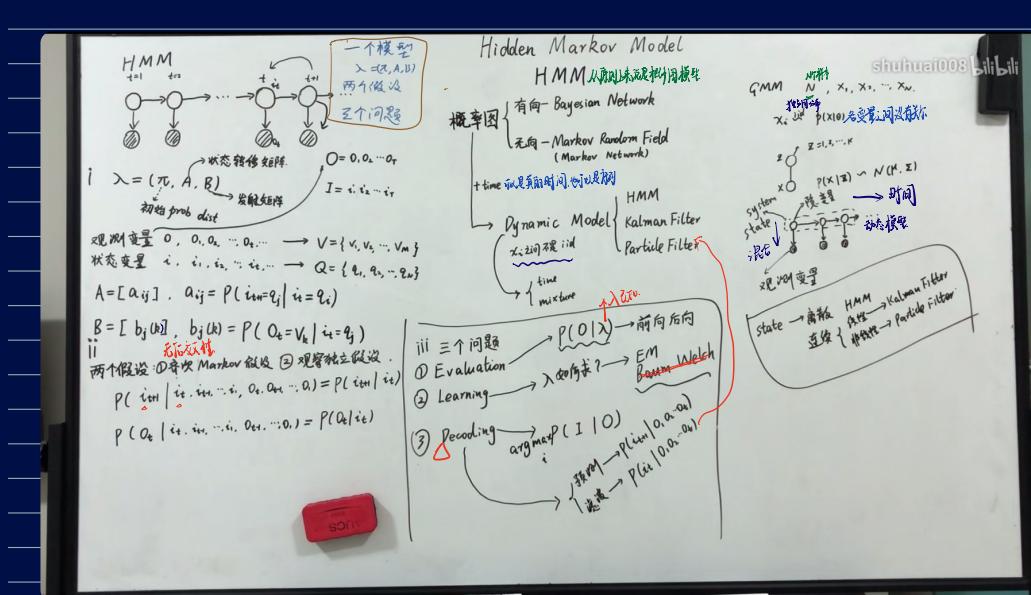


HMM在之前的NLP中使用了很多

补充一下概念圈模型



概率派 → 统计机器学习

十一

- ① 定义出一定模型
 - ② 策略, 即评估准则
 - ③ 求解 GDS, 特点进阶法

贝叶斯派 → 极端

推断 \rightarrow $P(Z|X)$

同上

↓
积分问题

HMM

$I = i_1, i_2, \dots, i_t \rightarrow$ 状态序列, $Q = \{q_{ij} | i, j \in V_I\}$ \rightarrow 状态值集合
 $O = o_1, o_2, \dots, o_t \rightarrow$ 观测序列, $V = \{v_1, v_2, \dots, v_m\}$ \rightarrow 观测值集合

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

π: 初始概率分布
 $A: [A_{ij}] \rightarrow$ 转移矩阵, $A_{ij} = P(i_{t+1} = q_j | i_t = q_i)$
 $B: [b_j(k)] \rightarrow$ 发射矩阵, $b_j(k) = P(O_k = v_k | i_t = q_j)$

(ii) 两个假设: ① 次 Markov ② 观测独立
 $P(i_{t+1} | i_1, i_2, \dots, i_t, o_1, \dots, o_t) = P(i_{t+1} | i_t)$
 $P(o_t | i_1, i_2, \dots, i_t, o_1, \dots, o_{t-1}) = P(o_t | i_t)$

(iii) 三个问题

- Evaluation: Given λ , find $P(O|\lambda)$ Bayesian Welch EM
- Learning: $\lambda_{MLE} = \arg \max_{\lambda} P(O|\lambda)$ Viterbi
- Decoding: $\hat{i} = \arg \max_i P(i|O, \lambda)$

Hidden Markov Model

HMM 问题:

