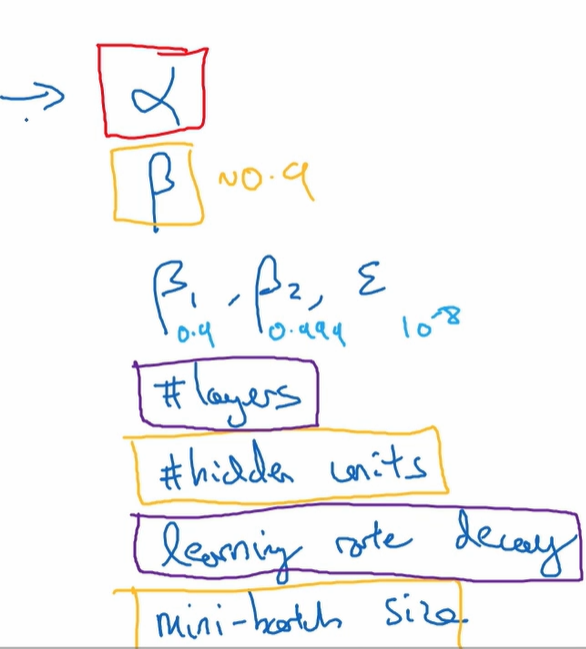
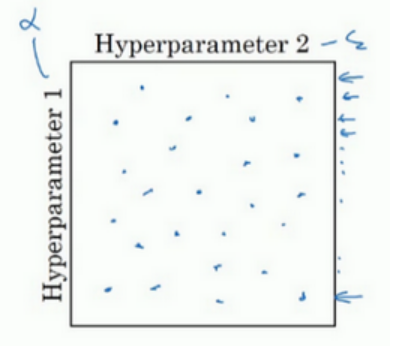
1. 超参数重要性排序

