HMM笔记

2020年2月18日 星期二 09:45

那么按照我们上一节H 型的定义，我们的观察集合是:

$$V = \{红，白\}，M = 2$$

我们的状态集合是:

$$Q = \{盒子1，盒子2，盒子3\}，N = 3$$

而观察序列和状态序列的长度为3.

初始状态分布为:

$$\Pi = (0.2, 0.4, 0.4)^T$$

状态转移概率分布矩阵为:

$$\boxed{A = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}}$$

观测状态概率矩阵为:

$$B = \begin{pmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \\ 0.7 & 0.3 \end{pmatrix}$$

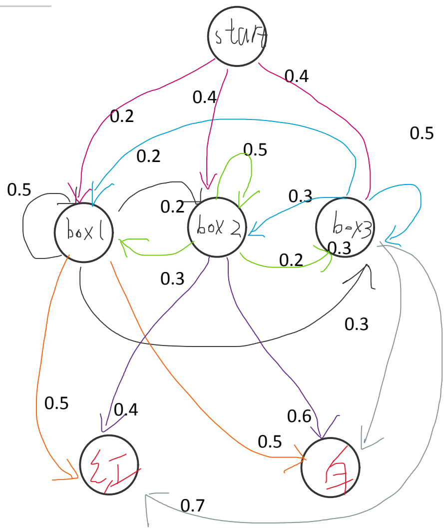


Fig 1. 抽球例子的概率转移图

π: 初始状态分布，表示抽到盒子1的概率为0.2，抽到盒子2的概率为0.4，抽到盒子3的概率为0.4。
A：状态转移概率，每一行表示在当前时刻抽到各个盒子的概率，每一列表示当前盒子在各个时刻的概率分布。
B:观测状态，行表示在当前时刻抽到红球、白球的概率分布，每列表示当前球色在不同时刻的概率分布。

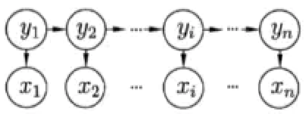


图 14.1 隐马尔可夫模型的图结构

该HMM模型满足下面两个假设：

- **齐次马尔科夫链假设。**即任意时刻的隐藏状态只依赖于它前一个隐藏状态
- **观测独立性假设。**即任意时刻的观察状态只仅仅依赖于当前时刻的隐藏状态

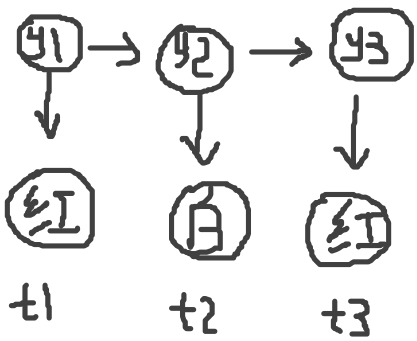
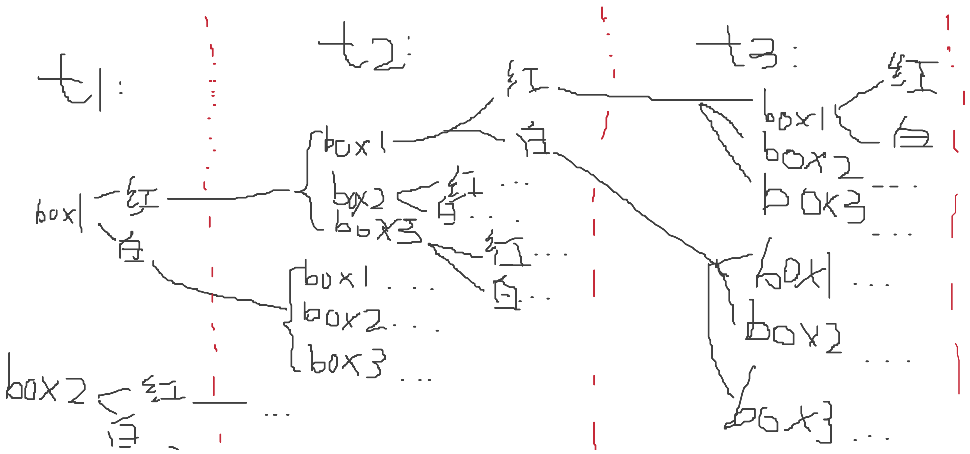


