模式识别课后作业二

2020E8013282019 李一帆

October 24, 2020

1 问题 1

1.1 第一小问

对于 θ 的似然函数 $P(D|\theta)$ 可以表示为：

$$\begin{aligned} P(D|\theta) &= \prod_{k=1}^{n} P(x_k|D) \\ &= \prod_{k=1}^{n} \frac{1}{\theta} \\ &= \frac{1}{\theta^n} \end{aligned}$$

将上述公式进行取对数可以得到

$$l(\theta) = lnP(D|\theta) = -nln\theta$$

将对数似然函数 $l(\theta)$ 进行求导运算后可以得到：

$$\frac{\partial l}{\partial \theta} = -\frac{n}{\theta}$$

由导数可知，$l(\theta)$ 为关于 θ 的单调递减函数，当 θ 最小时 $l(\theta)$ 最大。由题意可知，$\theta \geq \mathrm{x}$，因此当 $\theta = \max D$ 时，$l(\theta)$ 最大。

1.2 第二小问

似然函数 $P(D|\theta)$ 和对数似然函数 $l(\theta)$ 如 Figure 1 所示。由图可以看出，在 $\theta = 0.6$ 时，似然函数取得最大值。

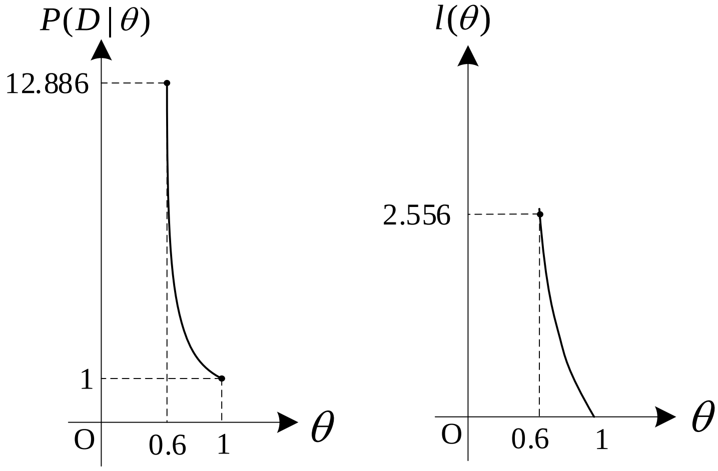

Figure 1: 错误率.

2 问题 2

2.1 第一小问

由题目可知，训练样本服从高斯分布 $p(\mathrm{x}|\mu) \sim N(\mu, \Sigma)$，未知变量 μ 服从高斯分布 $p(\mu) \sim N(\mu_0|\Sigma_0)$。待估计参数为 μ，在给定训练数据后，我们需要最大化的目标为如下公式。

$$l(\mu)p(\mu) = \ln\left[p(D|\mu)\right] p(\mu)$$

其中，$\ln\left[p(D|\mu)\right]$ 可以表示为如下形式。

$$\begin{aligned}
\ln\left[p(D|\mu)\right] &= \ln(\prod_{k=1}^{n} p(x_k|\mu)) \\
&= \sum_{k=1}^{n} \ln\left[p(x_k|\mu)\right] \\
&= \sum_{k=1}^{n} \ln\left[\frac{1}{(2\pi)^{d/2}|\Sigma|^{\frac{1}{2}}} \exp(-\frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu))\right] \\
&= \sum_{k=1}^{n} \left\{\ln\left[\frac{1}{(2\pi)^{d/2}|\Sigma|^{\frac{1}{2}}}\right] - \frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu))\right\} \\
&= -\frac{n}{2}\ln\left[(2\pi)^d\left|\Sigma\right|\right] - \sum_{k=1}^{n}\frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu))
\end{aligned}$$

其中，$p(\mu)$ 可以表示为如下形式。

$$p(\mu) = \frac{1}{(2\pi)^{d/2}|\Sigma_0|^{1/2}} \exp\left[-\frac{1}{2}(\mu-\mu_0)^t\Sigma_0^{-1}(\mu-\mu_0)\right]$$