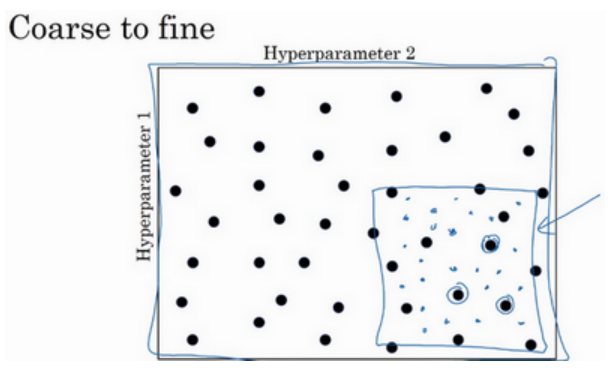


1. 网格搜索

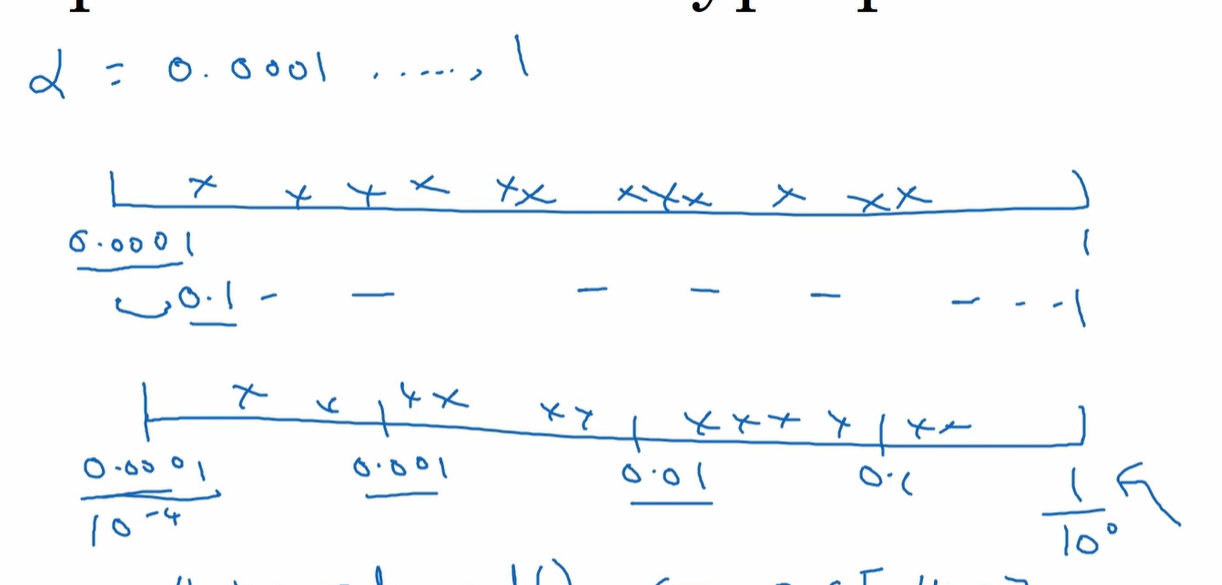


对于存在2个超参数的情况下，随机选取一系列的取值，并比较不同取值的效果。

在此基础上，再对一个更小的范围进行随机取值。

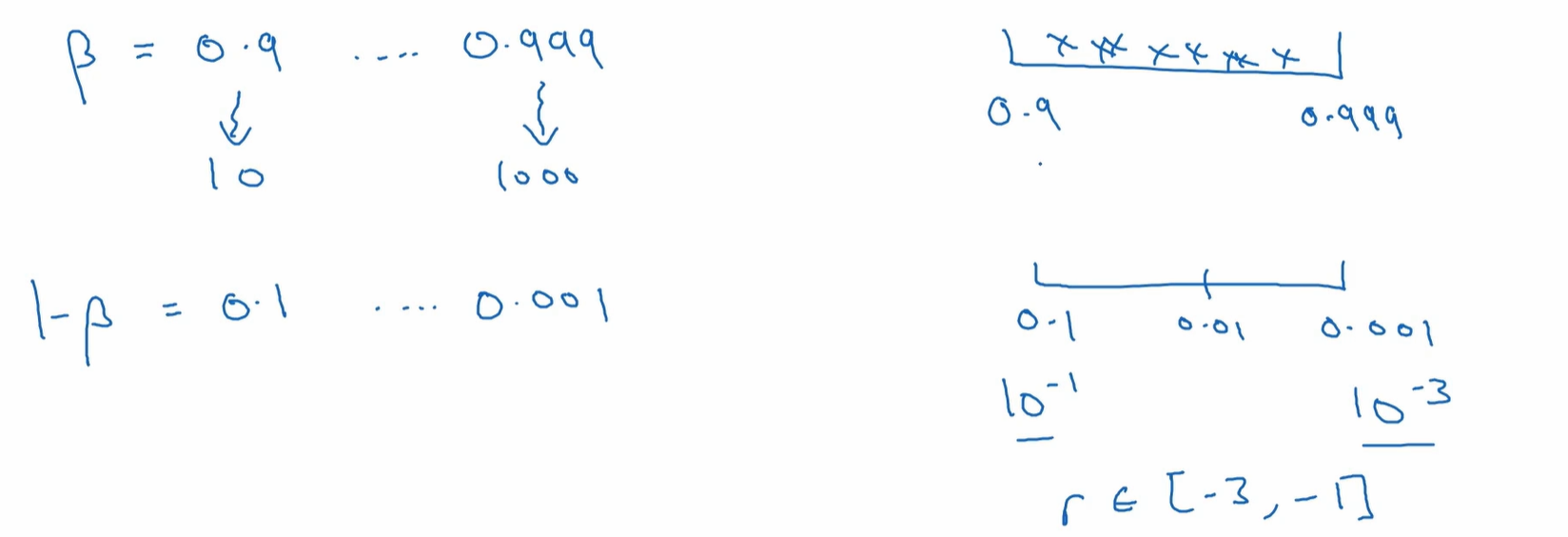


1. 学习率α的合适尺度



在对数尺度上进行随机选取

1. β的合适尺度

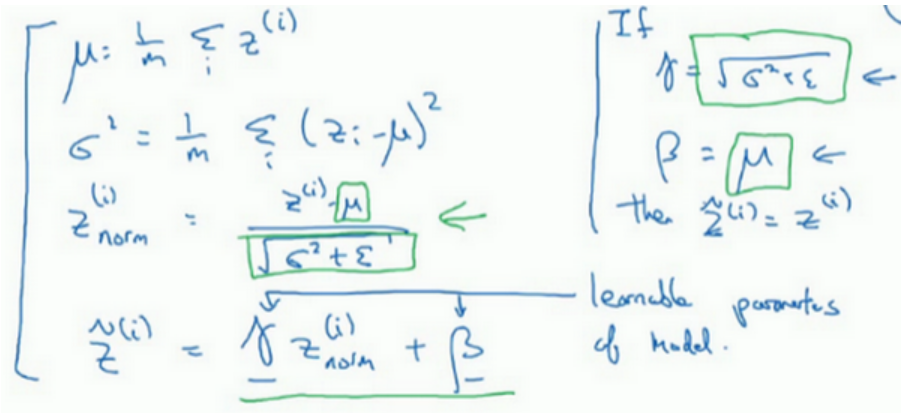


β本身的取值范围为0.9~0.999，因此对（1-β）在对数尺度上进行取值

1. Batch Norm

在逻辑回归中通过对输入参数进行归一化操作，将输入数据转化为均值为0，方差为1的数据，从而使得模型更快收敛。

在神经网络中也可以通过同样的方式对每一层的输入A[l]进行归一化操作，从而加快收敛速度。但是由于均值0，方差1的归一化并不适合各层输入数据，因此增加β和γ对归一化的尺度进行调整。



1. softmax回归

softmax通常用于多类别分类，拥有一个特殊的激活函数

