模式识别第二次作业　　2020E8017782046　　贺梦哲

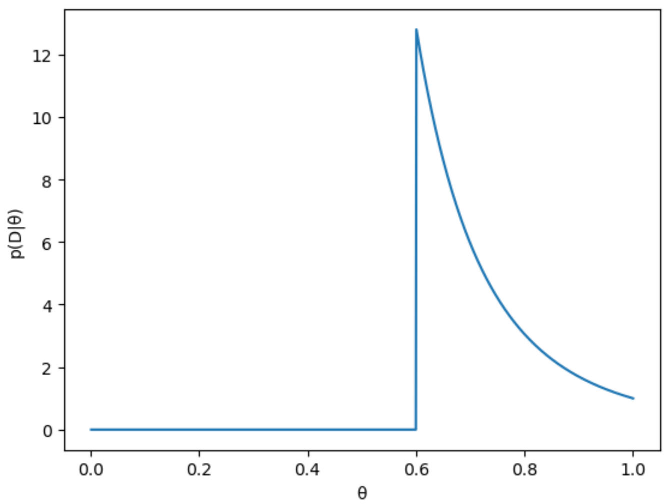
问题1

（1）可以写出最大似然函数为

$$l(\theta)=\begin{cases}\frac{1}{\theta^n} & 0\le x_k\le\theta\,(k=1,2,\ldots n)\\ 0 & others\end{cases}$$

易知，欲使最大似然函数取最大值，θ应取最最小值，同时，由于$\theta\ge x_k\,(k=1,2,\ldots n)$,故$\theta$的最大似然估计值为$max(D)$。

（2）由（1）知，$\theta=max(D)=0.6$时，最大似然估计取最大值，在n=5时，最大似然函数如下图所示，因为在$max(D)=0.6$成立的情况下，其他样本点的值不影响最大似然估计，因此无需知道。



问题2

（1）μ的后验概率为

$$\begin{aligned}p(u\mid D)&=\frac{p(D\mid u)p(u)}{\int p(D\mid u)p(u)du}\\&=\alpha\prod_{k=1}^{n}p\left(x_k\mid u\right)p(u)\\&=\alpha'\exp\left(-\frac{1}{2}\left(\sum_{k=1}^{n}\left(x_k-u\right)^t\Sigma^{-1}\left(x_k-u\right)+\left(u-m_0\right)^t\Sigma^{-1}\left(u-m_0\right)\right)\right)\\ \ln p(u\mid D)&=\ln\alpha'-\frac{1}{2}\left(\sum_{k=1}^{n}\left(x_k-u\right)^t\Sigma^{-1}\left(x_k-u\right)+\left(u-m_0\right)^t\Sigma^{-1}\left(u-m_0\right)\right)\end{aligned}$$

去掉无关项后，最大后验估计(MAP)为

$$\hat{u}=\arg\max_u\left(-\frac{1}{2}\left(\sum_{k=1}^{n}\left(x_k-u\right)^t\Sigma^{-1}\left(x_k-u\right)+\left(u-m_0\right)^t\Sigma^{-1}\left(u-m_0\right)\right)\right)\quad(1)$$

(2)在经过线性变换后，各个相关的参数变化为

$$
\begin{aligned}
\boldsymbol{\mu}' &= \mathbf{A}\boldsymbol{\mu} \\
\boldsymbol{\Sigma}' &= \mathcal{E}\left[\left(\mathbf{x}'-\boldsymbol{\mu}'\right)\left(\mathbf{x}'-\boldsymbol{\mu}'\right)^t\right] \\
&= \mathcal{E}\left[\left(\mathbf{A}\mathbf{x}-\mathbf{A}\boldsymbol{\mu}\right)\left(\mathbf{A}\mathbf{x}-\mathbf{A}\boldsymbol{\mu}\right)^t\right] \\
&= \mathcal{E}\left[\mathbf{A}\left(\mathbf{x}-\boldsymbol{\mu}\right)\left(\mathbf{x}-\boldsymbol{\mu}\right)^t\mathbf{A}^t\right] \\
&= \mathbf{A}\mathcal{E}\left[\left(\mathbf{x}-\boldsymbol{\mu}\right)\left(\mathbf{x}-\boldsymbol{\mu}\right)^t\right]\mathbf{A}^t \\
&= \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^t
\end{aligned}
$$