Evaluation: Given λ , find $P(O|\lambda)$

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

$$P(I|\lambda) = P(i_1, i_2, \dots, i_t, \lambda) = P(i_1 | \lambda) \cdot P(i_2 | i_1, \lambda) \cdot P(i_3 | i_2, i_3, \lambda) = \pi(i_1, \lambda) \prod_{t=2}^T \alpha_{i_{t-1}, i_t}$$

$$P(O|I, \lambda) = \prod_{k=1}^T b_{i_k}(o_k)$$

$$\therefore P(O|\lambda) = \sum_I \pi(i_1, \lambda) \prod_{t=2}^T \alpha_{i_{t-1}, i_t} \prod_{k=1}^T b_{i_k}(o_k) = \boxed{\sum_{i_1, i_2, \dots, i_T} \prod_{t=2}^T \alpha_{i_{t-1}, i_t} \prod_{k=1}^T b_{i_k}(o_k)}$$

相同步序求和

Forward Algorithm

记: $\alpha_t(i) = P(O_1, \dots, O_t, i_t = q_i | \lambda)$
 $\alpha_T(i) = P(O, i_T = q_i | \lambda)$

$$P(O|\lambda) = \sum_{i_1}^N P(O_1, i_1 = q_{i_1} | \lambda) \quad \text{相同步序求和}$$

$$= \sum_{i_1}^N \alpha_1(i_1) \cdot P(O_2 | O_1, i_1 = q_{i_1}, i_2 = q_{i_2}, \lambda) \cdot P(O_3 | O_2, i_2 = q_{i_2}, i_3 = q_{i_3}, \lambda) \cdots \alpha_T(i_T)$$

$$= \sum_{i_1}^N \alpha_1(i_1) \cdot P(O_2 | i_1 = q_{i_1}) \cdot P(O_3 | i_2 = q_{i_2}) \cdots P(O_T | i_{T-1} = q_{i_{T-1}})$$

$$= \sum_{i_1}^N P(O_{1:T} | i_1 = q_{i_1}) \cdot P(O_{2:T} | i_2 = q_{i_2}) \cdots P(O_{T:T} | i_{T-1} = q_{i_{T-1}})$$

$$\boxed{P(O_{1:T} | i_1 = q_{i_1}) \rightarrow P(O_{1:T} | i_1 = q_{i_1}) \cdot P(O_{2:T} | i_2 = q_{i_2}) \cdots P(O_{T:T} | i_{T-1} = q_{i_{T-1}}) \rightarrow P(O_{1:T} | i_1 = q_{i_1}) \cdot P(O_{2:T} | i_2 = q_{i_2}) \cdots P(O_{T:T} | i_{T-1} = q_{i_{T-1}})}$$

$$= \sum_{i_1}^N P(O_{1:T} | i_1 = q_{i_1}) \cdot P(O_{2:T} | i_2 = q_{i_2}) \cdots P(O_{T:T} | i_{T-1} = q_{i_{T-1}}) \alpha_T(i_T)$$

$$= \sum_{i_1}^N P(O_{1:T} | i_1 = q_{i_1}) \cdot P(O_{2:T} | i_2 = q_{i_2}) \cdots P(O_{T:T} | i_{T-1} = q_{i_{T-1}}) \alpha_T(i_T)$$

$$= \sum_{i_1}^N P(O_{1:T} | i_1 = q_{i_1}) \cdot P(O_{2:T} | i_2 = q_{i_2}) \cdots P(O_{T:T} | i_{T-1} = q_{i_{T-1}}) \alpha_T(i_T)$$