因此对于均值的最大后验概率估计可以表示为如下形式。

$$\hat{\mu} = \arg\max_{\mu}\left[-\frac{n}{2}\ln\left[(2\pi)^d\left|\Sigma\right|\right] - \sum_{k=1}^{n}\frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu))\right] \times$$

$$\frac{1}{(2\pi)^{d/2}|\Sigma_0|^{1/2}} \exp\left[-\frac{1}{2}(\mu-\mu_0)^t\Sigma_0^{-1}(\mu-\mu_0)\right]$$

2.2 第二小问

经过线性变换后的均值 μ' 和协方差 Σ' 可以表示为如下形式。

$$\mu' = \mathrm{E}\left[x'\right] = \mathrm{E}\left[\mathrm{A}x\right] = \mathrm{A}\mathrm{E}\left[x\right] = \mathrm{A}\mu$$

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

因此，经过对 x 的线性变换后的对数似然函数可以表示为如下形式。

$$\begin{aligned}
\ln\left[p(D'|\mu')\right] &= \ln\left(\prod_{k=1}^{n} p({x'}_k|\mu')\right) \\
&= \ln\left(\prod_{k=1}^{n} p(\mathrm{A}x_k|\mathrm{A}\mu)\right) \\
&= \sum_{k=1}^{n}\ln\left[\frac{1}{(2\pi)^{d/2}\left|\mathrm{A}\Sigma\mathrm{A}^t\right|}\exp\left(-\frac{1}{2}(\mathrm{A}x_k-\mathrm{A}\mu)^t(\mathrm{A}\Sigma\mathrm{A}^t)^{-1}(\mathrm{A}x_k-\mathrm{A}\mu)\right)\right] \\
&= -\frac{n}{2}\ln\left[(2\pi)^{d/2}\left|\mathrm{A}\Sigma\mathrm{A}^t\right|\right] - \sum_{k=1}^{n}\frac{1}{2}\left((\mathrm{A}x_k-\mathrm{A}\mu)^t(\mathrm{A}\Sigma\mathrm{A}^t)^{-1}(\mathrm{A}x_k-\mathrm{A}\mu)\right) \\
&= -\frac{n}{2}\ln\left[(2\pi)^{d/2}\left|\mathrm{A}\Sigma\mathrm{A}^t\right|\right] - \sum_{k=1}^{n}\frac{1}{2}((x_k-\mu)^t\mathrm{A}^t)((\mathrm{A}^{-1})^t\Sigma^{-1}\mathrm{A}^{-1})(\mathrm{A}(x_k-\mu)) \\
&= -\frac{n}{2}\ln\left[(2\pi)^{d/2}\left|\mathrm{A}\Sigma\mathrm{A}^t\right|\right] - \sum_{k=1}^{n}\frac{1}{2}(x_k-\mu)^t(\mathrm{A}^t(\mathrm{A}^t)^{-1})\Sigma^{-1}(\mathrm{A}^{-1}\mathrm{A})(x_k-\mu)) \\
&= -\frac{n}{2}\ln\left[(2\pi)^{d/2}\left|\mathrm{A}\Sigma\mathrm{A}^t\right|\right] - \sum_{k=1}^{n}\frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu))
\end{aligned}$$

