隐马尔可夫模型

W.J.Z

1 隐马尔可夫模型

隐马尔可夫模型是关于时序的概率模型，描述由一个隐藏的马尔可夫链随机生成不可观测的状态随机序列，再由各个状态生成一个观测而产生观测随机序列的过程。隐藏的马尔可夫链随机生成的状态的序列成为状态序列；每一个状态生成一个观测由此产生的观测的随机序列称为观测序列，序列的每一个位置可以看成一个时刻。

马尔可夫模型由初始状态概率向量 π、状态转移概率矩阵 A 和观测概率矩阵 B 决定。π 和 A 决定状态序列，B 决定观测序列，因此隐马尔可夫模型 λ 可以用三元符号表示

$$\lambda = (A, B, \pi)$$

隐马尔可夫模型做了两个基本假设：(1)、在任意时刻 t 的状态只依赖于其前一时刻的状态，与其他时刻的状态及观测无关。(2)、任意时刻的观测只依赖该时刻的马尔可夫链的状态，与其他观测和状态无关。

隐马尔可夫模型由 3 个基本问题：

1. 概率计算问题：给定模型 $\lambda = (A, B, \pi)$ 和观测序列 $O = (o_1, o_2, \ldots, o_T)$，计算在模型下 λ 观测序列 O 出现的概率 $P(O|\lambda)$.

2. 学习问题: 已知模型 $\lambda = (A, B, \pi)$ 和观测序列 $O = (o_1, o_2, \ldots, o_T)$，使得在该模型下观测序列概率 $P(O|\lambda)$ 最大，即用极大似然估计的方法估计参数。

3. 预测问题：已知模型 $\lambda = (A, B, \pi)$ 和观测序列 $O = (o_1, o_2, \ldots, o_T)$，求对给定观测序列条件概率 $P(I|O)$ 最大的状态序列 $I = (i_1, i_2, \ldots, _t)$。即给定观测序列，求最有可能的对应的状态序列。

2 概率计算算法

2.1 前向算法

前向算法实际基于” 状态序列的路径结构” 递推计算 $P(O|\lambda)$ 的算法，关键在于局部计算前向概率，然后利用路径结构将前向概率递推到全局。

Algorithm 1: 观测序列概率的前向算法

Input: 隐马尔可夫模型 λ, 观测序列 O

Output: 观测序列概率 P(O|λ)

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

2 递推对 t=1,2,...,T-1

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

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

Algorithm 2: 观测序列概率的后向算法

Input: 隐马尔可夫模型 λ, 观测序列 O

Output: 观测序列概率 $P(O|\lambda)$

1

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

2 对 $t = T-1, T-2, \ldots, 1$

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

3

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

2.2 后向算法

3 学习算法

3.1 监督学习

转移概率 a_{ij} 的估计：

设样本中时刻 t 处于状态 i 时刻 $t+1$ 转移到状态 j 的频数为 A_{ij}，name 状态转移概率 a_{ij} 的估计为

$$a_{ij} = \frac{A_{ij}}{\sum_{j=1}^{N} A_{ij}}$$

观测概率 $b_j(K)$ ：

设样本中状态为 j 并观测为 K 的频数是 B_{jk}, 那么状态为 j 观测为 K 的概率 $b_j(k)$ 的估计是

$$b_j(k) = \frac{B_{jk}}{\sum_{k=1}^{M} B_{jk}}$$

初始状态概率 π_i 的估计为 S 个样本中初始状态为 q_i 的频数。

3.2 无监督学习

概率模型为

$$P\left(O|\lambda\right)=\sum_I P\left(O|I,\lambda\right)P\left(I,\lambda\right)$$

数据 O 为观测数据，隐数据 I 为状态序列数据, 它的参数学习由 EM 算法实现。

Algorithm 3: Baum-Welch 算法

Input: 观测数据 O

Output: 隐马尔可夫模型参数

初始化：对 $n=0$，选取 $a_{ij}^{(0)}$,$b_j(K)^(0)$,π_i^0，得到模型 $\lambda^{(0)}=\left(A^0,B^{(0)},\pi^{(0)}\right)$

递推：对 $n=1,2,\ldots,$

$$a_{ij}^{n+1)}=\frac{\sum_{t=1}^{T-1}P\left(O,i_t=i,i_{t+1}=j|\lambda\right)}{2}$$

$$b_j\left(k\right)=\frac{\sum_{t=1}^{T}P\left(O,i_t=j|\lambda\right)I\left(o_t=v_k\right)}{\sum_{t=1}^{T}P\left(O,i_t=j|\lambda\right)}$$

$$\pi_i=\frac{P\left(O\ i_1=i|\lambda\right)}{P\left(O|\lambda\right)}$$

终止，得到模型参数

$$\lambda^{(n+1)}=\left(A^{(n+1)},B^{(n+1)},\pi^{(n+1)}\right)$$

4 预测算法

维特比算法使用动态规划解隐马尔可夫模型问题，相等于求概率最大路径。

Algorithm 4: 维特比算法

Input: 模型 $\lambda = (A, B, \pi)$ 和观测 O

Output: 最优路径 $I^* = (i_1^*, \ldots, i_T^*)$

初始化

$$\delta_t(i) = \pi_i b_i(o_1), i = 1.2. \ldots, N$$

$$\psi_1(i) = 0, i = 1, 2, \ldots, N$$

递推：对 $t = 2, 3, \ldots, T$

$$\delta_t(i) = \underset{1 \leq j \leq N}{max} \left[\delta_{t-1}(j) a_{ji}\right] b_i(o_t) \;\; i = 1, 2, \ldots, N$$

$$\psi_t(i) = \arg \underset{1 \leq j \leq N}{max} \left[\delta_{t-1}(j) a_{ji}\right], i = 1, 2, \ldots, N$$

终止

$$P^* = \underset{1 \leq i \leq N}{max} \delta_T(i)$$

$$i_T^* = \arg \underset{i \leq i \leq N}{max} \left[\delta_T(i)\right]$$

最优路径回溯：对 $t = T-1, T-2, \ldots, 1$

$$i_t^* = \psi_{t+1}\left(i_{t+1}^*\right)$$

求得最优路径 $I^* = (i_1^*, i_2^*, \ldots, i_T^*)$