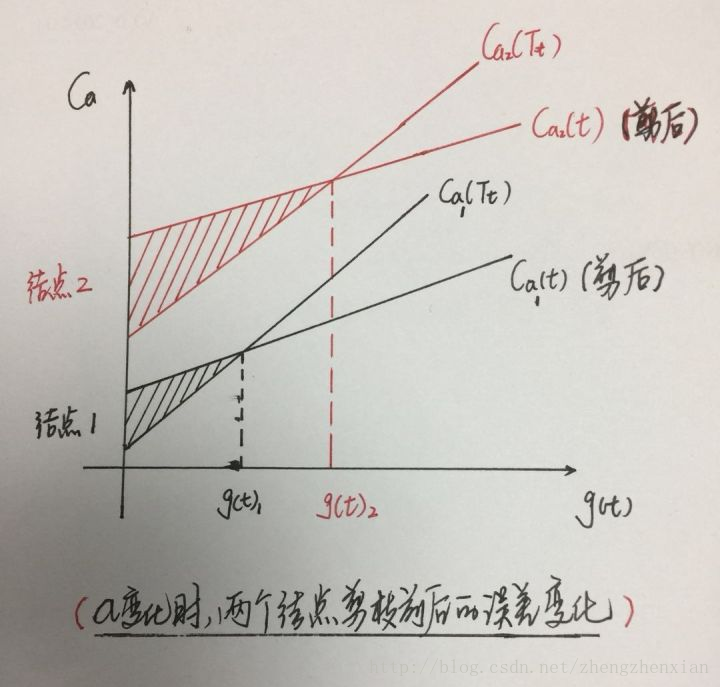
参考链接：<https://blog.csdn.net/zhengzhenxian/article/details/79083643>

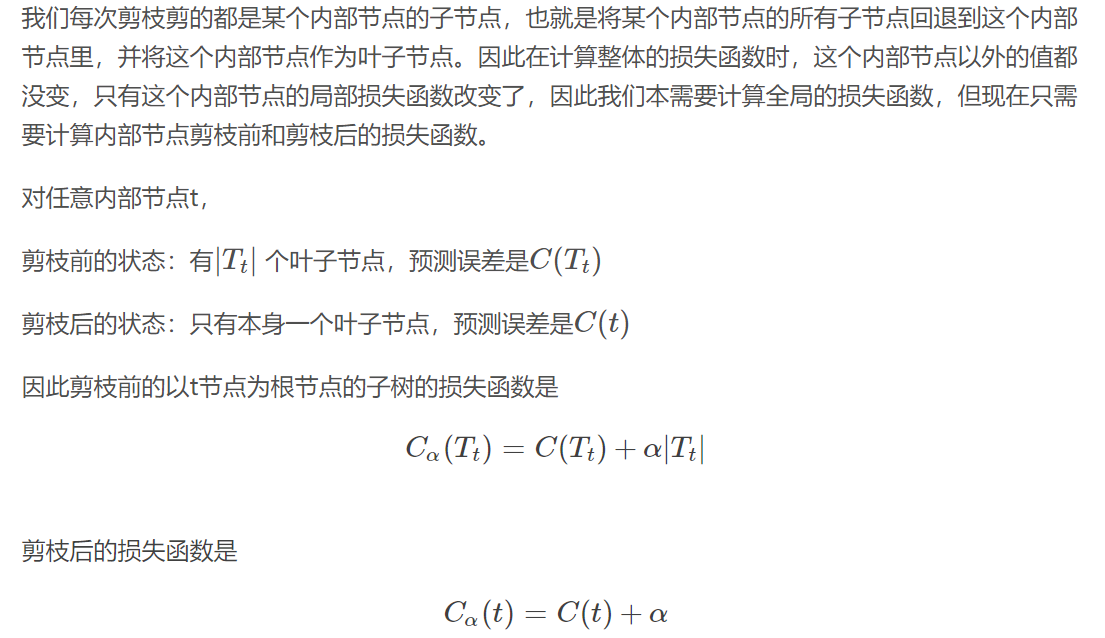
图解：



为什么要选择最小的g(t)呢？以图中两个点为例，结点1和结点2，g(t)2大于g(t)1, 假设在所有结点中g(t)1最小，g(t)2最大，两种选择方法：当选择最大值g(t)2，即结点2进行剪枝，

但此时结点1的不修剪的误差大于修剪之后的误差，即如果不修剪的话，误差变大，依次类推，对其它所有的结点的g(t)都是如此，从而造成整体的累计误差更大。反之，如果

选择最小值g(t)1，即结点1进行剪枝，则其余结点不剪的误差要小于剪后的误差，不修剪为好，且整体的误差最小。从而以最小g(t)剪枝获得的子树是该alpha值下的最优子树！



参考链接：

https://blog.csdn.net/wjc1182511338/article/details/76793164?locationNum=6&fps=1