

[click here for link](#)

概率图模型

inference $\rightarrow P(Z|X) \rightarrow$ 积分问题 (MCMC)

GMM: 样本之间是独立同分布

HMM: Dynamic Model

y: System state 隐变量

- state 离散: HMM
- state 线性: Kalman Filter
- state 非线性: Particle Filter

$$\lambda = (\underbrace{\pi}_{\text{初始prob dist}}, \underbrace{A}_{\text{状态转移矩阵}}, \underbrace{B}_{\text{发射矩阵emission}})$$

状态变量 $i: i_1, i_2, \dots, i_t \dots \rightarrow Q = q_1, q_2, \dots, q_m$

观测变量 $o: o_1, o_2, \dots, o_t \dots \rightarrow V = v_1, v_2, \dots, v_m$

$$A = [a_{ij}], a_{ij} = P(i_{t+1} = q_i | i_t = q_j)$$
$$B = [b_{jk}], b_{jk} = P(Q_t = v_k | i_t = q_j)$$

transition 和 emission probability 是 independent

两个假设:

- 齐次Markov 假设

$$P(i_{t+1} | i_t, t_{t-1}, \dots, t_1, o_t, o_{t-1}, \dots, o_1) = p(i_{t+1} | i_t)$$

- 观察独立假设

$$P(o_t | i_t, t_{t-1}, \dots, t_1, o_t, o_{t-1}, \dots, o_1) = p(o_t | i_t)$$

三个问题:

1. Evaluation: $P(O|\lambda) \Rightarrow$ 前向后向 Forward-backward
2. learning $\lambda = \arg \max P(O|\lambda)$ EM algorithm \ baum welch
3. Decoding $\lambda = \arg \max_i P(I|O)$
 1. 预测: $P(i_{t+1}|o_1, o_2, \dots, o_t)$
 2. 滤波: $P(i_t|o_1, o_2, \dots, o_t)$

HMM-Evaluation

Give λ , 求 $P(O|\lambda)$

$$P(O|\lambda) = \sum_1 P(I, O|\lambda) = \sum_1 P(O|I, \lambda) \cdot P(I|\lambda)$$

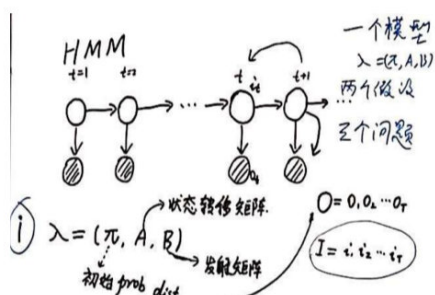
$$P(I|\lambda) = P(i_1, i_2, \dots, i_T|\lambda) = \underbrace{P(i_T|i_1, i_2, \dots, i_{T-1}, \lambda)}_{P(i_T|i_{T-1})=a_{i_{T-1}, i_T}} P(i_1, i_2, \dots, i_{T-1}|\lambda) = a_{i_{T-1}, i_T} \cdot a_{i_{T-2}, i_{T-1}} \cdots a_{i_1, i_2} \cdot \pi(i_1)$$

π 是初始分布

$$= \pi(a_i) \cdot \prod_{t=2}^T a_{i_{t-1}, i_t}$$

$$p(O|I, \lambda) = \prod_{t=1}^T b_{i_t}(O_t)$$

PPT



② 两个假设: ① 马尔可夫假设 ② 观察独立假设
 $P(i_{t+1} | i_t, i_{t-1}, \dots, i_1, o_1, o_2, \dots, o_t) = P(i_{t+1} | i_t)$
 $P(o_t | i_t, i_{t-1}, \dots, i_1, o_1, o_2, \dots, o_{t-1}) = P(o_t | i_t)$

Hidden Markov Model HMM

概率图: 有向-Bayesian Network
无向-Markov Random Field (Markov Network)

Dynamic Model: HMM, Kalman Filter, Particle Filter
 x_i 不一定是 iid
 time mixture

③ 三个问题
 ① Evaluation $P(O|\lambda) \rightarrow$ 前向后向
 ② Learning $\lambda = \arg \max P(O|\lambda) \rightarrow$ EM, Baum Welch
 ③ Decoding $I = \arg \max_i P(I|O) \rightarrow$ 预测 $\rightarrow P(i_{t+1}|o_1, o_2, \dots, o_t)$
 滤波 $\rightarrow P(i_t|o_1, o_2, \dots, o_t)$

