《统计学习方法》第十章读书笔记2

一、自己提出的问题的理解：

1. 提出的问题1：维特比算法中变量和前向算法中有什么区别？

**讨论后的理解**：这个问题主要是理解各个字母的物理含义，维特比变量是t时刻到达状态i的所有序列概率中最大的概率，我们通过计算出这个最大值，然后保留住这个状态作为时刻t的隐状态。前向算法中中是记录t时刻所有状态的概率最大值。

1. 提出的问题2：维特比算法动态规划原理的细节是什么？逆推的方式为什么可以得出最优的路径？

**讨论后的理解**：

1）首先该算法通过迭代的方式找到概率最大的路径，将该路径作为最优路径，因此将该路径的最后一个结点作为最终结点。

2）当确定最后一个结点时，因为输入参数为，所以在已知最后一个结点的情况下，为一定值，所以此时，

所以将使得此式最大的j作为倒数第二个结点。

因为最后一个结点已经固定，有公式可确定倒数第二个结点转移到最后一个结点，概率最大，所以最优。

由于回溯考虑了之后的所有状态，而正推只能考虑前几个状态，所以采用从后往前进行回溯。

二、别人提出的问题的理解：

1. 问题3：条件随机场模型（CRF）和隐马尔可夫模型（HMM）在序列标注时有什么异同？

**讨论后的理解：**CRF可以定义数量更多，种类更丰富的特征函数。HMM模型具有天然具有局部性，就是说，在HMM模型中，当前的单词只依赖于当前的标签，当前的标签只依赖于前一个标签。这样的局部性限制了HMM只能定义相应类型的特征函数，我们在上面也看到了。但是CRF却可以着眼于整个句子s定义更具有全局性的特征函数。

三、读书计划

1. 本周完成的内容章节：第十章剩余部分
2. 下周计划：第十一章

四、读书摘要及理解

隐马尔可夫模型根据训练数据是包括观测序列和对应的状态序列还是只有观测序列，分为监督学习和无监督学习。一般情况下，标注数据工作量比较大，所以会采用无监督学习，学习的算法是Baum-Welch算法，这个算法利用了，其中I是隐变量，O是观测变量，lambda是我们学习的参数，通过将隐变量显式地表示出来，利用EM算法的思想，先确定一组lambda值，利用其去求当前观测状态和隐状态的期望值，也就是Q函数；然后最大化这个期望值即可。

隐马尔可夫模型预测问题是给定观测状态，利用你的参数值，给出隐藏状态，解决算法有两种，一种是近似算法，另一种是维特比算法。近似算法有个缺点是不能保证预测的状态序列整体是最有可能的状态序列，因为预测状态序列可能有实际不发生的情况。维特比算法利用了动态规划的思想，即当前找到的状态时最优的解，它的思路时计算每个时刻的状态发生的最大概率，直到最后即可求得全局最大概率的状态值。