同理，$m_0{}' = Am_0$， $\Sigma_0{}' = A\Sigma_0 A^t$

故最大似然估计变为

$$
\begin{aligned}
\hat{u'} &= \arg\max_u \left(-\frac{1}{2}\left(\sum_{k=1}^n \left(Ax_k-Au\right)^t\left(A\Sigma A^t\right)^{-1}\left(Ax_k-Au\right)+\left(Au-Am_0\right)^t\Sigma_0^{-1}\left(Au-Am_0\right)\right)\right) \\
&= \arg\max_u \left(-\frac{1}{2}\left(\sum_{k=1}^n \left(x_k-u\right)^t A^t\left(\left(\mathbf{A}^{-1}\right)^t\boldsymbol{\Sigma}^{-1}\mathbf{A}^{-1}\right)\mathbf{A}\left(x_k-u\right)+\left(u-m_0\right)^t A^t\left(\left(\mathbf{A}^{-1}\right)^t\boldsymbol{\Sigma}_0^{-1}\mathbf{A}^{-1}\right)\mathbf{A}\left(u-m_0\right)\right)\right) \\
&= \arg\max_u \left(-\frac{1}{2}\left(\sum_{k=1}^n \left(x_k-u\right)^t\left(\mathbf{A}^t\left(\mathbf{A}^{-1}\right)^t\right)\boldsymbol{\Sigma}^{-1}\left(\mathbf{A}^{-1}\mathbf{A}\right)\left(x_k-u\right)+\left(u-m_0\right)^t\left(\mathbf{A}^t\left(\mathbf{A}^{-1}\right)^t\right)\boldsymbol{\Sigma}_0^{-1}\left(\mathbf{A}^{-1}\mathbf{A}\right)\left(u-m_0\right)\right)\right) \\
&= \arg\max_u \left(-\frac{1}{2}\left(\sum_{k=1}^n \left(x_k-u\right)^t\boldsymbol{\Sigma}^{-1}\left(x_k-u\right)+\left(u-m_0\right)^t\boldsymbol{\Sigma}_0^{-1}\left(u-m_0\right)\right)\right)
\end{aligned} \tag{2}
$$