Fig 2抽球例子HMM模型图



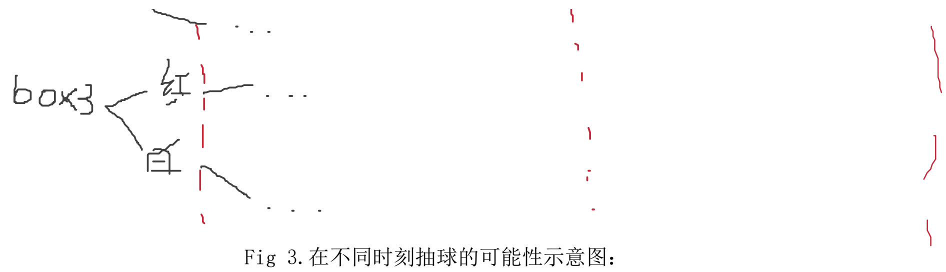


Fig 3.在不同时刻抽球的可能性示意图:

上面共有2*3*2*3*2*3=6^3=36*6=216种情况。

（1）HMM模型的三个基本问题：

1.评估观察序列概率问题-->前向后向算法：计算在观测序列O={o_1,o_2,...o_T}和估计模型λ=(A,B,π)下观测序列O出现的概率，**即计算P(O|λ)的概率。**

2.模型参数学习问题-->鲍姆-韦尔奇算法：即给定观测序列O={o_1,o_2,...o_T}和估计模型λ=(A,B,π)的参数值，使该模型下观测序列的**条件概率P(O|λ)最大。**

3.解码问题-->维特比算法：用来求给定观测序列O={o_1,o_2,...o_T}和估计模型λ=(A,B,π)的参数值下，**最可能出现的对应的状态序列。**

以下例子均以假设我们观测到的球色顺序是O={红，白，红}为推算结果：

（2）前向算法

前向算法基本思路是t1时刻观测序列为{红球}，t2时刻的观测序列为{红、白}，t3时刻下观测序列为{红，白，红}。已知前面的一个观测状态，依次往后计算。

时刻t1下，隐藏状态分别抽到盒子1、盒子2、盒子3是红色球的概率为：

box at t_1	probability
1（盒子1是红球的概率）	$0.2*0.5=0.1$
2（盒子2是红球的概率）	$0.4*0.4=0.16$
3（盒子3是红球的概率）	$0.4*0.7=0.28$

在时刻t2下，3个盒子转移到不同盒子的概率以及生成白球的概率：

时刻2是白色球，隐藏状态是盒子1的概率为：

$$\alpha_2(1)=\left[\sum_{i=1}^{3}\alpha_1(i)a_{i1}\right]b_1(o_2)=\underbrace{[0.1*0.5+0.16*0.3+0.28*0.2]}_{\text{抽到盒子1的概率}}\times\underbrace{0.5}_{\text{盒子1下白球的概率}}=0.077$$

隐藏状态是盒子2的概率为：

$$\alpha_2(2)=\left[\sum_{i=1}^{3}\alpha_1(i)a_{i2}\right]b_2(o_2)=[0.1*0.2+0.16*0.5+0.28*0.3]\times0.6=0.1104$$

隐藏状态是盒子3的概率为：

$$\alpha_2(3)=\left[\sum_{i=1}^{3}\alpha_1(i)a_{i3}\right]b_3(o_2)=[0.1*0.3+0.16*0.2+0.28*0.5]\times0.3=0.0606$$

时刻3是红色球，隐藏状态是盒子1的概率为：

$$\alpha_3(1)=\left[\sum_{i=1}^{3}\alpha_2(i)a_{i1}\right]b_1(o_3)=[0.077*0.5+0.1104*0.3+0.0606*0.2]\times0.5=0.04187$$

隐藏状态是盒子2的概率为：

$$\alpha_3(2)=\left[\sum_{i=1}^{3}\alpha_2(i)a_{i2}\right]b_2(o_3)=[0.077*0.2+0.1104*0.5+0.0606*0.3]\times0.4=0.03551$$

隐藏状态是盒子3的概率为：

$$\alpha_3(3)=\left[\sum_{i=1}^{3}\alpha_3(i)a_{i3}\right]b_3(o_3)=[0.077*0.3+0.1104*0.2+0.0606*0.5]\times0.7=0.05284$$