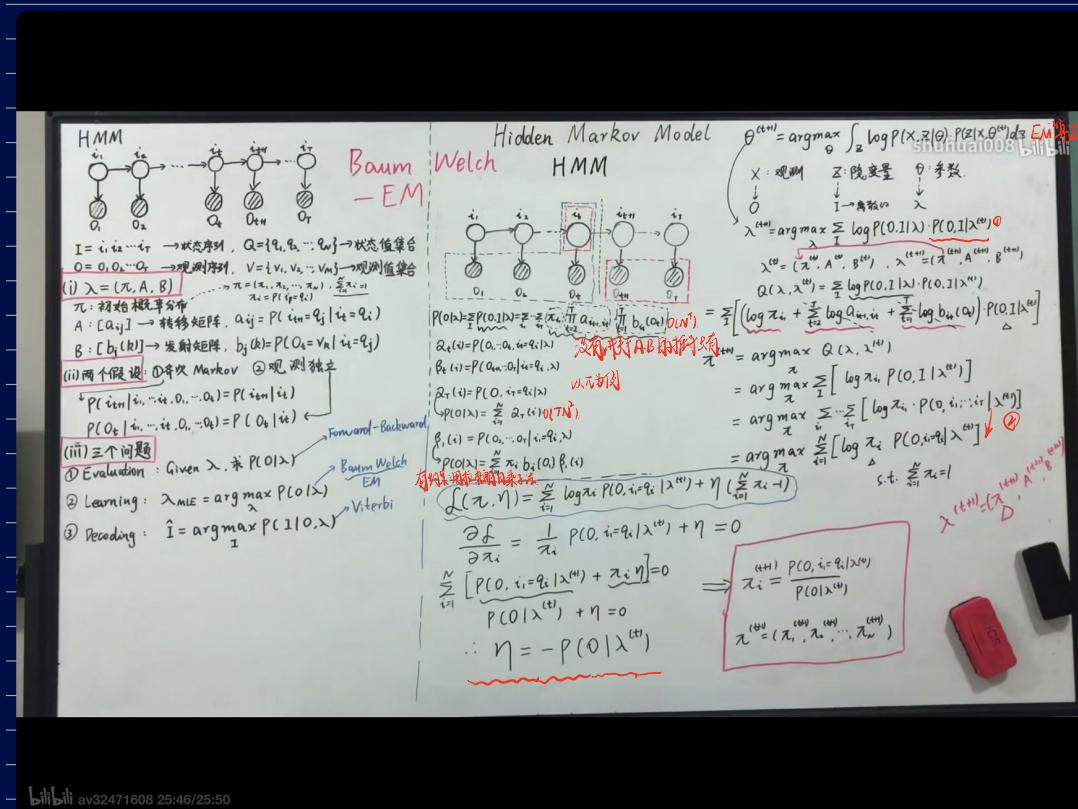
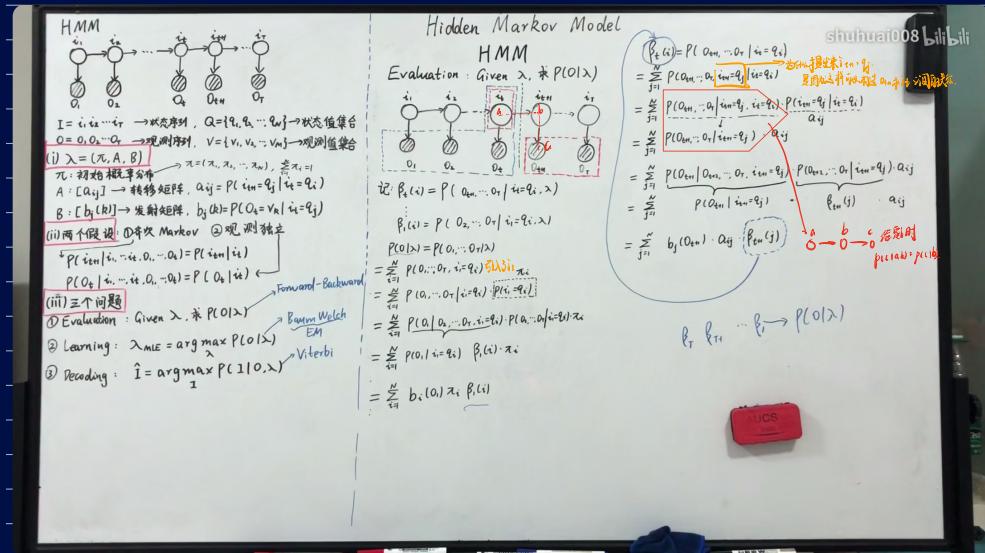
shuhua008 bilibili

