1. 评价问题：给定HMM模型M（即已知 A, B, π）和观测序列 $O=o_1 o_2 \cdots o_K$

计算模型M产生观测序列的概率（模型与观测序列匹配度）

经典应用：训练好模型 $M_1 \cdots M_c$（分别对应第1类…第c类）

针对观测序列O（测试数据），判别为哪一类？

方法1：前向算法（Forward recusion） → 数学归纳法

状态序列 $Q=q_1 \cdots q_K$

$q_k \in S$（状态集合）

$S=\{s_1, s_2 \cdots s_N\}$

$V=\{v_1, v_2 \cdots v_M\}$

（观测值集合）

① 初始化：令 $\alpha_1(i)=P(o_1\ q_1=s_i)=P(o_1|s_i)P(s_i)=b_i(o_1)\pi_i$ $(1\le i\le N)$

$$P(o_1)=\sum_j P(o_1\ q_1=s_j)=\sum_j \alpha_1(j)$$

② $\alpha_2(i)=P(o_1 o_2\ q_2=s_i)=\sum_j P(o_1 o_2\ q_1=s_j\ q_2=s_i)$

$$=\sum_j \underbrace{P(o_1|s_j)}_{b_j(o_1)} \underbrace{P(o_2|s_i)}_{b_i(o_2)} \underbrace{P(s_i|s_j)}_{a_{ji}} \underbrace{P(s_j)}_{\pi_j}$$ → 两个基本假设

$$=\sum_j b_j(o_1)\pi_j\ a_{ji}\ b_i(o_2)$$

$$=\sum_j \alpha_1(j)\ a_{ji}\ b_i(o_2)$$

$$\therefore P(o_1 o_2)=\sum_j P(o_1 o_2\ q_2=s_j)=\sum_j \alpha_2(j)$$

⋮

$$\alpha_{k+1}(i)=\sum_j \alpha_k(j)\ a_{ji}\ b_i(o_{k+1}) \quad 1\le j\le N,\ 1\le k\le K-1$$

③ 结束（终止）$P(o_1 o_2 \cdots o_K)=\sum_j P(o_1 o_2 \cdots o_K\ q_K=s_j)$

$$=\sum_j \alpha_K(j) \quad (1\le j\le N)$$

复杂度：$P(o_1 o_2 \cdots o_K)=\underbrace{\sum_j \underbrace{\alpha_K(j)}_{NK}}_{N^2K}$

方法2：后向算法（Backward recursion）→ 数列倒传

令 $\beta_k(i)=P(O_{k+1}O_{k+2}\cdots O_K \mid q_k=S_i)$

① 初始化 $\beta_K(i)=1 \qquad 1\le i\le N$

② $\beta_{K-1}(i)=P(O_K \mid q_{K-1}=S_i)=\sum_j P(O_K, q_K=S_j \mid q_{K-1}=S_i)$

$$=\sum_j \frac{P(O_K, q_K=S_j, q_{K-1}=S_i)}{P(q_{K-1}=S_i)}=\sum_j \frac{P(O_K \mid q_K=S_j, q_{K-1}=S_i)\,P(q_K=S_j, q_{K-1}=S_i)}{P(q_{K-1}=S_i)}$$

$$=\sum_j \frac{P(O_K \mid q_K=S_j)\,P(q_K=S_j \mid q_{K-1}=S_i)\,\cancel{P(q_{K-1}=S_i)}}{\cancel{P(q_{K-1}=S_i)}}=\sum_j b_j(O_K)\,a_{ij}\cdot 1$$

$$=\sum_j \beta_K(j)\,b_j(O_K)\,a_{ij}$$

$\beta_{K-2}(i)=P(O_{K-1}O_K \mid q_{K-2}=S_i)=\sum_j P(O_{K-1}O_K\, q_{K-1}=S_j \mid q_{K-2}=S_i)$

$$=\sum_j \frac{P(O_{K-1}O_K\, q_{K-1}=S_j\, q_{K-2}=S_i)}{P(q_{K-2}=S_i)}=\sum_j \frac{P(O_{K-1}O_K \mid q_{K-1}=S_j\, q_{K-2}=S_i)\,P(q_{K-1}=S_j, q_{K-2}=S_i)}{P(q_{K-2}=S_i)}$$