式（2）与式（1）相等，故对于u'也能给出正确的估计。

问题3

$$
\begin{aligned}
Q\left(\theta;\theta^0\right) &= \mathcal{E}_{x_{32}}\left[\ln p\left(\mathrm{x}_g,\mathrm{x}_b;\theta\right)\mid\theta^0,\mathcal{D}_g\right] \\
&= \int_{-\infty}^{\infty}\left(\ln p\left(\mathrm{x}_1\mid\theta\right)+\ln p\left(\mathrm{x}_2\mid\theta\right)+\ln p\left(\mathrm{x}_3\mid\theta\right)\right)p\left(x_{32}\mid\theta^0,x_{31}=2\right)dx_{32} \\
&= \ln p\left(\mathrm{x}_1\mid\theta\right)+\ln p\left(\mathrm{x}_2\mid\theta\right)+\int_{-\infty}^{\infty}\ln p\left(\mathrm{x}_3\mid\theta\right)\cdot p\left(x_{32}\mid\theta^0,x_{31}=2\right)dx_{32} \\
&= \ln p\left(\mathrm{x}_1\mid\theta\right)+\ln p\left(\mathrm{x}_2\mid\theta\right)+\int_{-\infty}^{\infty}\ln p\left(\begin{pmatrix}2\\x_{32}\end{pmatrix}\mid\theta\right)\cdot\frac{p\left(\begin{pmatrix}2\\x_{32}\end{pmatrix}\mid\theta^0\right)}{\int_{-\infty}^{\infty}p\left(\begin{pmatrix}2\\x'_{32}\end{pmatrix}\mid\theta^0\right)d'_{32}}dx_{32} \\
&= \ln p\left(\mathrm{x}_1\mid\theta\right)+\ln p\left(\mathrm{x}_2\mid\theta\right)+\frac{1}{4}\int_{-\infty}^{\infty}\ln\left(\frac{1}{\theta_1}e^{-\frac{2}{\theta_1}}\frac{1}{\theta_2}\right)dx_{32} \\
&= \ln\left(\frac{1}{\theta_1}e^{-\theta_1}\frac{1}{\theta_2}\right)+\ln\left(\frac{1}{\theta_1}e^{-\frac{3}{\theta_1}}\frac{1}{\theta_2}\right)+\frac{1}{4}\int_{-\infty}^{\infty}\ln\left(\frac{1}{\theta_1}e^{-\frac{2}{\theta_1}}\frac{1}{\theta_2}\right)dx_{32}
\end{aligned}
$$

由于已知的样本中$x_2=3$,故当θ_2<3时，不满足条件。

当$3\le\theta_2\le 4$时，

$$
\begin{aligned}
Q\left(\theta;\theta^0\right) &= \ln\left(\frac{1}{\theta_1}e^{-\frac{1}{\theta\,1}}\frac{1}{\theta_2}\right)+\ln\left(\frac{1}{\theta_1}e^{-\frac{3}{\theta\,1}}\frac{1}{\theta_2}\right)+\frac{1}{4}\int_0^{\theta_2}\ln\left(\frac{1}{\theta_1}e^{-\frac{2}{\theta_1}}\frac{1}{\theta_2}\right)dx_{32} \\
&= \ln\left(\frac{1}{\theta_1}e^{-\frac{1}{\theta_1}}\frac{1}{\theta_2}\right)+\ln\left(\frac{1}{\theta_1}e^{-\frac{3}{\theta_1}}\frac{1}{\theta_2}\right)+\frac{\theta_2}{4}\ln\left(\frac{1}{\theta_1}e^{-\frac{2}{\theta_1}}\frac{1}{\theta_2}\right) \\
&= -2\ln(\theta_1\theta_2)-\frac{4}{\theta_1}-\frac{\theta_2}{4}\ln(\theta_1\theta_2)-\frac{\theta_2}{2\theta_1} \\
\frac{\partial Q(\theta;\,\theta^0)}{\partial\theta_1} &= -\frac{2}{\theta\,1}+\frac{4}{\theta_1^2}-\frac{\theta_2}{4\theta_1}+\frac{\theta 2}{2\theta_1^2} \\
&= \frac{-8\theta_1+16-\theta_1\theta_2+2\theta_2}{4\theta_1^2} \\
&= \frac{(2-\theta_1)(\theta_2+8)}{4\theta_1^2}
\end{aligned}
$$

故当$\theta_1=2$时，取极值。此时

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

因为$Q\left(\theta;\theta^0\right)$在$\theta_2>0$时单调递减，故当$\theta_2=3$时，$Q\left(\theta;\theta^0\right)$取得最大值$-6.75$

当$\theta_2>4$时，

$$\begin{aligned}
Q\left(\theta;\theta^0\right) &= \ln\left(\frac{1}{\theta_1}e^{-\frac{1}{\theta_1}}\frac{1}{\theta_2}\right)+\ln\left(\frac{1}{\theta_1}e^{-\frac{3}{\theta_1}}\frac{1}{\theta_2}\right)+\frac{1}{4}\int_0^4\ln\left(\frac{1}{\theta_1}e^{-\frac{2}{\theta_1}}\frac{1}{\theta_2}\right)dx_{32}\\
&= -\frac{6}{\theta_1}-3\ln(\theta_1\theta_2)\\
\frac{\partial Q\left(\theta;\theta^0\right)}{\partial\theta_1} &= \frac{6-3\theta_1}{\theta_1^2}
\end{aligned}$$