最终我们求出观测序列:$O=\{$红，白，红$\}$的概率为：

$$P(O|\lambda)=\sum_{i=1}^{3}\alpha_3(i)=0.13022$$

（三）后向算法

和前向算法相反，这里是已知t3时刻观测序{红_t3}，t2观测序列为{白_t2，红_t3}，t1时刻的观测序列是{红_t1，白_t2，红_t3}，这里的红_t3表示时刻t3观测状态为红球,其它类似。**其实前向算法是从前往后计算，后向算法是从后往前计算；即向前算法是已知前面的i个观测状态来计算后面第i+1个观测状态的概率，而后向算法是已知后面的第i+1个观测状态来计算前面第i个观测状态的概率。**

因为最后一个观测数据已经确定，所以初始化三个状态的概率为1，表示从盒子1抽到球且观测为空的概率是1，盒子2抽到球且观测序列为空的概率为1，盒子3抽到球且观测序列为空的概率为1。
在t3时刻，观测序列分别在盒子1、2、3下的观测概率为空且抽到红球的概率如下：

box at t_T	color_probability
1（从盒子1中选一个球，观测序列为空的概率，即选中红球的概率）	$1*0.5=0.5$
2（从盒子2中选一个球，观测序列为空的概率，即选中红球的概率）	$1*0.4=0.4$
3（从盒子3中选一个球，观测序列为空的概率，即选中红球的概率）	$1*0.7=0.7$

在t2时刻下，t3时刻在盒子1、2、3中选到白球的概率分布：
抽到盒子1、2、3且白球：$\beta_2(1)=(0.5*0.5+0.4*0.2+0.7*0.3)*0.5=0.27$
抽到盒子1、2、3且白球：$\beta_2(2)=(0.5*0.3+0.4*0.5+0.7*0.2)*0.6=0.294$
抽到盒子1、2、3且白球：$\beta_2(3)=(0.5*0.2+0.4*0.3+0.7*0.5)*0.3=0.171$
在t1时刻下，选到红球的概率分布：
$\beta_1(1)=(0.27*0.5+0.294*0.2+0.171*0.3)*0.5=0.12255$
$\beta_1(2)=(0.27*0.3+0.294*0.5+0.171*0.2)*0.4=0.10488$
$\beta_1(3)=(0.27*0.2+0.294*0.3+0.171*0.5)*0.7=0.15939$
在初始状态π下：
0.12255*0.2+0.10488*0.4+0.15939*0.4=0.130218
所以把最后的概率分布加起来就是后向算法的观测序列O={红，白，红}生成的概率.这里和前向算法的结果其实是相等的，因为前向算法的数值进行了舍入。

下图是前向算法的伪代码：

我们的动态规划从时刻1开始，到时刻T结束，由于$\alpha_T(i)$表示在时刻T观测序列为$o_1,o_2,\dots o_T$，并且时刻T隐藏状态q_i的概率，我们只要将所有隐藏状态对应的概率相加，即$\sum_{i=1}^N \alpha_T(i)$就得到了在时刻T观测序列为$o_1,o_2,\dots o_T$的概率。

下面总结下前向算法。

输入：HMM模型$\lambda=(A,B,\Pi)$，观测序列$O=(o_1,o_2,\dots o_T)$

输出：观测序列概率$P(O|\lambda)$

1) 计算时刻1的各个隐藏状态前向概率：

$$\alpha_1(i)=\pi_i b_i(o_1),\ i=1,2,\dots N$$

2) 递推时刻$2,3,\dots T$时刻的前向概率：

$$\alpha_{t+1}(i)=\Big[\sum_{j=1}^N \alpha_t(j)a_{ji}\Big]b_i(o_{t+1}),\ i=1,2,\dots N$$

3) 计算最终结果：

$$P(O|\lambda)=\sum_{i=1}^N \alpha_T(i)$$

从递推公式可以看出，我们的算法时间复杂度是$O(TN^2)$，比暴力解法的时间复杂度$O(TN^T)$少了几个数量级。

下图是后向算法的伪代码：

现在我们总结下后向算法的流程,注意下和前向算法的相同点和不同点：

输入：HMM模型$\lambda=(A,B,\Pi)$，观测序列$O=(o_1,o_2,\dots o_T)$

输出：观测序列概率$P(O|\lambda)$

1) 初始化时刻T的各个隐藏状态后向概率：

$$\beta_T(i)=1,\ i=1,2,\dots N$$

2) 递推时刻$T-1,T-2,\dots 1$时刻的后向概率：

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(o_{t+1}) \beta_{t+1}(j),\ i = 1, 2, \ldots N$$

3）计算最终结果：

$$P(O|\lambda) = \sum_{i=1}^{N} \pi_i b_i(o_1) \beta_1(i)$$

此时我们的算法时间复杂度仍然是$O(TN^2)$。