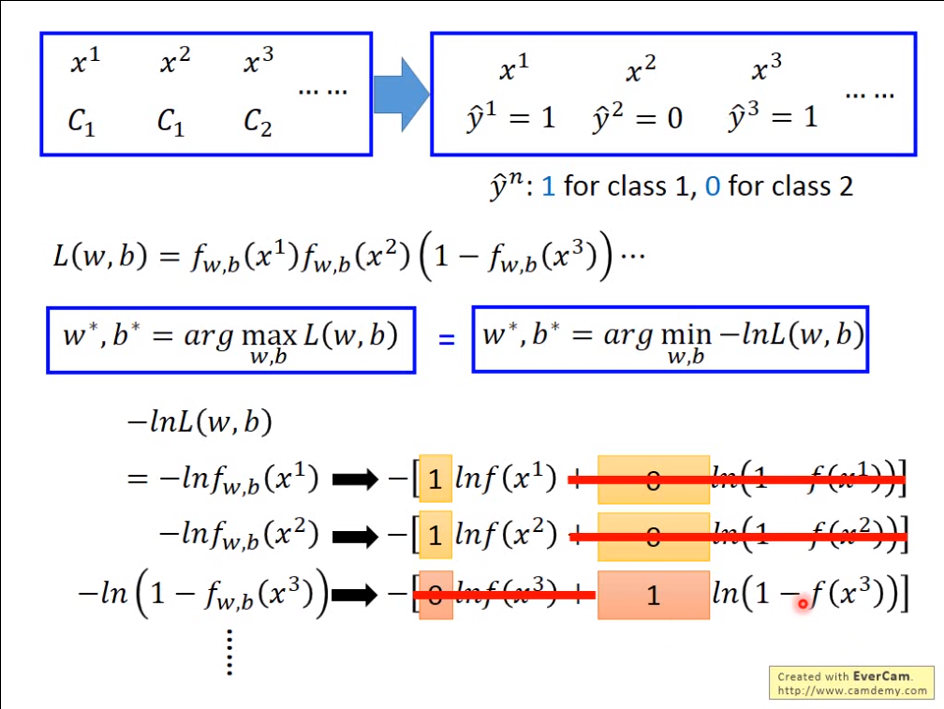
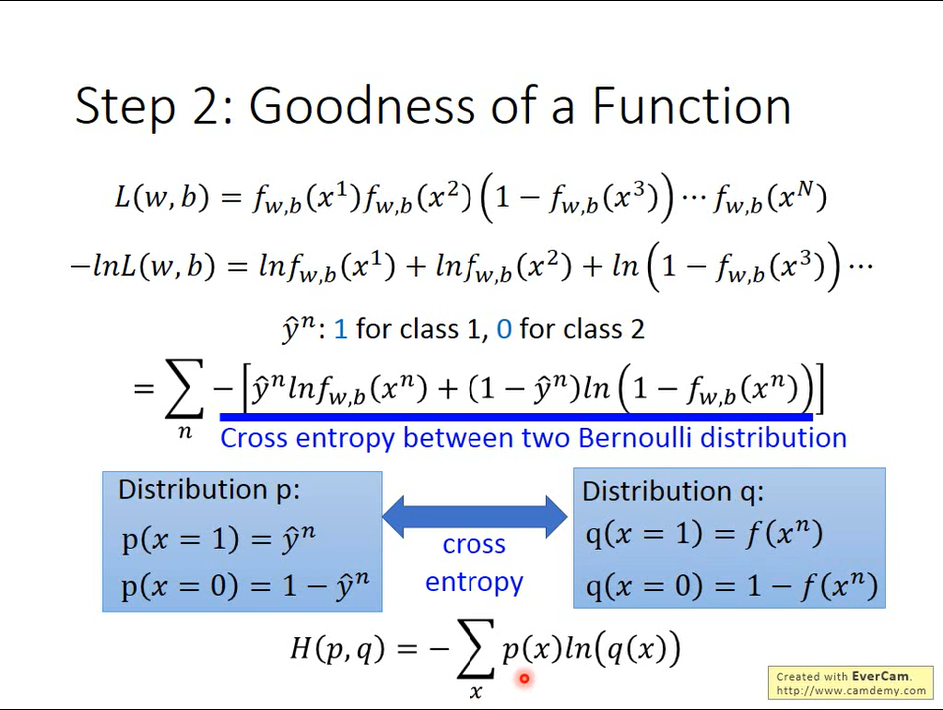