故当$\theta_1=2$时取得极值，此时$Q\left(\theta;\theta^0\right)==-3-3\ln(2\theta_2)$为减函数，故当$\theta_2=4$时取得最大值$-9.24$

1.在E-step中

$$Q\left(\theta;\theta^0\right)=\begin{cases}\text{无 意 义} & \theta_2<3\\ -2\ln(\theta_1\theta_2)-\frac{4}{\theta_1}-\frac{\theta_2}{4}\ln(\theta_1\theta_2)-\frac{\theta_2}{2\theta_1} & 3\le\theta_2\le4\\ -\frac{6}{\theta_1}-3\ln(\theta_1\theta_2) & \theta_2>4\end{cases}$$

2.M-step运算如上，综上$\theta=\begin{pmatrix}2\\3\end{pmatrix}$

问题4

$$\begin{aligned}
\hat{a}_{ij} &= \frac{\sum_{t=1}^{T}\gamma_{ij}(t)}{\sum_{t=1}^{T}\sum_k\gamma_{ik}(t)}\\
\hat{b}_{ij} &= \frac{\sum_{??}^{??}\gamma_{ij}}{\sum_{t=1}^{T}\gamma_{ij}(t)}\\
\gamma_{ij}(t) &= \frac{\alpha_i(t-1)a_{ij}b_{ij}\beta_i(t)}{P\left(V^T\mid M\right)}
\end{aligned}$$

$\alpha_i(t)'\mathrm{s}$ 和 $P\left(V^T\mid M\right)$通过前向算法计算得到，分别需要$O\left(c^2T\right)$的复杂度，

$$\begin{aligned}
&\text{For t} = \text{T to } 1(\text{by} - 1)\\
&\text{For i} = 1 \text{ to c}\\
&\beta_i(t)=\sum_j a_{ij}b_{jk}v(t+1)\beta_j(t+1)\\
&\text{End}
\end{aligned}$$

$\beta_i(t)$的计算由上式得到，也需要$O\left(c^2T\right)$的复杂度

此时，计算γ_{ij}也需要$O(c^2T)$的复杂度

$$\underbrace{O\left(c^2T\right)}_{\alpha_i(t)'\mathrm{s}}+\underbrace{O\left(c^2T\right)}_{\beta_i(t)'\mathrm{s}}+\underbrace{O\left(c^2T\right)}_{\gamma_{ij}(t)'\mathrm{s}}=O\left(c^2T\right)$$

因此共需要$O(c^2T)$的复杂度。

问题5

（1）根据题目可以写出，

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$
$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

故估计的期望值

$$\begin{aligned}
\bar{p}_n(x)=\mathcal{E}\left[p_n(x)\right]&=\frac{1}{nh_n}\sum_{i=1}^{n}\mathcal{E}\left[\varphi\left(\frac{x-x_i}{h_n}\right)\right]\\
&=\frac{1}{h_n}\int_{-\infty}^{\infty}\varphi\left(\frac{x-v}{h_n}\right)p(v)dv\\
&=\frac{1}{h_n}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-v}{h_n}\right)^2\right]\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right]dv\\
&=\frac{1}{2\pi h_n\sigma}\exp\left[-\frac{1}{2}\left(\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}\right)\right]\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}v^2\left(\frac{1}{h_n^2}+\frac{1}{\sigma^2}\right)-2v\left(\frac{x}{h_n^2}+\frac{\mu}{\sigma^2}\right)\right]dv\\
&=\frac{1}{2\pi h_n\sigma}\exp\left[-\frac{1}{2}\left(\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}\right)+\frac{1}{2}\frac{\alpha^2}{\theta^2}\right]\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{v-\alpha}{\theta}\right)^2\right]dv
\end{aligned}$$

