

387 人赞同了该回答

最高票的回答结果是正确的，但是并没有把viterbi算法讲明白，尤其是其中的dp思想。可以用相同的例子来解释：

1.题目背景：

从前有个村儿，村里的人的身体情况只有两种可能：健康或者发烧。假设这个村儿的人没有体温计或者百度这种神奇东西，他唯一判断他身体情况的途径就是到村头我的偶像金正月的小诊所询问。月儿通过询问村民的感觉，判断她的病情，再假设村民只会回答正常、头晕或冷。有一天村里奥巴马就去月儿那去询问了。第一天她告诉月儿她感觉正常。第二天她告诉月儿感觉有点冷。第三天她告诉月儿感觉有点头晕。那么问题来了，月儿如何根据阿驴的描述的情况，推断出这三天中阿驴的一个身体状态呢？为此月儿上百度搜 google，一番狂搜，发现维特比算法正好能解决这个问题。月儿乐了。

2.已知情况：

隐含的身体状态 = { 健康, 发烧 }
可观察的感觉状态 = { 正常, 冷, 头晕 }
月儿预判的阿驴身体状态的概率分布 = { 健康: 0.6, 发烧: 0.4 }
这就是初始状态序列。

月儿认为的阿驴身体健康状态的转换概率分布 = {
健康->健康: 0.7,
健康->发烧: 0.3,
发烧->健康: 0.4,
发烧->发烧: 0.6
}

这样就可以列出相应的状态转移矩阵。(人懒。。不想编辑公式了)

月儿认为的在相应健康状况条件下，阿驴的感觉的概率分布 = {
健康, 正常: 0.5, 冷: 0.4, 头晕: 0.1;
发烧, 正常: 0.1, 冷: 0.3, 头晕: 0.6
}

这样就可以列出相应的观测矩阵。

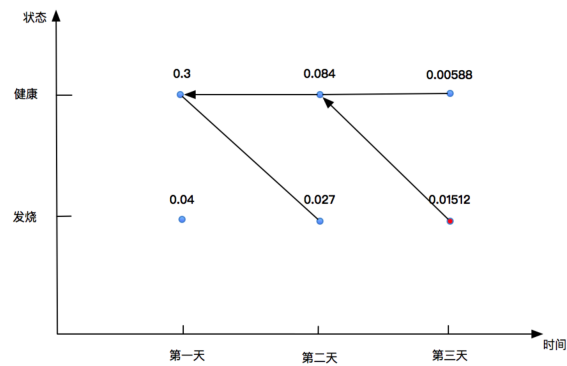
由上面我们可以发现，HMM的三要素都齐备了，下面就是解决问题了。阿驴连续三天的身体感觉依次是：正常、冷、头晕。

3.题目：

已知如上，求：阿驴这三天的身体健康状态变化的过程是怎样的？即已知观测序列和HMM模型的情况下，求状态序列。

4.过程：

- 初始化。第一天的时候，对每一个状态（健康或者发烧），分别求出第一天身体感觉正常的概率：P(第一天健康) = P(正常|健康)\*P(健康|初始情况) = 0.5 \* 0.6 = 0.3 P(第一天发烧) = P(正常|发烧)\*P(发烧|初始情况) = 0.1 \* 0.4 = 0.04
第二天的时候，对每个状态，分别求在第一天状态为健康或者发烧情况下观察到冷的最大概率。在维特比算法中，我们先求得路径的单个路径的最大概率，然后再乘上观测概率。P(第二天健康) = max(0.3\*0.7, 0.04\*0.4)\*0.4=0.3\*0.7\*0.4=0.084 此时我们需要记录概率最大的路径的前一个状态，即0.084路径的前一个状态，我们在小本本上记下，第一天健康。 P(第二天发烧)=max(0.3\*0.3, 0.04\*0.6)\*0.3=0.027, 同样的在0.027这个路径上，第一天也是健康的。
第三天的时候，跟第二天一样。P(第三天健康)=max(0.084\*0.7, 0.027\*0.4)\*0.1=0.00588, 在这条路径上，第二天是健康的。P(第三天发烧)=max(0.084\*0.3, 0.027\*0.6)\*0.6=0.01512, 在这条路径上，第二天是健康的。
最后一天的状态概率分布即为最优路径的概率，即P(最优)=0.01512，这样我们可以得到最优路径的终点，是发烧
由最优路径开始回溯。请看我们的小本本，在求得第三天发烧概率的时候，我们的小本本上面写的是第二天健康，好了，第二天就应该是健康的状态，然后在第二天健康的情况下，我们记录的第一天是健康的。这样，我们的状态序列逆推出来了。即为：健康，健康，发烧。
简略的画个图吧：



这儿的箭头指向就是一个回溯查询小本本的过程，我们在编写算法的时候，其实也得注意，每一个概率最大的单条路径上都要把前一个状态记录下来。

编辑于 2018-04-05 19:21