$$=\sum_j \frac{P(O_{K-1} \mid q_{K-1}=S_j)\,P(O_K \mid q_{K-1}=S_j)\,P(q_{K-1}=S_j \mid q_{K-2}=S_i)\,\cancel{P(q_{K-2}=S_i)}}{\cancel{P(q_{K-2}=S_i)}}$$

$$=\sum_j P(O_K \mid q_{K-1}=S_j)\,b_j(O_{K-1})\,a_{ij}=\sum_j \beta_{K-1}(j)\,b_j(O_{K-1})\,a_{ij}$$

$$\vdots$$

$$\beta_k(i)=P(O_{k+1}O_{k+2}\cdots O_K \mid q_k=S_i)=\sum_j \beta_{k+1}(j)\, b_j(O_{k+1})\,a_{ij}$$

③ 终止：$P(O_1O_2\cdots O_K)=\sum_j P(O_1O_2\cdots O_K, q_1=S_j)=\sum_j P(O_1\cdots O_K \mid q_1=S_j)\,P(q_1=S_j)$

$$=\sum_j P(O_2\cdots O_K \mid q_1=S_j)\,P(O_1 \mid q_1=S_j)\,P(q_1=S_j)$$

$$=\sum_j \beta_1(j)\,b_j(O_1)\,\pi_j$$

2. 解码问题（预测问题）：给定HMM模型 $M=(A,B,\pi)$ 和观测序列 $O=o_1 o_2 \cdots o_K$

计算产生此观测序列 $O$ 的最可能的状态序列 $Q$.

└→ 经典应用：在语音识别任务中，观测值为语音信号，隐藏状态为文字，

目标是根据观测信号推断最有可能的状态序列（即对应的文字）

△ 维特比(Viterbi)算法 —— 实际是用动态规划求概率最大路径.（见PPT第27页）



定义在时刻 $k$，状态为 $q_k$ 的（$q_k=S_j$）所有路径中概率最大值为，

$$\delta_k(j)=\max P(o_1 o_2 \cdots o_k, q_1 q_2 \cdots q_{k-1}, q_k=S_j)$$

$$=\max_i \underbrace{P(o_1 o_2 \cdots o_{k-1}, q_1 q_2 \cdots q_{k-1}=S_i)}_{\text{时刻}k-1\text{，状态为}q_{k-1}=S_i\text{所有路径中概率最大值}} \underbrace{P(q_k=S_j \mid q_{k-1}=S_i)}_{\text{状态}q_{k-1}=S_i\text{转移到状态}q_k=S_j\text{概率}} \underbrace{P(O_k \mid q_k=S_j)}_{\text{状态}q_k=S_j\text{时观测值为}O_k}$$

$$=\max_i \left[a_{ij}\cdot b_j(O_k)\cdot \max P(o_1 o_2 \cdots o_{k-1}, q_1 q_2 \cdots q_{k-1}=S_i)\right]$$

$$=\max_i \left[a_{ij}\, b_j(O_k)\, \delta_{k-1}(i)\right] \qquad 1\le i\le N,\quad 2\le k\le K$$

① 初始化 $\delta_1(i)=\max P(o_1, q_1=S_i)=\max P(o_1 \mid q_1=S_i)P(q_1=S_i)=\max\{\pi_i b_i(o_1)\}$

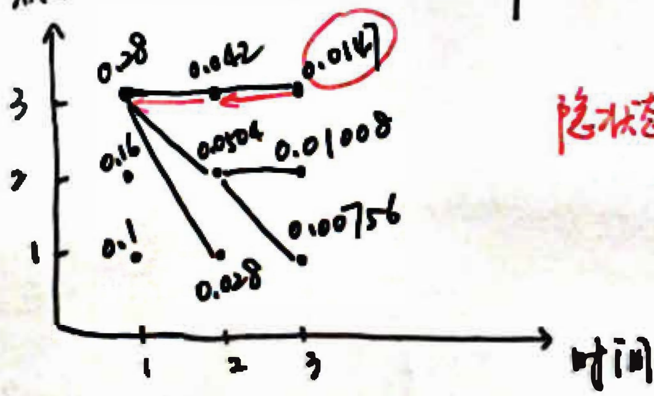
② ~~前后~~ 前向迭代 $\delta_k(i)=\max_j\left[a_{ji}\, b_i(O_k)\, \delta_{k-1}(j)\right]$