这里的fw,b(x)是针对L1的，因为sigmoid function的范围为0~1，0->L2，1->L1

Logistic Regression使用goodness of function来衡量模型的好坏，最大似然估计。

## LogisticRegressionVSLinearRegression





改变符号，

这样可以

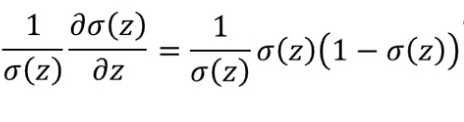
更快找到

Function，

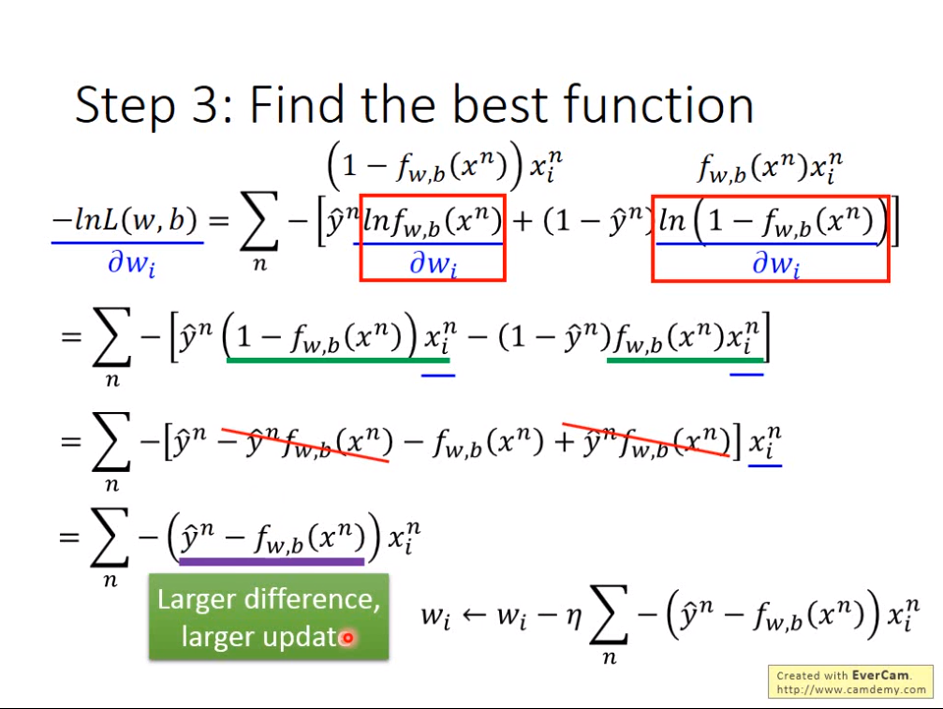
并且将G

oodness of function 转化为loss of function，利用gradient

交叉熵（cross entropy）

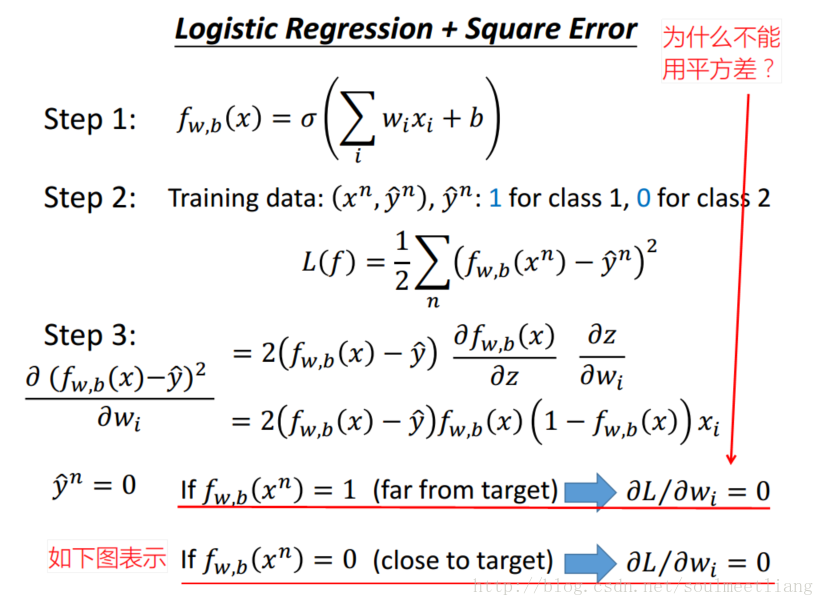


Sigmoid的导数=s\*(1-s)

