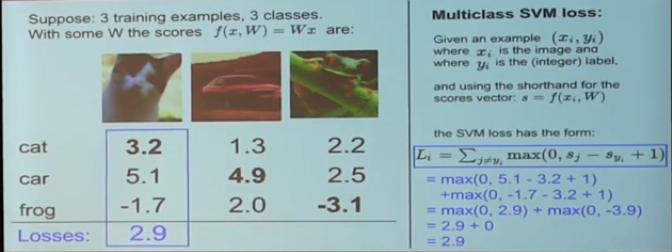
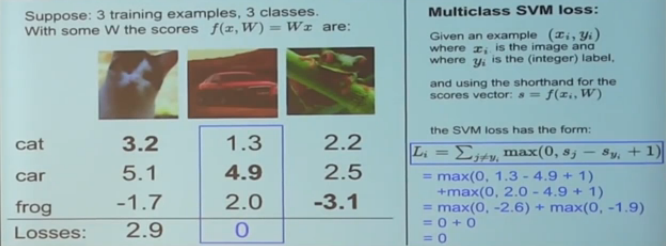
# 线性分类器的损失函数和最优化

损失函数，举个例子，如SVM loss：

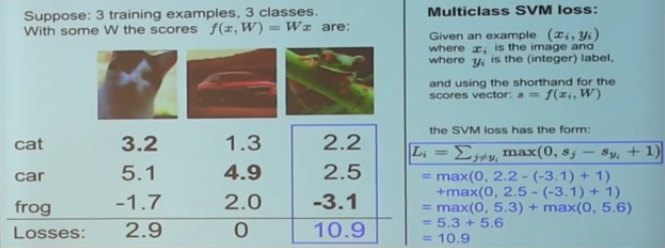


则第一幅图片的loss为2.9，反映了分类的不满意程度。

SVM loss中的1为安全系数，因此，可接受的其他不是cat类的得分上限为2.2（3.2-1），而car的5.1远超2.2，因此添加了loss的值；而frog的-1.7结果可以，没有提供loss

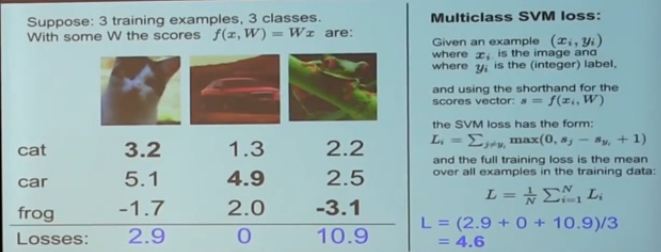


Loss等于0，说明结果还不错



Loss=10.9，糟糕的结果

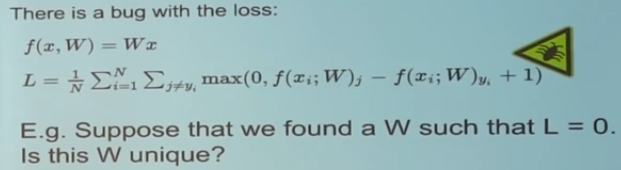
定义最终的loss结果：



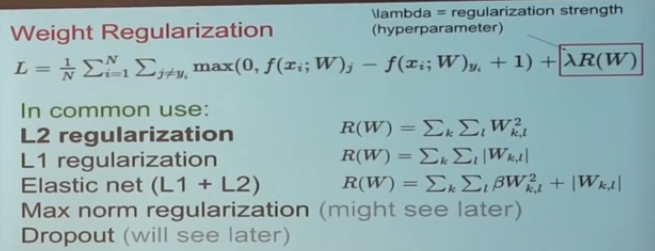
上述的loss被称为hinge loss：

加上平方，称为square hinge loss：

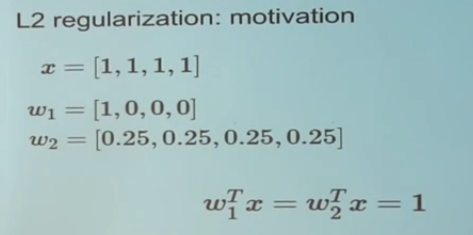
最终的损失函数（有bug，W不唯一）



解决方法：引入正则化项



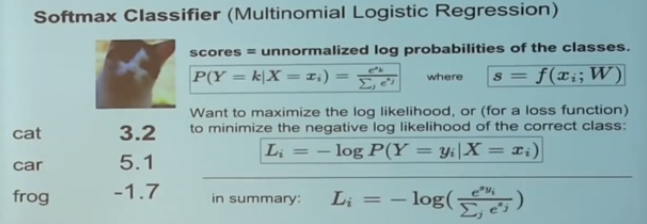
举个例子：

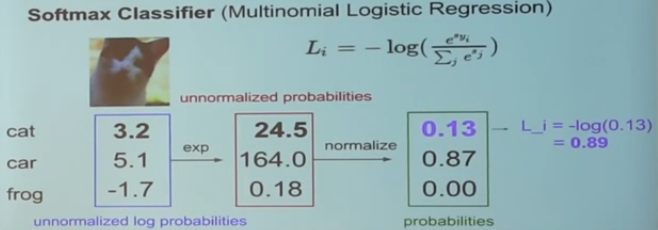


不加正则化项之前，w1，w2的结果是一样的；

加L2之后，w2的结果将好于w1（因为w2考虑了x中的所有维度，而w1只取第一维）

**Softmax**



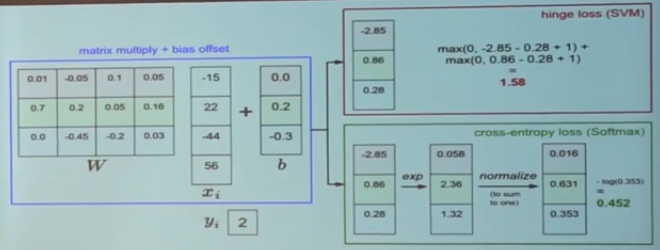


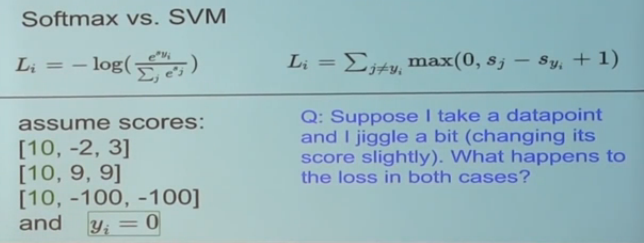
Li的取值范围是0-正无穷

同样的，初始的时候(第一次优化)，W的值应该是比较小的，这个时候Li的值应该是接近

-log(1/N)，N为类的数目

小结：





当数据点上下波动时，SVM（具有稳定性）得出的结果可能是相同的，因为SVM对得分低的点不感兴趣；而Softmax则会综合考虑每一个点，因此得出的结果会有所不同，这是两者最大的差异。

**优化策略**

1. 随机搜索——漫无目的
2. 利用梯度
3. 利用动量