③ 终止 $\max_i \delta_K(i) \longrightarrow$ 时刻K，隐藏状态为 $S_{i^*}$，$i^*=\arg\max_i \delta_K(i)$

K-1
⋮
1

回溯路径

举例




隐状态序列.{3.3.3}

$\boxed{\text{几个结论}}$

$$\alpha_k(i) = P(O_1 O_2 \cdots O_k, q_k = S_i) \qquad 1 \le i \le N$$

$$\beta_k(i) = P(O_{k+1} O_{k+2} \cdots O_K \mid q_k = S_i)$$

$$P(O_1 O_2 \cdots O_K, q_k = S_i) = P(\underbrace{O_1 O_2 \cdots O_k}, \underbrace{O_{k+1}, \cdots O_K} \mid q_k = S_i)\, P(q_k = S_i)$$

给定 $q_k$, 则 $O_1 \cdots O_k$, $O_{k+1} \cdots O_K$ 是条件独立的

$\Downarrow$

$$\therefore \quad P(O_1 O_2 \cdots O_k \mid q_k = S_i)\, P(O_{k+1} \cdots O_K \mid q_k = S_i)$$

因此 $P(O, q_k = S_i) = P(O_1 O_2 \cdots O_K, q_k = S_i)$

$$= P(O_1 O_2 \cdots O_k \mid q_k = S_i)\, \underbrace{P(O_{k+1} \cdots O_K \mid q_k = S_i)}_{\beta_k(i)}\, P(q_k = S_i)$$

$\alpha_k(i)$

$$\therefore \boxed{P(O, q_k = S_i) = \alpha_k(i)\beta_k(i)}$$ 结论1.

△ 当给定HMM模型 $M = (A, B, \pi)$, 已知观测序列 $O$, 令 $\gamma_k(i)$ 表示在时刻k状态为 $S_i$ 的概率, 即

结论2

$$\gamma_k(i) = P(q_k = S_i \mid O) = \frac{P(O, q_k = S_i)}{\underbrace{P(O)}_{= \sum_i P(O, q_k = S_i)}} = \boxed{\frac{\alpha_k(i)\beta_k(i)}{\sum_i \alpha_k(i)\beta_k(i)}}$$

$$\beta_k(i) = P(O_{k+1} O_{k+2} \cdots O_K \mid q_k = S_i) = \sum_j P(O_{k+1} \cdots O_K, q_{k+1} = S_j \mid q_k = S_i)$$

$$= \sum_j \beta_{k+1}(j) b_j(O_{k+1}) a_{ij}$$

$$\therefore P(O_{k+1} \cdots O_K, q_{k+1} = S_j \mid q_k = S_i) = \beta_{k+1}(j) b_j(O_{k+1}) a_{ij}$$

因此，$P(O, q_k = S_i, q_{k+1} = S_j) = P(O_1 \cdots O_k, O_{k+1}, O_{k+2} \cdots O_K, q_k = S_i, q_{k+1} = S_j)$

$$= P(O_1 \cdots O_k, O_{k+1} \cdots O_K, q_{k+1} = S_j \mid q_k = S_i) P(q_k = S_i)$$

给定 $q_k = S_i$，则 $O_1 \cdots O_k$ 与 $O_{k+1} \cdots O_K, q_{k+1} = S_j$ 是条件独立的

$$= P(O_1 \cdots O_k \mid q_k = S_i) P(O_{k+1} \cdots O_K, q_{k+1} = S_j \mid q_k = S_i) P(q_k = S_i)$$

$$= \underbrace{P(O_1 \cdots O_k, q_k = S_i)}_{\alpha_k(i)} \underbrace{P(O_{k+1} \cdots O_K, q_{k+1} = S_j \mid q_k = S_i)}_{a_{ij} b_j(O_{k+1}) \beta_{k+1}(j)}$$

$$\therefore \boxed{P(O, q_k = S_i, q_{k+1} = S_j) = \alpha_k(i) a_{ij} b_j(O_{k+1}) \beta_{k+1}(j)}$$

