题目1

(1) $p(z_{i1},\cdots z_{in}|p(w_i))=\prod_{k=1}^{n}p(z_{ik}|p(w_i))$

$p(z_{ik}=1|p(w_i))=p(w_i)\quad p(z_{ik}=0|p(w_i))=1-p(w_i)$

$\Downarrow$

$p(z_{ik}|p(w_i))=p(w_i)^{z_{ik}}[1-p(w_i)]^{1-z_{ik}}$

$\therefore p(z_{i1}\cdots z_{in}|p(w_i))=\prod_{k=1}^{n}p(w_i)^{z_{ik}}[1-p(w_i)]^{1-z_{ik}}$

(2) $l=\ln p(z_{i1}\cdots z_{in}|p(w_i))$

$=\sum_{k=1}^{n}z_{ik}\log p+(1-z_{ik})\log(1-p)$

$\frac{\partial l}{\partial p}=\frac{1}{p}\sum_{k=1}^{n}z_{ik}+\frac{1}{p-1}\sum_{k=1}^{n}(1-z_{ik})=0$

$(p-1)\sum_{k=1}^{n}z_{ik}+p\sum_{k=1}^{n}(1-z_{ik})=0$

$np=\sum_{k=1}^{n}z_{ik}$

$p=\frac{1}{n}\sum_{k=1}^{n}z_{ik}$

极大似然估计的结果就是用频率近似概率

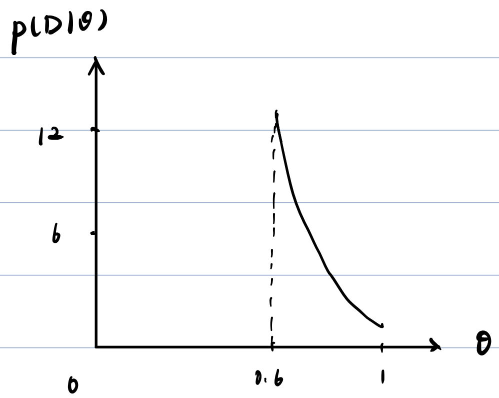
题目 2

(1) $p(x_1 \cdots x_n) = \prod_{i=1}^{n} \frac{1}{\theta} = \frac{1}{\theta^n} \qquad 0 \le \max[D] \le \theta$

$\therefore \max_{\theta} p(x_1 \cdots x_n) = \frac{1}{\max[D]^n}$

$\arg\max_{\theta} p(x_1 \cdots x_n) = \max[D]$

(2) $p(D|\theta) = \begin{cases} \frac{1}{\theta^5} & \max(x_k) \le \theta \le 1 \\ 0 & 0 \le \theta < \max(x_k) \end{cases}$



题目 3

$D_{KL}(p_2 \| p_1) = \int p_2(x) \ln \frac{p_2(x)}{p_1(x)} dx$

$= \int p_2(x) \ln p_2(x) dx - \int p_2(x) \ln p_1(x) dx$

$p_1(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(- \frac{(x-\mu)^t \Sigma^{-1} (x-\mu)}{2} \right)$

$D_{KL}(p_2 \| p_1) = c + \int \frac{1}{2} p_2(x) \left[d \ln 2\pi + \ln|\Sigma| + (x-\mu)^t \Sigma^{-1} (x-\mu) \right] dx$

$\frac{\partial D_{KL}}{\partial \mu} = \frac{\partial}{\partial \mu} \left(\int \frac{1}{2} p_2(x) (x-\mu)^t \Sigma^{-1} (x-\mu) \, dx \right) = -\int p_2(x) \Sigma^{-1} (x-\mu) dx = 0$

$$\Sigma^{-1}\int p_2(x)(x-\mu)\,dx = 0$$

$$E_2(x-\mu) = 0$$

$$E_2(x) = \mu$$

$$\frac{\partial D_{KL}}{\partial \Sigma} = \frac{\partial}{\partial \Sigma}\left(\int \frac{1}{2}p_2(x)\left[\ln|\Sigma| + (x-\mu)^t\Sigma^{-1}(x-\mu)\right]dx\right)$$

$$= \int \frac{1}{2}p_2(x)\left[\Sigma^{-1} - \Sigma^{-1}(x-\mu)(x-\mu)^t\Sigma^{-1}\right]dx = 0$$