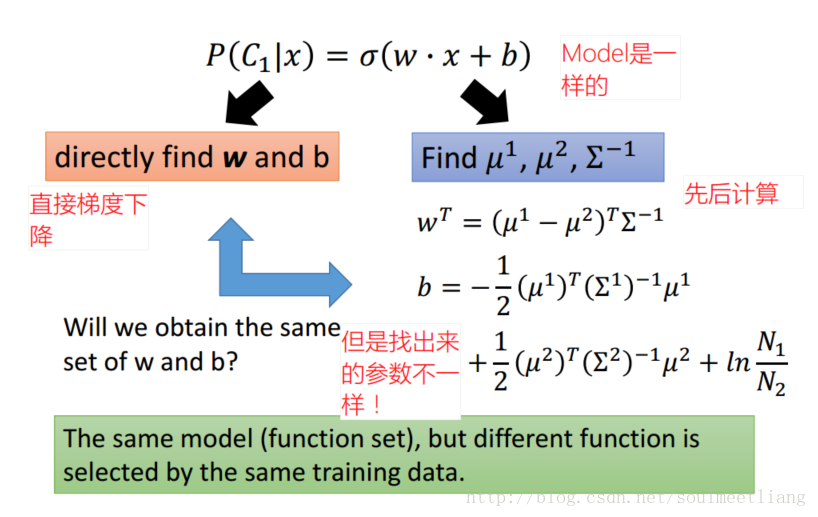


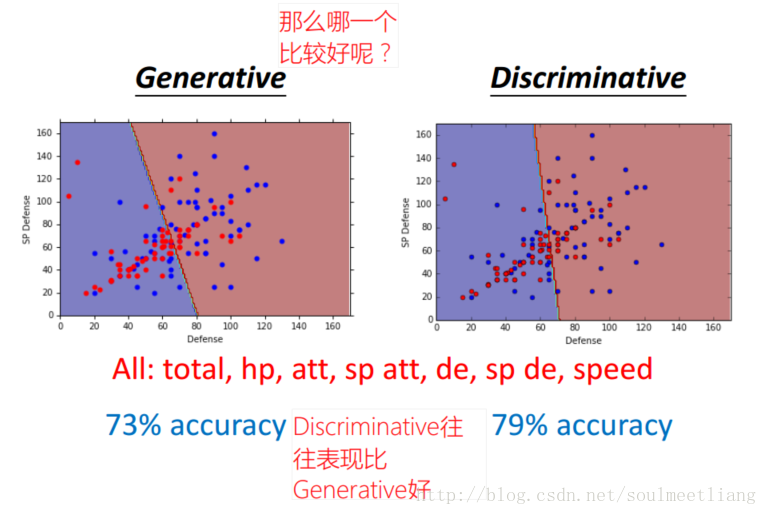
Gradient Descent

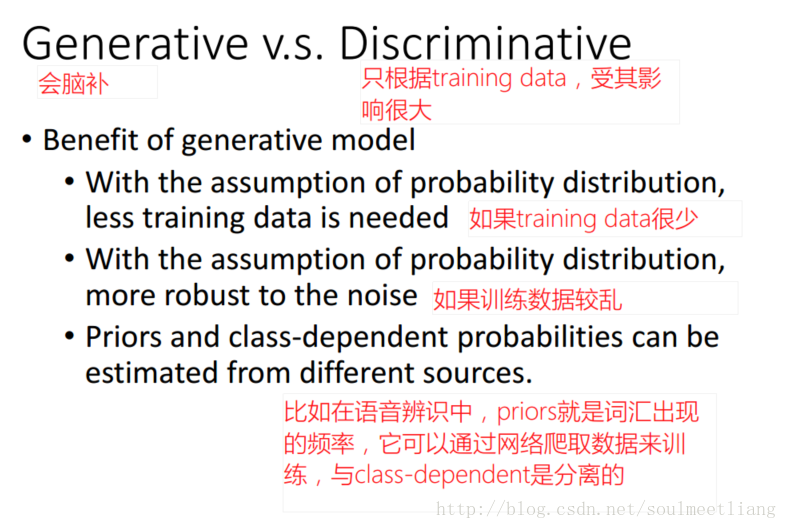
