模式识别 第二次作业

学号：2021E8018782022

姓名：樊宇

Question 1

(1)

因为样本$D = x_1, x_2, \ldots, x_n$都独立地服从分布$p(x|\theta)$

所以有似然函数：

$$\begin{aligned} p(D|\theta) &= \prod_{k=1}^{n} p(x_k|\theta) \\ &= \begin{cases} \frac{1}{\theta^n} & \forall x_k \in [0, \theta] \\ 0 & else \end{cases} \end{aligned}$$

要使似然函数最大，θ就要越小，但又要包含全部的样本

因此

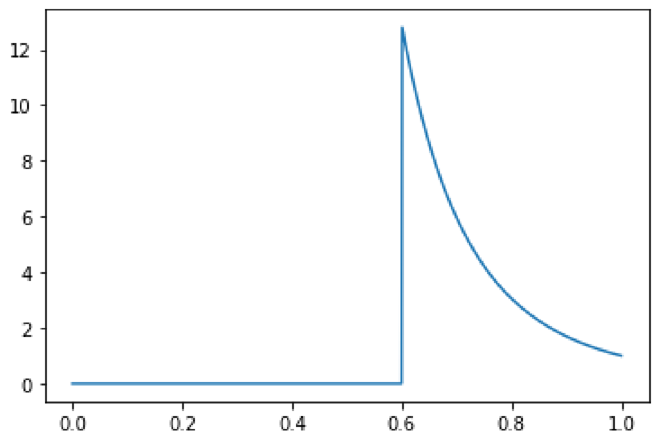
$$\min_{\theta} p(D|\theta) \Leftrightarrow \theta = max[D]$$

(2)

由（1）问可知，当$\theta \leq max[D]$时，似然函数为：0

当$\theta \geq max[D]$时，似然函数为：$\frac{1}{\theta^5}$

因此，当$0 \leq \theta \leq 1$时，似然函数$p(D|\theta)$有：



根据概率密度性质，θ必包含所有的样本，即必大于最大的样本，因此不需要知道其他四个点的值

Question 2

(1)

μ的后验概率为：

$$p(\mu|D) = \frac{p(D|\mu)p(\mu)}{\int p(D|\mu)p(\mu)d\mu}$$
$$= \alpha \prod_{k=1}^{n} p(x_k|\mu)p(\mu)$$

其中α为归一化因子

因为其中样本服从高斯分布，且μ的先验分布为$\mathrm{N}(m_0, \Sigma_o)$，将上式展开：

$$p(\mu|D) = \alpha \prod_{k=1}^{n} \frac{1}{(2\pi)^(\frac{D}{2})|\Sigma|^{\frac{1}{2}}} exp(-\frac{1}{2}(x_k-\mu)^T\Sigma^{-1}(x_k-\mu)) \frac{1}{(2\pi)^(\frac{D}{2})|\Sigma_0|^{\frac{1}{2}}} exp(-\frac{1}{2}(\mu-m_0)^T\Sigma_0^{-1}(\mu-m_0))$$
$$= \alpha' exp(-\frac{1}{2}(\mu^T(n\Sigma^{-1}+\Sigma_0^{-1})\mu - 2\mu^T(\Sigma^{-1}\sum_{k=1}^{n} x_k + \Sigma_0^{-1}m_0))$$

又因为μ的分布为$\mathrm{N}(m_0, \Sigma_o)$，可以写成：

$$p(\mu|D) = \frac{1}{(2\pi)^(\frac{D}{2})|\Sigma_n|^{\frac{1}{2}}} exp(-\frac{1}{2}(\mu-m_n)^T\Sigma_n^{-1}(\mu-m_n))$$

μ_n在均值处取得最大后验概率

于是有：

$$\mu_n = \Sigma_0(\Sigma_0 + \frac{1}{n}\Sigma)^{-1}\hat{\mu} + \frac{1}{n}\Sigma(\Sigma_0 + \frac{1}{n}\Sigma)^{-1}m_0$$
$$\Sigma_n = \Sigma_0(\Sigma_0 + \frac{1}{n}\Sigma)^{-1}\frac{1}{n}\Sigma$$

