

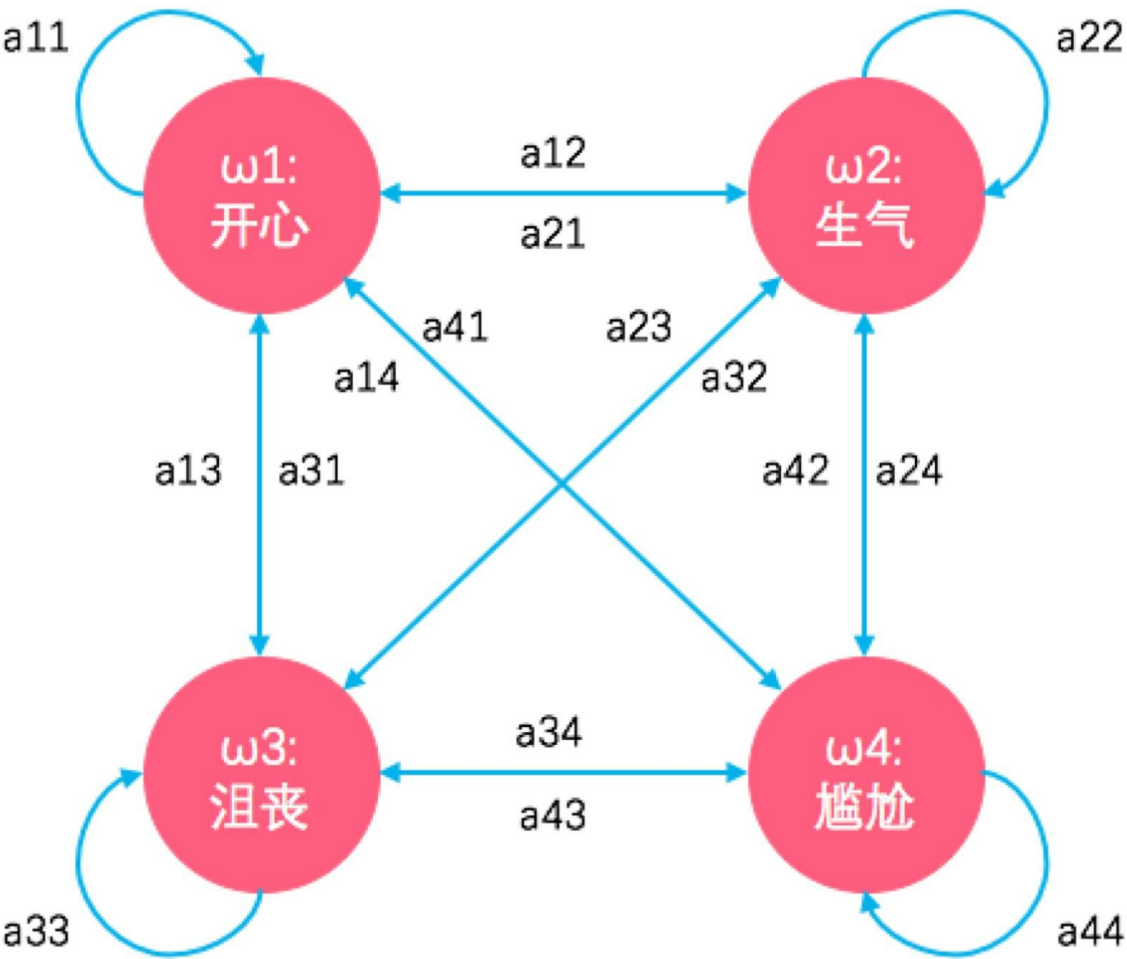
如果你跟夕小瑶恋爱了... (下)