测试值与训练值的差距



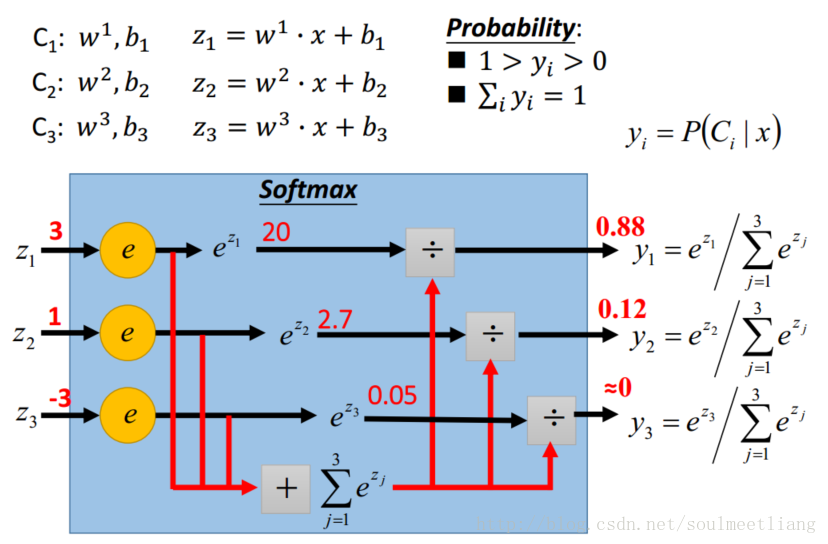
# 这里写图片描述Discriminative（Logstic） v.s. Generative（Gaussion）



