

(a) 问题:  $X \sim P(x|\theta)$ , 已知样本  $X_{1:n}$  iid  $X$

当新来一个  $X = t$  来自总体  $X$ , 求其出现概率。

后续问题均为 (a) 是特例

解: Training:  $P(\theta | X_{1:n}) = \eta P(X_{1:n}|\theta) P(\theta)$  ① 可由训练得出

Prediction:  $P(t | X_{1:n}) = \int P(t | X_{1:n}, \theta) P(\theta | X_{1:n}) d\theta$

② 当  $\theta$  已知时,  $X_{1:n}$  对求  $P(t)$  无帮助,

$\therefore P(t | X_{1:n}, \theta) = P(t | \theta)$  可以求出,

(b)  $X \sim \text{Bern}(M)$ , 即  $P(X=H)=M$ ,  $P(X=T)=1-M$

而参数  $M \sim \text{Beta}(M | a_0, b_0)$ ,  $a_0=50$ ,  $b_0=50$

样本  $X_{1:5} = [H, H, T, T, T]$  iid  $X$

解: ①  $M$  的后验分布为  $P(M | X_{1:5}) = \text{Beta}(M | 52, 53)$

②  $\therefore$  对  $X$  而言  $M_{ML} = \frac{52}{52+53} = \frac{52}{105}$

即新来一个  $X_6 = t$ ,  $P(t=H) = M_{ML}$

(c)  $X \sim N(\mu, \Sigma)$ , 参数  $M$  的先验分布  $M \sim N(\mu_0, \Sigma_0)$

参数  $M$  的先验分布

已知  $\mu \sim N(\mu_0, \Sigma_0)$  且  $X_i | \mu \sim N(\mu, \Sigma)$

现在观测到  $X_{1:n}$  iid  $X$ 。

① 求  $P(\mu | X_{1:n})$       ② 求  $P(X=t | X_{1:n})$

① Training:  $P(\mu | X_{1:n}) \propto P(X_{1:n} | \mu) P(\mu)$

由于高斯分布的特性, 此式可简化为

$$\begin{cases} \Sigma_{MAP}^{-1} = \Sigma_0^{-1} + \Sigma_{1:n}^{-1} \\ \Sigma_{MAP}^{-1} \mu_{MAP} = \Sigma_0^{-1} \mu_0 + \Sigma_{1:n}^{-1} \mu_{1:n} \end{cases}$$

其中  $P(\mu | X_{1:n}) \sim N(\mu | \mu_{MAP}, \Sigma_{MAP})$

$$P(X_{1:n} | \mu) \sim N(\mu | \mu_{1:n}, \Sigma_{1:n})$$

②  $P(X=t | X_{1:n}) \sim N(t | \mu_{MAP}, \Sigma_{MAP} + n \Sigma)$

(d)  $Y = AX + \eta$ ,  $\eta \sim N(0, Q)$  已知,  $X$  有先验分布  $X \sim N(\mu_0, \Sigma_0)$   
现在观测  $Y_{1:n}$ ,  $n=1$ 。

① 求  $P(X | Y_{1:n})$  这一后验分布      ② 求  $P(Y | Y_{1:n})$   
与极大似然估计分布

① 由  $Y = AX + \eta$  可设  $X = HY + v$ ,  $v \in N(0, L)$

$$P(x|y) = P(y|x) P(x)$$

给定  $X$ ,  $P(Y|X) \sim N(y|AX, Q)$ , 而又已知  $X \sim N(\mu_0, \Sigma_0)$

$$\begin{cases} \Sigma_{MAP}^{-1} = Q^{-1} + \Sigma_0^{-1} \\ \Sigma_{MAP}^{-1} \mu_{MAP} = Q^{-1} AX + \Sigma_0^{-1} \mu_0 \end{cases}$$

而取  $\Sigma_0^{-1} = 0$  可得  $m_L$ :

$$\begin{cases} L = (A^T Q^{-1} A)^{-1} = \Sigma_{ML} \\ Hy = (A^T Q^{-1} A)^{-1} A^T Q^{-1} y = m_{ML} \end{cases}$$

$$(2) \quad Y = AX + n$$

$$P(Y|Y:n) = N(y | A\mu_X, A\Sigma_X A^T + Q)$$

$$\text{其中 } \mu_X, \Sigma_X = \mu_{MAP}, \Sigma_{MAP}$$