其中：

$$\hat{\mu}_n = \frac{1}{n}\sum_{k=1}^{n} x_k$$

(2)

$$E(x') = E(Ax) = AE(x) = A\mu_n$$
$$D(x') = E((x'-\mu')(x'-\mu')^T) = A\Sigma_n A^T$$

根据MAP，有：

$$\mu_n' = \Sigma_0'(\Sigma_0' + \frac{1}{n}\Sigma')^{-1}\frac{1}{n}\sum_{k=1}^{n} x_k' + \frac{1}{n}\Sigma'(\Sigma_0' + \frac{1}{n}\Sigma')^{-1}m_0'$$

简化上式可得

$$\mu_n' = A\mu_n$$

因此MAP可以正确地对变换之后的μ'进行估计，当A是非奇异矩阵的情况下，从原来的μ_n到μ_n'相当于也做了线性变换

Question 3

(1)

$$
\begin{aligned}
Q(\theta;\theta^0) &= E_{x_{32}}(lnp(x_g,x_b;\theta|\theta^0;D_g)) \\
&= \int_{-\infty}^{+\infty}[\sum_{k=1}^{2} lnp(x_k|\theta) + lnp(x_3|\theta)]p(x_{32}|\theta^0;x_{31}=2)dx_{32} \\
&= \sum_{k=1}^{2} lnp(x_k|\theta) + \int_{-\infty}^{+\infty} lnp(x_3|\theta)\frac{p(x_3|\theta^0)}{\int p((2,x'_{32})^T|\theta^0)dx'_{32}}dx_{32} \\
&= \sum_{k=1}^{2} lnp(x_k|\theta) + 2e^4\int_{-\infty}^{+\infty} lnp(x_3|\theta)p(x_3|\theta^0)dx_{32} \\
&= -4\theta_1 - 2ln\theta_1 - 2ln\theta_2 - \frac{1}{4}\int_{-\infty}^{+\infty}(2\theta_1 + ln\theta_1 + ln\theta_2)dx_{32}
\end{aligned}
$$

其中:

$$
Q(\theta;\theta^0) = \begin{cases} -4\theta_1 - 2ln\theta_1 - 2ln\theta_2 - \frac{1}{4}\int_0^{\theta_2}(2\theta_1 + ln\theta_1 + ln\theta_2)dx_{32} & 3 \le \theta_2 \le 4 \\ -4\theta_1 - 2ln\theta_1 - 2ln\theta_2 - \frac{1}{4}\int_0^{4}(2\theta_1 + ln\theta_1 + ln\theta_2)dx_{32} & 4 < \theta_2 \end{cases}
$$

又因为:

$$
\begin{aligned}
\int_{-\infty}^{+\infty} p(x_1)dx_1 &= \int_{0}^{+\infty} \frac{1}{\theta_1}e^{-x_1\theta_1}dx_1 \\
&= \frac{1}{\theta_1^2} \\
&= 1
\end{aligned}
$$

因此:

$$
\theta_1 = 1
$$

(2)

当$3 \le \theta_2 \le 4$时:

$$
\nabla_{\theta_2} Q(\theta;\theta^0) = -\frac{2}{\theta_2} - \frac{3}{4} - \frac{ln\theta_2}{4}
$$

当$4 < \theta_2$时:

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

由上可知$\nabla_{\theta_2} Q(\theta;\theta^0) < 0$，因此:

$$
\begin{aligned}
\theta_1 &= 1 \\
\theta_2 &= 3
\end{aligned}
$$

Question 4

因为:

$$
\begin{aligned}
\gamma_{ij}(t) &= \frac{\alpha_i(t-1)a_{ij}b_{ij}\beta_i(t)}{P(V^T|\theta)} \\
\hat{a}_{ij} &= \frac{\sum_{t=1}^{T}\gamma_{ij}(t)}{\sum_{t=1}^{T}\sum_k \gamma_{ij}(t)} \\
\hat{b}_{jk} &= \frac{\sum_{v(t)=k}\sum_l \gamma_{ij}(t)}{\sum_{t=1}^{T}\sum_k \gamma_{ij}(t)}
\end{aligned}
$$

所以:

$$
\begin{aligned}
T(\gamma) &= O(c^2 T) \\
T(a) &= O(c^2 T) \\
T(b) &= O(c^2 T)
\end{aligned}
$$

因此总的时间复杂度为$O(c^2T)$

Question 5

(1)

$$
\begin{aligned}
\bar{p}_n(x) &= E(p_n(x)) \\
&= \frac{1}{n}\sum_{i=1}^{n} E(\frac{1}{V_n}\varphi(\frac{x-x_i}{h_n})) \\
&= \int \frac{1}{h_n}\varphi(\frac{x-v}{h_n})p(v)dv \\
&= \frac{1}{2\pi h_n \sigma} exp(-\frac{1}{2}(\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}))\int exp(-\frac{1}{2}(v^2(\frac{\sigma^2+h_n^2}{h_n^2\sigma^2})-2v(\frac{x}{h_n^2}+\frac{\mu}{\sigma^2})))dv
\end{aligned} \tag{*}
$$

其中积分项可看作正态分布$N(\frac{\sigma^2 x+\mu h_n^2}{\sigma^2+h_n^2}, \frac{h_n^2\sigma^2}{\sigma^2+h_n^2})$的一部分，则可以计算出其积分为：

$$
\int exp(-\frac{1}{2}(v^2(\frac{\sigma^2+h_n^2}{h_n^2\sigma^2})-2v(\frac{x}{h_n^2}+\frac{\mu}{\sigma^2})))dv = \frac{1}{\frac{1}{\sqrt{2\pi}\Sigma}exp(-\frac{1}{2}\frac{m^2}{\Sigma^2})}
$$

其中：

$$
\begin{aligned}
\Sigma^2 &= \frac{h_n^2\sigma^2}{\sigma^2+h_n^2} \\
m &= \frac{\sigma^2 x+\mu h_n^2}{\sigma^2+h_n^2}
\end{aligned}
$$

因此，*式可写作：

$$
\frac{1}{\sqrt{2\pi}\sqrt{h_n^2+\sigma^2}} exp(-\frac{1}{2}(\frac{(x-\mu)^2}{h_n^2+\sigma^2}))
$$

得证

(2)

$$
\begin{aligned}
Var[p_n(x)] &= \sum_{i=1}^{n} E[(\frac{1}{nV_n}\varphi(\frac{x-x_i}{h_n})-\frac{1}{n}\bar{p}_h(x))^2] \\
&= nE(\frac{1}{n^2V_n^2}\varphi^2(\frac{x-x_i}{h_n}))-\frac{1}{n}\bar{p}_n^2(x) \\
&= \frac{1}{nh_n^2}\int \varphi^2(\frac{x-v}{h_n})p(v)dv-\frac{1}{n}\bar{p}_n^2(x)
\end{aligned} \tag{**}
$$

其中积分项与第(1)问同理，可写作：

$$
\frac{\frac{h_n}{\sqrt{2}}}{2\pi\sqrt{(\frac{h_n^2}{2}+\sigma^2)}} exp(-\frac{1}{2}\frac{(x-\mu)^2}{\frac{h_n^2}{2}+\sigma^2})
$$

因此，**式可写作：

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

整理得：

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

其中h_n本来就是一个比较小的量，在此忽略掉其二次项以及上式第二项有：

$$Var[p_n(x)]\approx\frac{1}{2\sqrt{\pi}nh_n}p(x)$$

得证