其中$\theta^2=\frac{1}{1/h_n^2+1/\sigma^2}=\frac{h_n^2\sigma^2}{h_n^2+\sigma^2}$，$\alpha=\theta^2\left(\frac{x}{h_n^2}+\frac{\mu}{\sigma^2}\right)$

$$\begin{aligned}
\bar{p}_n(x)&=\frac{\sqrt{2\pi}\theta}{2\pi h_n\sigma}\exp\left[-\frac{1}{2}\left(\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}\right)+\frac{1}{2}\frac{\alpha^2}{\theta^2}\right]\\
&=\frac{1}{\sqrt{2\pi}h_n\sigma}\frac{h_n\sigma}{\sqrt{h_n^2+\sigma^2}}\exp\left[-\frac{1}{2}\left(\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}-\frac{\alpha^2}{\theta^2}\right)\right]
\end{aligned}$$

由于

$$\begin{aligned}
\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}-\frac{\alpha^2}{\theta^2}&=\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}-\frac{\theta^4}{\theta^2}\left(\frac{x}{h_n^2}+\frac{\mu}{\sigma^2}\right)^2\\
&=\frac{x^2h_n^2}{\left(h_n^2+\sigma^2\right)h_n^2}+\frac{\mu^2\sigma^2}{\left(h_n^2+\sigma^2\right)\sigma^2}-\frac{2x\mu}{h_n^2+\sigma^2}\\
&=\frac{(x-\mu)^2}{h_n^2+\sigma^2}
\end{aligned}$$

代入原式得到

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

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

（2)

$$\begin{aligned}
\mathrm{Var}[p_n(x)]&=\mathrm{Var}\left[\frac{1}{nh_n}\sum_{i=1}^{n}\varphi\left(\frac{x-x_i}{h_n}\right)\right]\\
&=\frac{1}{n^2h_n^2}\sum_{i=1}^{n}\mathrm{Var}\left[\varphi\left(\frac{x-x_i}{h_n}\right)\right]\\
&=\frac{1}{nh_n^2}\mathrm{Var}\left[\varphi\left(\frac{x-v}{h_n}\right)\right]\\
&=\frac{1}{nh_n^2}\left\{\mathcal{E}\left[\varphi^2\left(\frac{x-v}{h_n}\right)\right]-\left(\mathcal{E}\left[\varphi\left(\frac{x-v}{h_n}\right)\right]\right)^2\right\}
\end{aligned}$$

$$
\begin{aligned}
\mathcal{E}\left[\varphi^2\left(\frac{x-v}{h_n}\right)\right] &= \int \varphi^2\left(\frac{x-v}{h_n}\right)p(v)dv \\
&= \int_{-\infty}^{\infty}\frac{1}{2\pi}\exp\left[-\left(\frac{x-v}{h_n}\right)^2\right]\exp\left[-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right]\frac{1}{\sqrt{2\pi}\sigma}dv \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-v}{h_n/\sqrt{2}}\right)^2\right]\exp\left[-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right]\frac{1}{\sqrt{2\pi}\sigma}dv \\
&= \frac{h_n/\sqrt{2}}{2\pi\sqrt{h_n^2/2+\sigma^2}}\exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{h_n^2/2+\sigma^2}\right]
\end{aligned}
$$