$$\Downarrow$$

$$\int p_2(x)\left[\Sigma - (x-\mu)(x-\mu)^t\right]dx = 0$$

$$E_2\left(\Sigma - (x-\mu)(x-\mu)^t\right) = 0$$

$$E_2\left[(x-\mu)(x-\mu)^t\right] = \Sigma$$

题目4

(1) EM算法：在数据缺失的情况下进行参数估计，已知参数θ^i的情况下，对缺失数据求期望，最大化该期望值估计新参数θ^{i+1}

(2) $$Q(\theta;\theta^0) = E\left[\ln p(x_1, x_2, x_3 | \theta)\right]$$

$$= \int \left[\ln p(x_1|\theta) + \ln p(x_2|\theta) + \ln p(x_3|\theta)\right] p(x_{32}|\theta^0; x_{31}=2)\,dx$$

$$= \ln p(x_1|\theta) + \ln p(x_2|\theta) + \int \ln p\left[\begin{pmatrix}2\\x_{12}\end{pmatrix}\Big|\theta\right] \frac{p\left[\begin{pmatrix}2\\x_{12}\end{pmatrix}\Big|\theta^0\right]}{\int p\left[\begin{pmatrix}2\\x_{32}\end{pmatrix}\Big|\theta^0\right]dx_{32}}\,dx_{32}$$

$$\int p[(\begin{smallmatrix}2\\x_{32}\end{smallmatrix})|\theta^0]\,dx_{32}$$

$$=\int_0^4 \frac{1}{2}e^{-1}\cdot\frac{1}{4}\,dx \quad = \frac{1}{2e}$$

$$Q(\theta;\theta^0) = \ln p(x_1|\theta) + \ln p(x_2|\theta) + \frac{1}{4}\int_{-\infty}^{+\infty} \ln\left(\frac{1}{\theta_1\theta_2}e^{-\frac{2}{\theta_1}}\right) dx_{32}$$

① 若 $\theta_2 \geq 4$ $\quad \frac{1}{4}\int_{-\infty}^{+\infty} \ln\left(\frac{1}{\theta_1\theta_2}e^{-\frac{2}{\theta_1}}\right) dx_{32} = -\frac{2}{\theta_1} - \ln\theta_1\theta_2$

$$Q(\theta;\theta^0) = \ln\left(\frac{1}{\theta_1\theta_2}e^{-\frac{1}{\theta_1}}\right) + \ln\left(\frac{1}{\theta_1\theta_2}e^{-\frac{3}{\theta_1}}\right) - \frac{2}{\theta_1} - \ln\theta_1\theta_2$$

$$= -\frac{6}{\theta_1} - 3\ln\theta_1\theta_2$$

$$\frac{\partial Q}{\partial\theta_1} = 6\theta_1^{-2} - 3\frac{1}{\theta_1\theta_2}\cdot\theta_2 = 0 \qquad \frac{\partial Q}{\partial\theta_2} = -3\frac{1}{\theta_1\theta_2}\cdot\theta_1 = -\frac{3}{\theta_2}$$

$\Downarrow$

$\theta_1 = 2,\ \theta_2 = 4$ 时 $Q(\theta;\theta^0)$ 最大

② $3 \leq \theta_2 \leq 4$ $\quad \frac{1}{4}\int_{-\infty}^{+\infty} \ln\left(\frac{1}{\theta_1\theta_2}e^{-\frac{2}{\theta_1}}\right) dx_{32} = -\frac{\theta_2}{4}\left(\frac{2}{\theta_1} + \ln\theta_1\theta_2\right)$

$$Q(\theta;\theta^0) = -\frac{4}{\theta_1} - 2\ln\theta_1\theta_2 - \frac{\theta_2}{2\theta_1} - \frac{1}{4}\theta_2\ln\theta_1\theta_2$$

$$\frac{\partial Q}{\partial\theta_1} = 4\theta_1^{-2} - 2\theta_1^{-1} + \frac{\theta_2}{2}\theta_1^{-2} - \frac{\theta_2}{4}\theta_1^{-1} \qquad \theta_2\uparrow\ Q(\theta;\theta^0)\downarrow$$

$\Downarrow$

$\theta_1 = 2 \quad \theta_2 = 3$ 时 Q 最大

题目5

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{t=1}^{T-1} \alpha_t(i) \beta_t(i)}$$

$$\hat{b}_{jk} = \frac{\sum_{t=1, O_t = v_k}^{T} \alpha_t(j) \beta_t(j)}{\sum_{t=1}^{T} \alpha_t(j) \beta_t(j)}$$

$\alpha_t(i)$ $\quad 1 \le t \le T \quad 1 \le i \le c \quad \alpha_t(i) = \sum_{i=1}^{N} (\ ) \quad O(c^2 T)$

$\beta_t(j)$ $\quad 1 \le t \le T \quad 1 \le j \le c \quad \beta_t(j) = \sum_{j=1}^{N} (\ ) \quad O(c^2 T)$

计算 $\hat{a}_{ij}$ 过程 $1 \le i \le c, \ 1 \le j \le c$, $\hat{a}_{ij} = \frac{\sum_t (\ )}{\sum_t (\ )} \quad O(c^2 T)$

$\therefore O(c^2 T)$

题目6.