原创 夕小瑶 夕小瑶的卖萌屋 2017-03-24

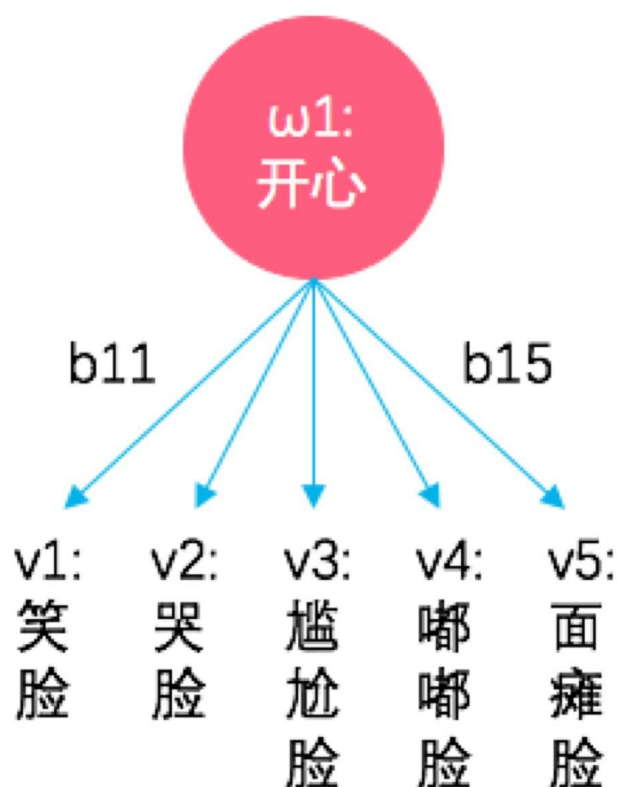
AB派

在上一篇文章中, 你成功的将“挽回夕小瑶”的任务卡进了隐马尔可夫模型(HMM)中。那么我们来规范化的整理一下已经有的信息和需要计算得到的信息。

还记得这两个图嘛?这就是我们建立好的模型。



(隐状态的转移图)



(每个隐状态 ω_i 都有概率发出5种可以观测到的信号)

对于第一张图, 这么多的参数看起来也蛮乱的, 那就将所有的状态转移概率 a_{ij} 存储到一个矩阵A中:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

矩阵A中的每个元素 a_{ij} 就代表当前状态为 ω_i 时, 下一状态为 ω_j 的概率(即状态 ω_i 到状态 ω_j 的转移概率)。

对于第二张图, 描述的是当(隐)状态为 ω_j 时, 发出信号 v_k 的概率。所以用 b_{jk} 来表示 ω_j 发出信号 v_k 的概率。将 b_{jk} 存储到矩阵B中:

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \end{bmatrix}$$

好~A矩阵和B矩阵显然就是我们需要计算出的模型参数啦。但是参数中仅仅是A和B就够了吗？

想一下，虽然A可以描述从某个状态转移到某个状态的概率，但是每个状态序列总要有个开头呀~这个开头是什么样子的，在A矩阵和B矩阵中都没有描述。

所以模型还有一个描述初始状态的参数，也就是描述每个隐状态 ω_i 作为初始状态的概率，记为 π_i 。也就是向量 π ：

$$\pi = [\pi_1 \ \pi_2 \ \pi_3 \ \pi_4]$$

整理完毕~**A、B、 π** 就是我们全部要计算出的模型参数。

开始训练！

而我们已经有了夕小瑶好多天的隐状态序列和对应的观测序列的数据了，那么我们如何用它们来训练出模型的参数呢。

其实这样就很简单很简单啦，直接搬出似然函数，找出使得似然函数最大的参数值嘛，也就是最大化似然函数。

顾名思义，“然”是这样的意思，所以似然函数就是用来描述当前的模型参数取值的合理性，间接反映当前模型参数对手中数据集的解释程度，因此使得似然函数最大，意思就是使得模型参数取的最合理，使得手中数据集在模型的解释下变得合理。

你陪小夕度过了300天。因此你记录下了300段隐状态序列，记为

$Q = q_1 \ q_2 \ q_3 \ \dots \ q_T$ (其实 q_i 就是之前表示的 ω_i)，其中 $T=6$ (每天经历6个时间点，化妆-吃饭-聊天-自习-上课-要抱抱)。

同时对应着300段观测序列，记为：

$O = O_1 \ O_2 \ O_3 \ \dots \ O_T$ ，同样 $T=6$ 。

然后根据极大似然估计的思想，直接用样本集估计出HMM的各个参数啦。

所以HMM的各个参数应该这样估计：

$$\pi_i = \frac{\text{初始状态为 } q_i \text{ 的次数}}{\text{序列总数 (即 300)}}$$

$$\bar{a}_{ij} = \frac{Q \text{ 中从状态 } q_i \text{ 转移到 } q_j \text{ 的次数}}{Q \text{ 中所有从状态 } q_i \text{ 转移到另一状态 (包括 } q_j \text{ 自身) 的总数}}$$

$$\bar{b}_j(k) = \frac{Q \text{ 中从状态 } q_j \text{ 输出符号 } v_k \text{ 的次数}}{Q \text{ 到达 } q_j \text{ 的总次数}}$$

