秦旭-20200301

一、 自己提出的问题的理解：

1、 算法8.4中，（2.b）中的J如何理解？（2.c）中寻找区域值怎么理解？

对残差进行拟合，得到新的回归树，回归树中的叶节点区域划分的数目就是J。

对区域进行划分以后，需要为区域赋值，此时选择合适的值，使损失函数最小。

2、 为什么会有Gradient boosting?

对于一般损失函数，直接优化不容易，就需要通过梯度提升计算。

负梯度是函数下降最快的方向，所以用梯度去拟合函数。只有在MSE的情况下，梯度才近似于残差的拟合。本质还是通过梯度寻求最优。

3、 GDBT怎么处理多酚类任务？

采取一对多的方法。1 vs n-1。

对每个类别训练M个分类器。假设有K个类别，那么训练完之后总共有M\*K颗树。

二、 别人提出的问题的理解：

4、 提升树模型第1次和第n次的损失函数有什么区别 (n>1) ?

因为第一次训练时没有前述模型，只能和真实值进行比较，所以采用对y的损失函数；

在后续的训练过程中，损失函数的计算会使当前模型拟合梯度（数据的残差是,其中是前m-1项的模型的预测和），从而整体优化模型。

5、 梯度提升改进了什么地方？

改进了优化的方式，简化了优化方法。同时特别注意，拟合的是梯度。

6. 梯度提升法对比提升树，为什么更work一点？

同上。

三、 读书计划

本周8.4

下周9