▲ 当给定HMM模型 $M = (A, B, \pi)$，已知观测序列 $O$，令 $\xi_k(i,j)$ 表示在时刻 $k$ 处于 $S_i$ 状态、时刻 $k+1$ 处于 $S_j$ 状态的概率，即

$$\xi_k(i,j) = P(q_k = S_i, q_{k+1} = S_j \mid O) = \frac{P(O, q_k = S_i, q_{k+1} = S_j)}{P(O)}$$

$$= \sum_i \sum_j P(O, q_k = S_i, q_{k+1} = S_j)$$

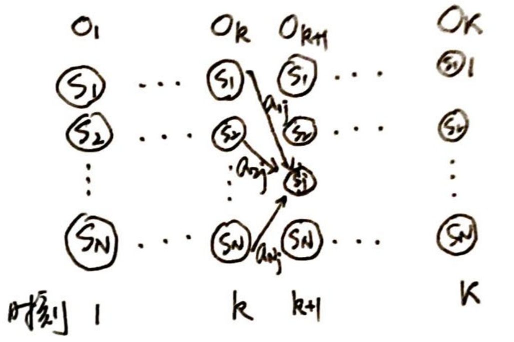
$$= \frac{\alpha_k(i) a_{ij} b_j(O_{k+1}) \beta_{k+1}(j)}{\sum_i \sum_j \alpha_k(i) a_{ij} b_j(O_{k+1}) \beta_{k+1}(j)}$$

结论3

我们总结一下:

$$\gamma_k(i) = P(q_k = s_i \mid O_1 O_2 \cdots O_K)$$

$$\xi_k(ij) = P(q_k = s_i \quad q_{k+1} = s_j \mid O_1 O_2 \cdots O_K)$$




栅格网(见PPT24页)

** 划重点 (PPT第37页)

1. Expected number of transitions from state $s_i$ to state $s_j$: $\sum_{k=1}^{K} \xi_k(i,j)$

   从状态 $s_i$ 到状态 $s_j$ 的转移个数.

2. Expected number of transitions out of state $s_i$: $\sum_{k=1}^{K} \gamma_k(i)$

   从状态 $s_i$ 出发的转移个数

3. Expected number of times observation $v_m$ occurs in state $s_i$: $\sum_{k=1,\ 且 O_k=v_m}^{K} \gamma_k(i)$

   在状态 $s_i$ 观察到的观测值为 $v_m$ 的次数

4. Expected frequency in state $s_i$ at time $k$: $\gamma_k(i) \longrightarrow \sum_{k=1}^{K} \gamma_k(i)$

   在时刻 k 处于状态 $s_i$ 的概率(频率). 　　整段时间处于状态 $s_i$

** 从定义的角度看 $a_{ij}$ $b_i(v_m)$ 和 $\pi_i$ (PPT第38页)

$$a_{ij} = \frac{\text{从状态 i 转移到状态 j 的个数}}{\text{所有从状态 i 出发的个数}} = \frac{\sum_{k=1}^{K} \xi_k(i,j)}{\sum_{k=1}^{K} \gamma_k(i)}$$

$$b_i(v_m) = \frac{\text{在状态 } s_i \text{ 观察到的观测值为 } v_m \text{ 的次数}}{\text{整段时间处于状态 } s_i \text{ 的次数}} = \frac{\sum_{k=1,\ O_k=v_m}^{K} \gamma_k(i)}{\sum_{k=1}^{K} \gamma_k(i)}$$

$$\pi_i = \underbrace{\text{在时刻 } k=1}_{\text{初始时刻}}\text{, 处于状态 } s_i \text{ 的概率} = \gamma_1(i)$$