这样我们就得到了所有的模型参数 (向量 π 、矩阵 A 、矩阵 B)。

看, 由此, 我们轻而易举的把夕小瑶这个隐马尔可夫模型建立完成了。

有了这个模型, 我们就完全看透夕小瑶了! 所以, 下面开始得到本任务的最终目标——预测夕小瑶在耍小脾气当天的情绪状态序列 (隐状态序列)!

「看透你了!」

量化一下我们要做的事情: 在给定HMM模型 (即已知全部参数的HMM模型, 记为 μ) 和观测序列 O 情况下, 求最大概率的 (隐) 状态序列:

$$\hat{Q} = \arg \max_Q p(Q | O, \mu)$$

怎么计算呢? **Viterbi算法!**

这个算法有点绕, 直接贴出来, 如果直接看看不懂的话可以看后面小夕的解释哦。

(1) 初始化: $\delta_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N$

(2) 递推计算:

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(O_t), \quad 2 \leq t \leq T, \quad 1 \leq j \leq N$$

(3) 结束:

$$\hat{Q}_T = \operatorname{argmax}_{1 \leq i \leq N} [\delta_T(i)], \quad \hat{p}(\hat{Q}_T) = \max_{1 \leq i \leq N} \delta_T(i)$$

(4) 通过回溯得到路径（状态序列）

这个算法的思想就是设立一个小人 δ (读作delta)，这个小人从时刻1，一直走到时刻T。

这个小人的意义就是记录下自己在每个时间点t的每个隐状态j的概率（注意不是从全局的观测序列计算出的概率，是他自己从t=1的时刻一步步的观察每个时刻的观测值所得到的那一时刻的累积概率）

算法的第一步：初始化这个小人，利用已知的模型参数 π_i 和 $b_i(O_1)$ (即状态i下，发出观测值 O_1 的概率，其中 O_1 为t=1时刻的观测值)得到t=1时刻每个隐状态i的 δ 值。

第二步：这个小人在每个时刻t都会将每个隐状态里呆一会。在每个隐状态j里，它都会抬头看看此刻的观测值 O_t ，并分别假设自己处于前一时刻t-1的每个隐状态i中，并用前一时刻t-1假设的隐状态的累计概率值乘以前一时刻假设的隐状态转移到当前时刻的当前隐状态的转移概率，然后算出使得当前时刻的当前隐状态的总概率最大化的前一时刻的隐状态，这个最优隐状态记为m吧。这个前一时刻的最优隐状态的累积概率 $\delta_{t-1}(m)$ 乘以前一时刻m状态转移到当前时刻的当前j状态的转移概率，再乘以当前时刻t的当前状态j发出观测值 O_t 的概率，即为 $\delta_t(j)$ 。

第三步：小人 δ 走完最后一个时刻T的最后一个隐状态后，就可以从最后时刻的全部隐状态中挑出最后时刻，也就是全局的最大累积概率 δ_T 啦，记下这个最大概率对应的隐状态 Q_T

第四部：一步步的回溯呀，记下T时刻走到最优隐状态 Q_T 的前一时刻的最优隐状态（回顾一下第二步），得到前一时刻最优隐状态 Q_{T-1} 。然后同样，再往前回溯得到 Q_{T-2} ，一直回溯到 Q_1 ！然后 $Q_1 Q_2 Q_3 \dots Q_T$ 就是全局最优隐状态序列啦！也就是夕小瑶耍小脾气时的情绪状态序列！

在你猜出后，夕小瑶瞬间（被自己）感动哭了...（内心os：竟然花了这么大的功夫教你怎么猜我的心思...还不如直接告诉你呢嘤嘤嘤...）

蟹蟹你o(≥v≤)o



 微信支付



Transfer to 夕小瑶

声明：pdf仅供学习使用，一切版权归原创公众号所有；建议持续关注原创公众号获取最新文章，学习愉快！