因此

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

当$h_n\to 0$时，$\sqrt{h_n^2/2+\sigma^2}\simeq\sigma$上式可以近似为

$$
\begin{aligned}
\frac{1}{nh_n^2}\mathcal{E}\left[\varphi^2\left(\frac{x-v}{h_n}\right)\right] &\simeq \frac{1}{nh_n 2\sqrt{\pi}}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \\
&= \frac{1}{2nh_n\sqrt{\pi}}p(x)
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{nh_n^2}\left\{\mathcal{E}\left[\varphi\left(\frac{x-v}{h_n}\right)\right]\right\}^2 &= \frac{1}{nh_n^2}h_n^2\left(\frac{1}{\sqrt{2\pi}\sqrt{h_n^2+\sigma^2}}\exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{h_n^2+\sigma^2}\right]\right)^2 \\
&= \frac{h_n}{nh_n}\frac{1}{2\pi(h_n^2+\sigma^2)}\exp\left[\frac{(x-\mu)^2}{h_n^2+\sigma^2}\right] \\
&\simeq \frac{h_n}{nh_n}\frac{1}{2\pi\sigma^2}\exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right] \simeq 0
\end{aligned}
$$

因此，题目得证：

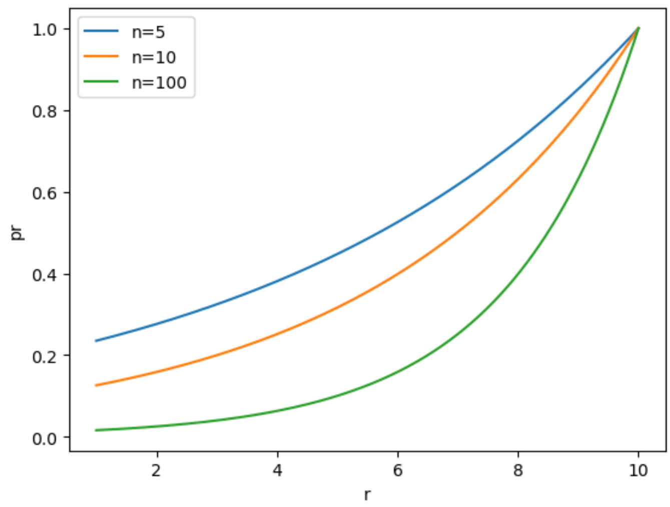
$$
\mathrm{Var}[P_n(x)] \simeq \frac{p(x)}{2nh_n\sqrt{\pi}}
$$

问题6

（1）假设n个样本点相互独立且均匀分布，则每个维度有$n^{\frac{1}{d}}$个样本点，其中r个维度有$n^{\frac{r}{d}}$个样本点。选中正确的数据作为最近邻的概率为

$$
p = \frac{n^{\frac{r}{d}}}{n} = n^{\frac{r}{d}-1}
$$

故当d=10时$p = n^{\frac{r}{10}-1}$，取不同的n画图得到概率p关于r的曲线如下：



（2）贝叶斯决策面为

$$
\begin{gathered}
g_1(x)=g_2(x) \\
p\left(\omega_1 \mid x\right)=p\left(\omega_2 \mid x\right) \\
\frac{p\left(\boldsymbol{x} \mid \omega_1\right) p\left(\omega_1\right)}{p(\boldsymbol{x})}=\frac{p\left(\boldsymbol{x} \mid \omega_2\right) p\left(\omega_2\right)}{p(\boldsymbol{x})} \\
p\left(\boldsymbol{x} \mid \omega_1\right) p\left(\omega_1\right)=p\left(\boldsymbol{x} \mid \omega_2\right) p\left(\omega_2\right)
\end{gathered}
$$

由于两类样本数均为$n/2$,故可以假设两类先验概率相等,即 $p(w_1)=p(w_2)$

由题目知概率密度为斜坡函数的乘积，故可以得到下式

$$
p\left(x \mid \omega_1\right)=\prod_{i=1}^d k_i x_i
$$

$$
p\left(x \mid \omega_2\right)=\prod_{i=1}^d k_i\left(1-x_i\right)
$$

因此贝叶斯决策面简化为

$$
\prod_{i=1}^d x_i=\prod_{i=1}^d\left(1-x_i\right)
$$