同理，经过线性变换后的高斯密度函数 $p(\mu')$ 可以表示为如下形式。

$$\begin{aligned}
p(\mu') &= \frac{1}{(2\pi)^{d/2}|{\Sigma'}_0|^{1/2}}\exp\left[-\frac{1}{2}(\mu'-\mu_0)^t\Sigma_0^{-1}(\mu'-\mu_0)\right] \\
&= \frac{1}{(2\pi)^{d/2}|{\Sigma'}_0|^{1/2}}\exp\left[-\frac{1}{2}(\mathrm{A}\mu-\mathrm{A}\mu_0)^t(\mathrm{A}\Sigma_0\mathrm{A}^t)^{-1}(\mathrm{A}\mu-\mathrm{A}\mu_0)\right] \\
&= \frac{1}{(2\pi)^{d/2}|{\Sigma'}_0|^{1/2}}\exp\left[-\frac{1}{2}(\mu-\mu_0)^t\mathrm{A}^t(\mathrm{A}^{-1})^t\Sigma_0^{-1}\mathrm{A}^{-1}\mathrm{A}(\mu-\mu_0)\right] \\
&= \frac{1}{(2\pi)^{d/2}|{\Sigma'}_0|^{1/2}}\exp\left[-\frac{1}{2}(\mu-\mu_0)^t\Sigma_0^{-1}(\mu-\mu_0)\right]
\end{aligned}$$

因此对于 μ' 的最大后验概率估计可以表示为下式。

$$\hat{\mu}' = \arg\max_{\mu}\left[-\frac{n}{2}\ln\left[(2\pi)^{d/2}\left|\mathrm{A}\Sigma\mathrm{A}^t\right|\right]-\sum_{k=1}^{n}\frac{1}{2}(x_k-\mu)^t\Sigma^{-1}(x_k-\mu))\right]$$
$$\times\frac{1}{(2\pi)^{d/2}|\Sigma'_0|^{1/2}}\exp\left[-\frac{1}{2}(\mu-\mu_0)^t\Sigma_0^{-1}(\mu-\mu_0)\right]$$

3 问题 3

3.1 第 1 小问

E 步可以表示为如下形式。

$$\begin{aligned}
Q(\theta|\theta^0) &= \mathrm{E}_{x_{32}}\left[\ln p(x_g,x_b;\theta)|\theta^0,D_g\right]\\
&= \mathrm{E}_{x_{32}}\left[\ln\prod_{k=1}^{3}p(x_k|\theta)\right]\\
&= \mathrm{E}_{x_{32}}\left[\sum_{k=1}^{2}\ln p(x_k|\theta)+\ln p(x_3|\theta)\right]\\
&= \int_{-\infty}^{+\infty}\left[\sum_{k=1}^{2}\ln p(x_k|\theta)+\ln p(x_3|\theta)\right]p(x_{32}|\theta^0,x_{31}=2)dx_{32}\\
&= \sum_{k=1}^{2}\ln p(x_k|\theta)+\int_{-\infty}^{+\infty}\ln p(\begin{pmatrix}2\\x_{32}\end{pmatrix}\Bigg|\theta^0)\cdot\frac{p(\begin{pmatrix}2\\x_{32}\end{pmatrix}\Bigg|\theta^0)}{\int_{-\infty}^{+\infty}p(\begin{pmatrix}2\\x'_{32}\end{pmatrix}\Bigg|\theta^0)dx'_{32}}dx_{32}\\
&= \sum_{k=1}^{2}\ln p(x_k|\theta)+\int_{-\infty}^{+\infty}\ln p(\begin{pmatrix}2\\x_{32}\end{pmatrix}\Bigg|\theta)\cdot\frac{p(\begin{pmatrix}2\\x_{32}\end{pmatrix}\Bigg|\theta^0))}{p(x_1=2|\,\theta_1=2)\int_{0}^{+\infty}p(x'_{32}|\,\theta_2=4)dx'_{32}}dx_{32}\\
&= \sum_{k=1}^{2}\ln p(x_k|\theta)+2e\int_{-\infty}^{+\infty}\ln p(\begin{pmatrix}2\\x_{32}\end{pmatrix}\Bigg|\theta)\cdot p(\begin{pmatrix}2\\x_{32}\end{pmatrix}\Bigg|\theta^0)dx_{32}
\end{aligned}$$