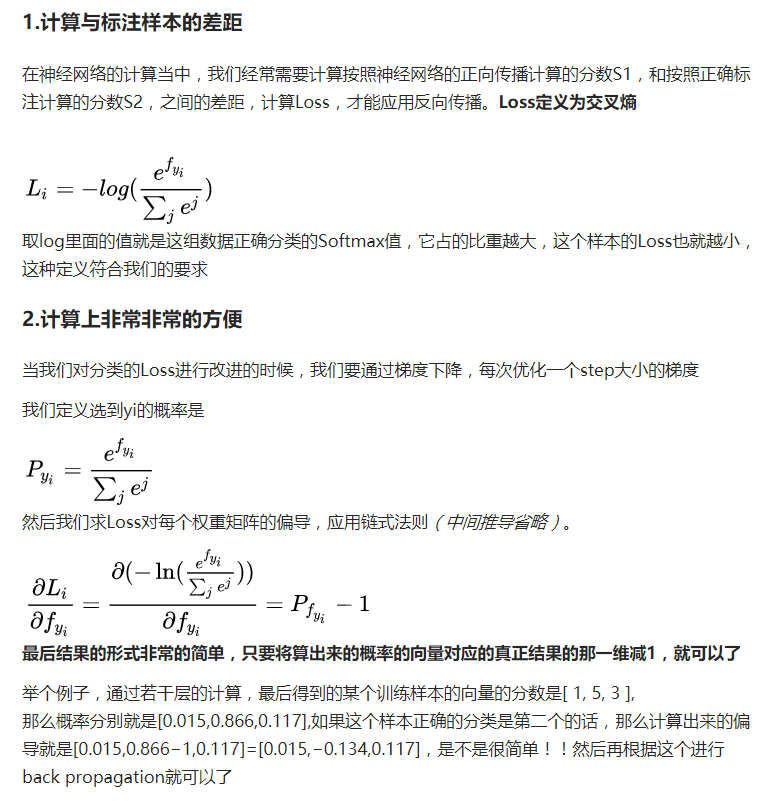
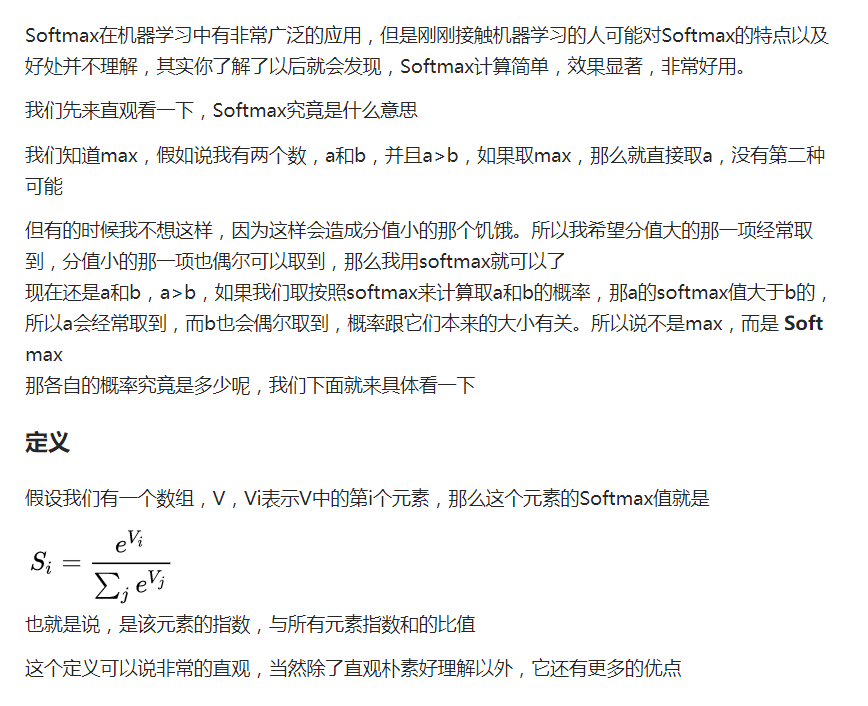


但Generative也有优势，discriminative不是完全比generative好

### Multi-class Classification

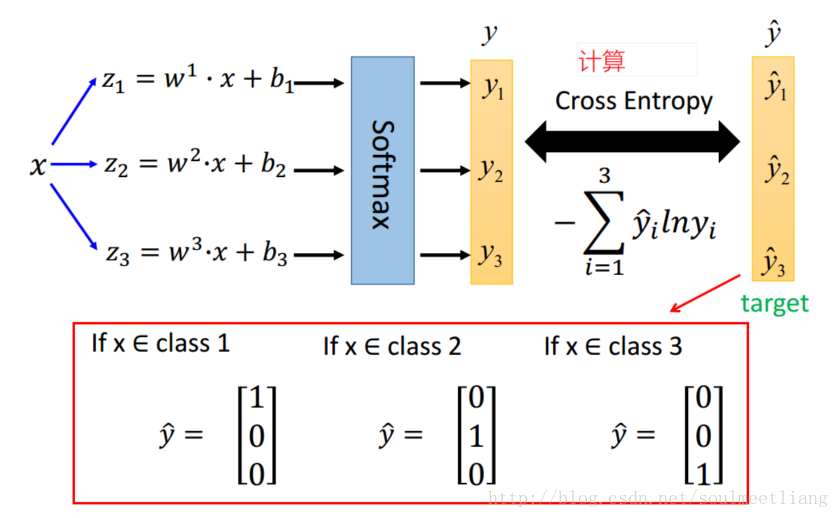


Softmax:



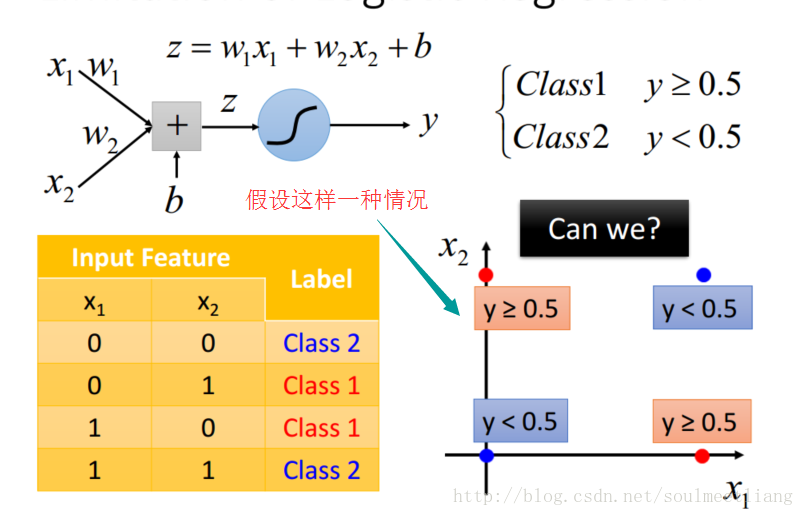
<https://www.zhihu.com/question/23765351>

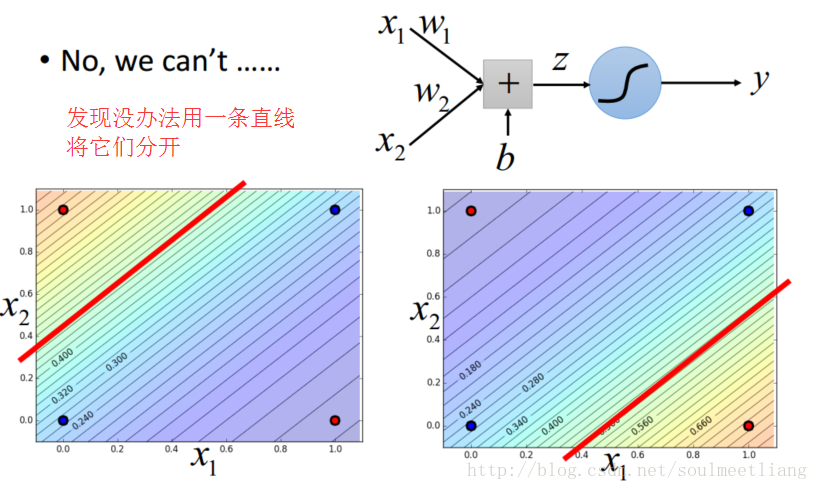
### Loss function of multi-classfication

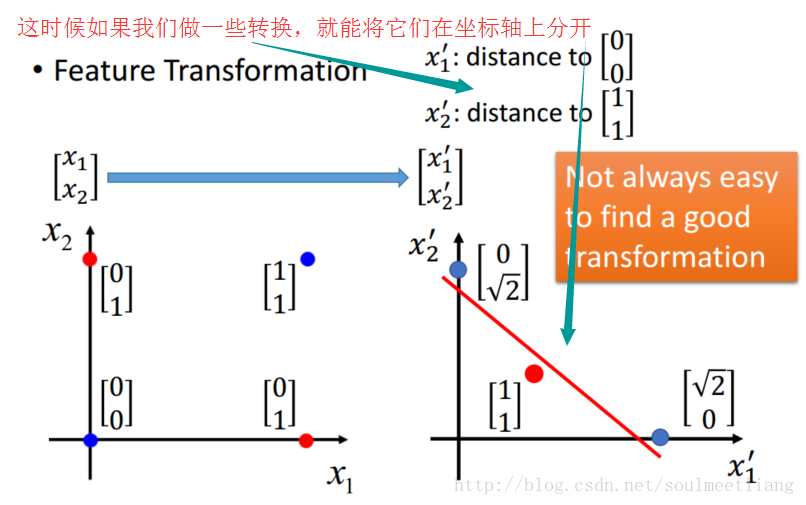


用向量表示

### Limitation of Logistic Regression







Cascading logistic regression models就是一种转换的普适方法。

