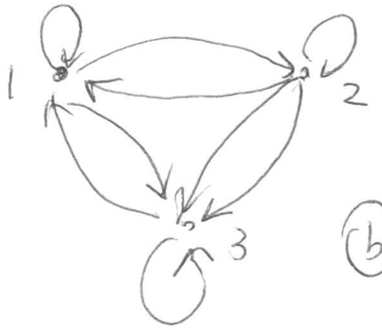
隐含马尔可夫模型
Hidden Markov Models

① Markov 模型



ⓐ 状态 $\{S_1, S_2, S_i\}$ $i=1\sim N$

ⓑ $a_{ij}=P(q_{t+1}=S_j|q_t=S_i)$
转移概率（矩阵）

$P=\{a_{ij}\}_{i=1\sim N, j=1\sim N}$

ⓒ 初始概率 π_q

② 隐含 Markov 模型

Observation

$$P(O_j|q_i)=b_{q_i}(O_j)$$

q: $q_1\ q_2 \cdots q_N$

O: $O_1\ O_2 \cdots O_N$

语音识别中 $P(O_i|q_i)$ 可取为 Gaussian Mixture

设 $\lambda=\{A, B, \pi\}$

隐马尔可夫模型的三个问题

① Given $O=O_1O_2\cdots O_T$, and Model $\lambda(A,B,\pi)$，计算 $P(O|\lambda)$

$$P(O|\lambda)=\sum_{q_T}\cdots\sum_{q_2}\sum_{q_1}\pi_{q_1}b_{q_1}(O_1)a_{q_1q_2}b_{q_2}(O_2)\cdots a_{q_{T-1}q_T}b_{q_T}(O_T)$$

② Given $O = O_1 O_2 O_3 \cdots O_T$, Given $\lambda(A, B, \pi)$，选择 Q $q_1 q_2 \cdots q_T$，使之在某种意义下最优

$$\underset{q_1 q_2 \cdots q_T}{\text{Maximize}}\ \pi_{q_1} b_{q_1}(O_1) a_{q_1 q_2} \cdots a_{q_{T-1} q_T} b_{q_T}(O_T)$$

③ Given $O = O_1 O_2 \cdots O_T$，怎么估计参数 $\lambda(A, B, \pi)$？

问题1解法

1. ① 定义 $\alpha_t(i) = P(O_1 O_2 \cdots O_t, q_t = S_i \mid \lambda)$

2. $\alpha_1(i) = P(O_1, q_1 = S_i) = \pi_i b_i(O_1)$

3. $\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1})$

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

另一种解法

定义：ⓐ $\beta_t(i) = P(O_{t+1} O_{t+2} \cdots O_T \mid q_t = S_i, \lambda)$

ⓑ $\beta_T(i) = 1 \quad (1 \le i \le N)$

ⓒ $\beta_t(i) = \sum_{j} a_{ij} \beta_{t+1}(j) b_j(O_{t+1})$

$t = T-1, T-2 \cdots 1$

ⓓ $P(O \mid \lambda) = \sum_{j=1}^{N} \pi_j b_j(O_1) \beta_1(j)$

问题2 解法

求 $q_1 q_2 \cdots q_T$，使

$$\pi_{q_1} b_{q_1}(O_1) a_{q_1 q_2} b_{q_2}(O_2) \cdots a_{q_{T-1} q_T} b_{q_T}(O_T)$$

最大

~~定义~~: Viterbi Algorithm

定义 $$\delta_t(i) = \max_{q_1 q_2 \cdots q_{t-1}} P(q_1, q_2 \cdots q_t = i\ O_1 O_2 \cdots O_t)$$

$$\delta_{t+1}(j) = \max_i [\delta_t(i) a_{ij}] b_j(O_{t+1})$$

1、 $\delta_1(i) = \pi_i b_i(O_1)$

$\psi_1(i) = 0$

2、 $$\delta_t(j) = \max_i [\delta_{t-1}(i) a_{ij}] b_j(O_t)$$

$$\psi_t(j) = \arg\max_{1 \le i \le N} [\delta_{t-1}(i) a_{ij}]$$

3、 $$p^* = \max_{1 \le i \le N} [\delta_T(i)]$$

$$q_T^* = \arg\max_{1 \le i \le N} [\delta_T(i)]$$

$$q_t^* = \psi_{t+1}(q_{t+1}^*)$$

问题3求解

$$\xi_t(i,j) = P(q_t = S_i, q_{t+1} = S_j | O, \lambda)$$

$$\xi_t(i,j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}$$

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j)$$

$\sum_{t=1}^{T-1} \gamma_t(i) =$ 从 ~~S_i~~ 到达 S_i 的次数均值

$\sum_{t=1}^{T-1} \xi_t(i,j) =$ 从 S_i 到 S_j 次数均值

$$\overline{\pi_i} = \gamma_1(i)$$

$$\overline{a_{ij}} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$\overline{b_j(k)} = \frac{\sum_{\substack{t=1 \\ O_t = v_k}}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)}$$

Baum-Welch 算法 (EM)

设观测序列 $\{O_1, O_2 \cdots O_S\}$

$O = (o_1\, o_2 \cdots o_T)$, $I = (i_1\, i_2 \cdots i_T)$

$(O, I) = (o_1\, o_2 \cdots o_T\ i_1\, i_2 \cdots i_T)$

λ为待求参数，包括 π_i, a_{ij}, $b_j(O)$等。

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

EM算法

$$P(O, I|\lambda) = \pi_{i_1} b_{i_1}(o_1) a_{i_1 i_2} b_{i_2}(o_2) \cdots a_{i_{T-1} i_T} b_{i_T}(o_T)$$

$$Q(\lambda, \bar\lambda) = \sum_I \log P(O, I|\lambda) P(O, I|\bar\lambda)$$

$$= \sum_I \log \pi_{i_1} P(O, I|\bar\lambda) + \sum_I \left(\sum_{t=1}^{T-1} \log a_{i_t i_{t+1}}\right) P(O, I|\bar\lambda) + \sum_I \left(\sum_{t=1}^{T} \log b_{i_t}(o_t)\right) P(O, I|\bar\lambda)$$

极大化 Q

$$\pi_i = \frac{P(O, i_1 = i|\bar\lambda)}{P(O|\bar\lambda)}$$

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

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