其中，在 θ_2 取不同值时，上式第二项可以分别表示为如下形式。

当 $3\le\theta_2<4$ 时：

$$\begin{aligned}
2e\int_{-\infty}^{+\infty}\ln p(\begin{pmatrix}2\\x_{32}\end{pmatrix}\Bigg|\theta)\cdot p(\begin{pmatrix}2\\x_{32}\end{pmatrix}\Bigg|\theta^0)dx_{32} &= 2e\int_{0}^{\theta_2}\ln(\frac{1}{\theta_1}e^{\frac{-2}{\theta_1}}\cdot\frac{1}{\theta_2})\frac{1}{2e}\cdot\frac{1}{4}dx_{32}\\
&= \frac{1}{4}\theta_2\ln(\frac{1}{\theta_1}e^{\frac{-2}{\theta_1}}\cdot\frac{1}{\theta_2})
\end{aligned}$$

当 $\theta_2 \geq 4$ 时：

$$2e\int_{-\infty}^{+\infty} \ln p(\left(\begin{array}{c} 2 \\ x_{32} \end{array}\right)\Bigg|\theta)\cdot p(\left(\begin{array}{c} 2 \\ x_{32} \end{array}\right)\Bigg|\theta^0)dx_{32} = 2e\int_0^4 \ln(\frac{1}{\theta_1}e^{\frac{-2}{\theta_1}}\cdot\frac{1}{\theta_2})\frac{1}{2e}\cdot\frac{1}{4}dx_{32}$$
$$= \ln(\frac{1}{\theta_1}e^{\frac{-2}{\theta_1}}\cdot\frac{1}{\theta_2})$$

当 $\theta_2 < 3$ 时，无意义（与数据集相悖）。

因此，最终对于似然的期望 $Q(\theta;\theta^0)$ 可以表示为如下形式。

$$Q(\theta;\theta^0) = \begin{cases} -6\frac{1}{\theta_1} - 3\ln(\theta_1\theta_2), & \theta_2 \geq 4 \\ -(\frac{\theta_2}{2}+4)\frac{1}{\theta_1} - (2+\frac{\theta_2}{4})\ln(\theta_1\theta_2), & 3 \leq \theta_2 < 4 \\ 0, & \theta_2 < 3 \end{cases}$$

3.2 第 2 小问

(a) 当 $\theta_2 \geq 4$ 时，似然函数期望对于 θ_1 的导数为：

$$\frac{\partial Q}{\partial\theta_1} = \frac{6}{\theta_1^2} - \frac{3}{\theta_1}$$

令导数为 0 可以得到 θ_1 的取值为 2。同理可得似然函数期望对于 θ_2 的导数为下式，并且 θ_2 恒负。

$$\frac{\partial Q}{\partial\theta_2} = \frac{-3}{\theta_2}$$

因此，当 θ_2 取最小值 4 时，$Q(\theta;\theta_0)$ 最大。此时，$\theta = \left(\begin{array}{c} 2 \\ 4 \end{array}\right)$，最大值 $\max Q(\theta;\theta_0) = -9.238$。

(b) 当 $3 \leq \theta_2 < 4$ 时，Q 对于 θ_1 的导数为如下形式。

$$\frac{\partial Q}{\partial\theta_1} = (\frac{\theta_2}{2}+4)\frac{1}{\theta_1^2} - (\frac{\theta_2}{4}+2)\frac{1}{\theta_1}$$

令导数为 0 可以得到 $\theta_1 = 2$。将 $\theta_1 = 2$ 代入 Q，并对 θ_2 求偏导，可以得到如下公式。可以看出在 $3 \leq \theta_2 < 4$ 的条件下，$\frac{\partial Q}{\partial\theta_2}$ 为负，最大值在 $\theta_2 = 3$ 处取得。

$$\frac{\partial Q}{\partial\theta_2} = -\frac{1}{2} - \frac{1}{4}\ln(2\theta_2) - \frac{1}{\theta_2}$$

因此，最终使得 $Q(\theta;\theta^0)$ 最大的 θ 可以表示为 $\theta = \left(\begin{array}{c} 2 \\ 3 \end{array}\right)$，最大值 $\max Q(\theta;\theta_0) = -7.667$。