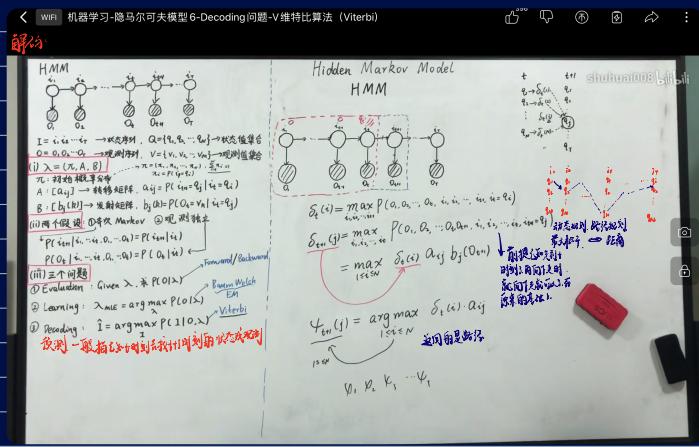
△在泥法上有点问题。
 $\Delta T = P(0, \bar{t} = q|1\lambda)?$

（一小部分的预测）
雨 (P�|入) 在给定参数的情况下预测的天气
预测降雨概率是很有办法的，所以
根据雨的预测去预测天气。

$$\begin{aligned}
 &= \sum_{j_1, j_2} P(O_{1+q_1} | O_1, \dots, O_{q_1-1}, i_1=j_1, i_2=j_2, \lambda) \\
 &\quad P(O_1, \dots, O_{q_1-1}, i_1=j_1, i_2=j_2, \lambda) \\
 &= \frac{1}{\binom{q_1}{j_1}} P(O_{1+q_1} | O_1, \dots, O_{q_1-1}, q_1=j_1, \lambda) P(O_1, \dots, O_{q_1-1}, q_1=j_1, \lambda) \\
 &= \frac{1}{\binom{q_1}{j_1}} P(O_{1+q_1} | O_1, \dots, O_{q_1-1}, q_1=j_1, \lambda) P(O_1, \dots, O_{q_1-1}, q_1=j_1, \lambda) \\
 &\quad \cdot \underbrace{P(O_{1+q_1} | O_1, \dots, O_{q_1-1}, q_1=j_1, \lambda)}_{P(O_{1+q_1} | O_1, \dots, O_{q_1-1}, q_1=j_1, \lambda)} \\
 &= P(O_1, \dots, O_{q_1-1}, i_1=j_1, \lambda) P(O_{1+q_1} | O_1, \dots, O_{q_1-1}, i_1=j_1, \lambda)
 \end{aligned}$$

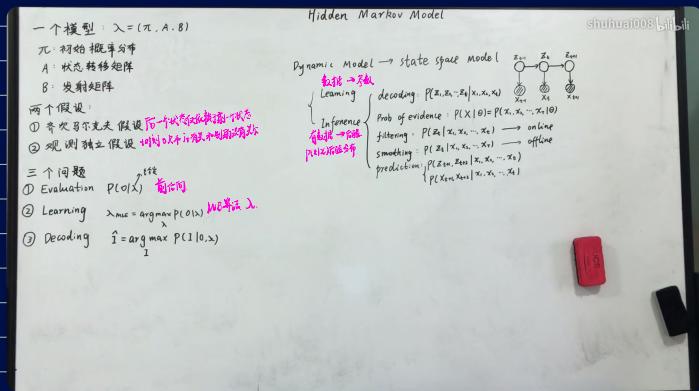
前向传播通过定义 $\delta_t(i)$ $\delta_{t+1}(j) = p_1(o_1, \dots, o_t, i; q_1 | \lambda)$





HMM: mixture + time

混合 HMM $\xrightarrow{\text{1. 从观察角度}} \text{混合高斯}$ $\xrightarrow{\text{2. 从参数角度}} \text{发射矩阵}$ $\xrightarrow{\text{3. 规则}} \text{HMM} \rightarrow \text{HMM+time}$



没听完打草稿再听