(1) $p_n(x) = \frac{1}{n h_n} \sum_{i=1}^{n} \varphi\left(\frac{x - x_i}{h_n}\right) \qquad \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$$\bar{p}_n(x) = \frac{1}{n h_n} \sum_{i=1}^{n} E\left[\varphi\left(\frac{x - x_i}{h_n}\right)\right]$$

$$= \frac{1}{h_n} \int_{-\infty}^{+\infty} \varphi\left(\frac{x - v}{h_n}\right) p(v) dv$$

$$= \frac{1}{h_n} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - v}{h_n}\right)^2\right] \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2}\left(\frac{v - \mu}{\sigma}\right)^2\right] dv$$

$$= \frac{1}{2\pi h_n \sigma} \exp\left[-\frac{1}{2}\left(\frac{x^2}{h_n^2} + \frac{\mu^2}{\sigma^2}\right) + \frac{1}{2}\frac{\alpha^2}{\theta^2}\right] \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\left(\frac{v - \alpha}{\theta}\right)^2\right] dv$$

其中 $\theta = \frac{h_n^2 \sigma^2}{h_n^2 + \sigma^2}$ $\alpha = \theta^2 \left(\frac{x}{h_n^2} + \frac{\mu}{\sigma^2} \right)$

$\Downarrow$

$$= \frac{1}{\sqrt{2\pi}\sqrt{h_n^2 + \sigma^2}} \exp\left[-\frac{1}{2} \frac{(x-\mu)^2}{h_n^2 + \sigma^2} \right]$$

$\bar{p}_n(x) \sim N(\mu, h_n^2 + \sigma^2)$

(2) $Var[p_n(x)] = Var\left[\frac{1}{nh_n} \sum_{i=1}^{n} \varphi\left(\frac{x - x_i}{h_n} \right) \right]$

$$= \frac{1}{n^2 h_n^2} \sum Var\, \varphi\left(\frac{x - x_i}{h_n} \right)$$

$$= \frac{1}{nh_n^2} Var\left[\varphi\left(\frac{x-v}{h_n} \right) \right]$$

$$= \frac{1}{nh_n^2} \left(E\left[\varphi^2\left(\frac{x-v}{h_n} \right) \right] - E^2\left[\varphi\left(\frac{x-v}{h_n} \right) \right] \right)$$

$E\left[\varphi^2\left(\frac{x-v}{h_n} \right) \right] = \int \varphi^2\left(\frac{x-v}{h_n} \right) p(v)\, dv$

$$= \int \frac{1}{2\pi} \exp\left[-\left(\frac{x-v}{h_n} \right)^2 \right] \exp\left[-\frac{1}{2} \left(\frac{v-\mu}{\sigma} \right)^2 \right] \frac{1}{\sqrt{2\pi}\sigma} dv$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{h_n^2/2 + \sigma^2}} \exp\left[-\frac{1}{2} \frac{(x-\mu)^2}{h_n^2/2 + \sigma^2} \right)$$

$\frac{1}{nh_n^2} E\left[\varphi^2\left(\frac{x-v}{h_n} \right) \right] = \frac{1}{2nh_n\sqrt{\pi}} p(x)$

$\frac{1}{nh_n^2} E^2\left[\varphi\left(\frac{x-v}{h_n} \right) \right] = \frac{h_n}{nh_n} \frac{1}{\sqrt{2\pi}\sqrt{h_n^2 + \sigma^2}} \exp\left[-\frac{1}{2} \frac{(x-\mu)^2}{h_n^2 + \sigma^2} \right]$

$$\approx \frac{h_n}{nh_n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2} \right] \Rightarrow 0$$

$$Var[p_n(x)] \approx \frac{1}{nh_n^2} E\left[\varphi^2\left(\frac{x-v}{h_n}\right)\right] - \frac{1}{nh_n^2} E^2\left[\varphi\left(\frac{x-v}{h_n}\right)\right]$$

$$= \frac{1}{2nh_n\sqrt{\pi}} p(x)$$