4 问题 4

$\hat{a}_{ij}, \hat{b}_{jk}, v_{ij}(t)$ 由下列公式计算得出。

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T} v_{ij}(t)}{\sum_{t=1}^{T}\sum_k v_{ik}(t)} \qquad \hat{b}_{jk} = \frac{\sum_{t=1}^{T}\sum_l v_{il}(t)}{\sum_{t=1}^{T}\sum_l v_{jl}(t)} \qquad v_{ij}(t) = \frac{\alpha_i(t-1)a_{ij}b_{jk}\beta_j(t)}{P(V^T|\theta)}$$

$\alpha_i(t-1)$ 和 $P(V^T \big| \theta)$ 通过前向传播算法进行计算，需要的时间复杂度为 $O(c^2T)$。$\beta_i(t)$ 所需要的时间复杂度为 $O(c^2T)$，$v_{ij}(t)$ 的复杂度主要通过 α_i, a_{ij}, b_{jk} 进行计算，一共花费的时间复杂度为 $O(c^2T)$。

5 问题 5

5.1 第 1 小问

$\bar{p}_n(x)$ 的推导过程如下。

$$\begin{aligned}
\bar{p}_n(x) &= \mathrm{E}\left[p_n(x)\right] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[\frac{1}{V_n}\varphi(\frac{x-x_i}{h_n})\right] = \frac{1}{V_n}\mathrm{E}\left[\varphi(\frac{x-v}{h_n})\right] \\
&= \int\limits_{-\infty}^{+\infty}\frac{1}{V_n}\varphi(\frac{x-v}{h_n})\cdot p(v)dv \\
&= \int\limits_{-\infty}^{+\infty}\frac{1}{V_n}\cdot\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{\left(\frac{x-v}{h_n}\right)^2}{2}\right]\cdot\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(v-\mu)^2}{2\sigma^2}\right]dv \\
&= \frac{1}{2\pi V_n\sigma}\cdot\int\limits_{-\infty}^{+\infty}\exp\left(-\frac{x^2-2vx+v^2}{2h_n^2}\right)\cdot\exp\left(-\frac{v^2-2v\mu+\mu^2}{2\sigma^2}\right)dv \\
&= \frac{1}{2\pi V_n\sigma}\cdot\exp\left[-\left(\frac{x^2}{2h_n^2}+\frac{\mu^2}{2\sigma^2}\right)\right]\cdot\int\limits_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2}\left[v^2\left(\frac{1}{h_n^2}+\frac{1}{\sigma^2}\right)-2v\left(\frac{\mu}{\sigma^2}+\frac{x}{h_n^2}\right)\right]\right\}dv \\
&= \frac{1}{2\pi V_n\sigma}\cdot\exp\left[-\left(\frac{x^2}{2h_n^2}+\frac{\mu^2}{2\sigma^2}\right)\right]\cdot\int\limits_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2}\left(\frac{1}{h_n^2}+\frac{1}{\sigma^2}\right)\left[v^2-2v\left(\frac{\frac{\mu}{\sigma^2}+\frac{x}{h_n^2}}{\frac{1}{h_n^2}+\frac{1}{\sigma^2}}\right)\right]\right\}dv \\
&= \frac{1}{2\pi V_n\sigma}\cdot\exp\left[-\left(\frac{x^2}{2h_n^2}+\frac{\mu^2}{2\sigma^2}\right)+\frac{1}{2}\cdot\frac{a^2}{b^2}\right]\cdot\int\limits_{-\infty}^{+\infty}\exp\left[-\frac{1}{2}\left(\frac{v-a}{b}\right)^2\right]dv
\end{aligned}$$

其中，a, b 分别表示为如下形式。

$$b^2 = \frac{1}{h_n^2}+\frac{1}{\sigma^2} = \frac{\sigma^2 h_n^2}{\sigma^2+h_n^2} \qquad a = \frac{\frac{\mu}{\sigma^2}+\frac{x}{h_n^2}}{\frac{1}{h_n^2}+\frac{1}{\sigma^2}} = b^2\left(\frac{\mu}{\sigma^2}+\frac{x}{h_n^2}\right)$$

$\bar{p}_n(x)$ 可以继续进行合并，如下式所示。

$$\begin{aligned}
\bar{p}_n(x) &= \frac{1}{2\pi V_n\sigma}\cdot\exp\left[-\frac{1}{2}\left(\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}\right)+\frac{1}{2}\cdot\frac{a^2}{b^2}\right]\cdot\sqrt{2\pi}b\cdot\int\limits_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}b}\exp\left[-\frac{1}{2}\left(\frac{v-a}{b}\right)^2\right]dv \\
&= \frac{b}{\sqrt{2\pi}V_n\sigma}\cdot\exp\left[-\frac{1}{2}\left(\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}-\frac{a^2}{b^2}\right)\right]
\end{aligned}$$

上式小括号中的内容可以进行合并，如下式所示。

$$\begin{aligned}\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}-\frac{a^2}{b^2}&=\frac{x^2}{h_n^2}+\frac{x^2}{\sigma^2}-\frac{\sigma^2h_n^2}{\sigma^2+h_n^2}\cdot\left(\frac{\mu}{\sigma^2}+\frac{x}{h_n^2}\right)^2\\&=\frac{x^2}{h_n^2}+\frac{\mu^2}{\sigma^2}-\frac{h_n^2\mu^2}{\sigma^2(\sigma^2+h_n^2)}-\frac{2\mu x}{\sigma^2+h_n^2}-\frac{\sigma^2x^2}{{h_n}^2(\sigma^2+h_n^2)}\\&=\frac{\mu^2(\sigma^2+h_n^2)-h_n^2\mu^2}{\sigma^2(\sigma^2+h_n^2)}+\frac{x^2(\sigma^2+h_n^2)-\sigma^2x^2}{h_n^2(\sigma^2+h_n^2)}-\frac{2\mu x}{\sigma^2+h_n^2}\\&=\frac{\mu^2}{\sigma^2+h_n^2}+\frac{x^2}{\sigma^2+h_n^2}-\frac{2\mu x}{\sigma^2+h_n^2}=\frac{(\mu-x)^2}{\sigma^2+h_n^2}\end{aligned}$$

因此，将上式代入 $\bar{p}_n(x)$，有如下公式成立。

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

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

5.2 第二小问

$p_n(x)$ 的方差可以表示为如下形式。

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

其中，$\mathrm{E}\varphi^2\left(\frac{x-v}{h_n}\right)$ 可以表示为如下形式。

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

由第一小问得出的结论可知，$\int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}h_n/\sqrt{2}} \exp\left[-\frac{1}{2}\left(\frac{x-v}{h_n/\sqrt{2}}\right)^2\right] \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-v}{\sigma}\right)^2\right] dv \sim N(\mu, \sigma^2 + h_n^2/2)$ ，因此，$\mathrm{E}\varphi^2\left(\frac{x-v}{h_n}\right)$ 可以表示为下式。

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

当 h_n 比较小的时候，$\sigma^2 + h_n^2/2 \simeq \sigma^2$。因此，$\mathrm{E}\varphi^2\left(\frac{x-v}{h_n}\right)$ 可以写成如下形式。

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

由第一问可知，有如下公式成立。

$$\bar{p}_n(x) = \frac{1}{V_n}\mathrm{E}\left[\varphi(\frac{x-v}{h_n})\right] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right]$$

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

因此，$\left[\mathrm{E}\varphi\left(\frac{x-v}{h_n}\right)\right]^2$ 可以表示为如下形式。

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

$p_n(x)$ 的方差可最终表示为如下形式。

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

其中，$\frac{1}{2\pi n\sigma^2} \exp\left[-\frac{(x-\mu)^2}{\sigma^2}\right] \simeq 0$，因此 $\mathrm{Var}\left[p_n(x)\right] = \frac{p(x)}{2nh_n\sqrt{\pi}}$ 成立。

6　问题 6

6.1　第一小问

假设 n 个点均匀分布在 d 维超空间内。那么平均每个维度上的点数为 $n^{1/d}$，那么在 r 维空间内存在的点数可以表示为 $n^{r/d}$。由于空间内点是均匀分布的，因此当最近邻点落在 r 维

空间的时候可以使得部分维度距离达到最近邻距离。因此，利用部分距离达到最近邻的概率可以表示为 $Pr = \frac{n^{r/d}}{n} = n^{r/d-1}$。

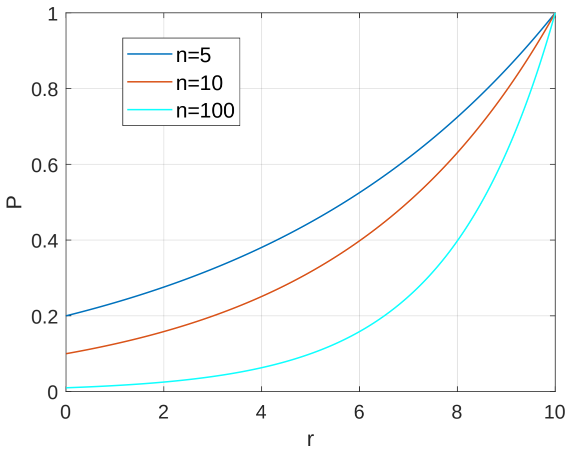

Figure 2: 部分距离达到最近邻概率.

6.2 第二小问

一维情况下的贝叶斯决策边界可以表示为 $x_1^* = 0.5$，二维下的决策边界可以表示为 $x_2^* = 1 - x_1$，三维情况下的决策边界满足下列等式。最终效果如图 Figure 3 所示

$$
\begin{aligned}
& x_1x_2x_3 = (1-x_1)(1-x_2)(1-x_3) \\
& \Rightarrow x_1x_2 = (1-x_1)(1-x_2)(\frac{1}{x_3}-1) \\
& \Rightarrow \frac{1}{x_3} = \frac{x_1x_2}{(1-x_1)(1-x_2)} + 1 \\
& \Rightarrow \frac{1}{x_3} = \frac{x_1x_2 + (1-x_1)(1-x_2)}{(1-x_1)(1-x_2)} \\
& \Rightarrow x_3 = \frac{(1-x_1)(1-x_2)}{1-x_1-x_2+2x_1x_2}
\end{aligned}
$$

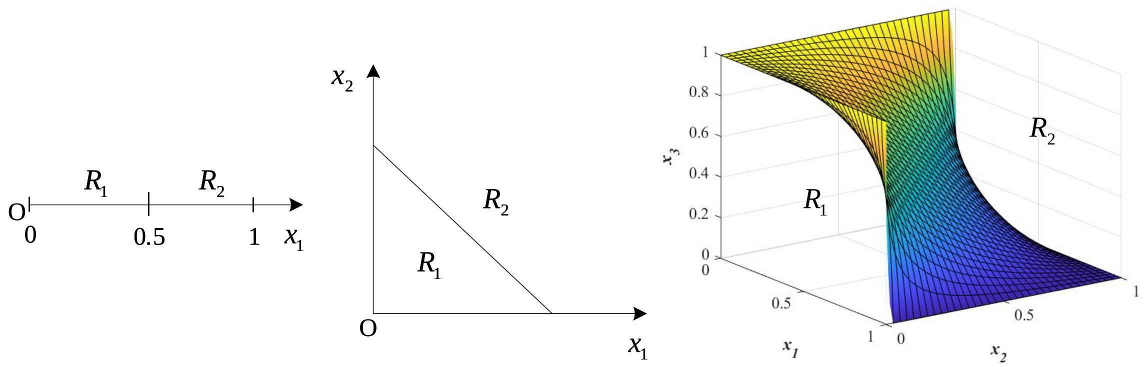

Figure 3: 贝叶斯决策边界.