模式识别
作业1

中国科学院大学
University of Chinese Academy of Sciences

2020E8017782032
蒲尧

Question 1.

证明：$\because P(w_i|\boldsymbol{x}) \geq P(w_j|\boldsymbol{x})$ 所以在 w_i 有最大后验概率.

(1) ① 若在特征集 $\boldsymbol{x}$，则判别非 i 类而所有损失为：$\lambda_s \sum_{j\neq i} P(w_j|x) = \lambda_s[1-P(w_i|\boldsymbol{x})]$

若选为非i类的k类，风险为 $\lambda_s \sum_{j\neq k} P(w_j|\boldsymbol{x}) = \lambda_s[1-P(w_k|\boldsymbol{x})] \geq \lambda_s[1-P(w_i|\boldsymbol{x})]$

$\therefore$ 在非拒绝情况下，判为 w_i 风险损失最小.

② 若 $\lambda_r \leq \lambda_s[1-P(w_i|\boldsymbol{x})]$ 即 $P(w_i|\boldsymbol{x}) \leq 1-\frac{\lambda_r}{\lambda_s}$ 则拒绝风险最小.

综上：在 $P(w_i|x) \geq P(w_j|x)$ 且 $P(w_i|x) \geq 1-\lambda_r/\lambda_s$ 判为 w_i，否则拒绝。

(2) 若 $\lambda_r = 0$ 则拒绝情况下的风险也为0，即对任意 j，判 w_i 或拒绝均可.

(3) 若 $\lambda_r > \lambda_s$ 则恒有 $P(w_i|x) > 1-\frac{\lambda_r}{\lambda_s}$，则判为 w_i 风险最小，不会拒绝.

Question 2

(a) $$P_e = P(|x-\mu_1| > |x-\mu_2| \,|\, w_1)P(w_1) + P(|x-\mu_1| < |x-\mu_2| \,|\, w_2)P(w_2)$$

$$= p(x > \frac{\mu_1+\mu_2}{2} | w_1)P(w_1) + p(x < \frac{\mu_1+\mu_2}{2} | w_2)P(w_2)$$

$$= p(x > \frac{\mu_1+\mu_2}{2} | w_1)$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{\frac{\mu_1+\mu_2}{2}}^{\infty} e^{-\frac{1}{2}(\frac{x-\mu_1}{\sigma})^2} dx \quad (令\ u = \frac{x-\mu_1}{\sigma})$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\frac{|\mu_2-\mu_1|}{2\sigma}}^{\infty} e^{-\frac{u^2}{2}} \cdot u' du \quad (u' = \frac{1}{\sigma})$$

$$= \frac{1}{\sqrt{2\pi}}\int_a^{\infty} e^{-\frac{u^2}{2}} du \quad (a = \frac{|\mu_2-\mu_1|}{2\sigma})$$

(b) 当 $\frac{|\mu_1-\mu_2|}{\sigma} \to \infty$ 则 $a \to \infty$，$\frac{1}{\sqrt{2\pi}a} e^{-\frac{a^2}{2}} \to 0$

$\because P_e = \frac{1}{\sqrt{2\pi}}\int_a^{\infty} e^{-\frac{u^2}{2}} du \leq \frac{1}{\sqrt{2\pi}}\int_a^{\infty} \frac{u}{a} e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}a} e^{-\frac{a^2}{2}}$ $\therefore P_e \to 0$

Question 3

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

$$\mu = \mathcal{E}[x] = \int x p(x) dx$$

年 月 日

$$\Sigma = \mathcal{E}[(x-\mu)(x-\mu)^t] = \int (x-\mu)(x-\mu)^t p(x) dx$$

(b) (a) Σ_i = 任意，$g_i(x)=x^t W_i x + w_i^t x + w_{i0}$

$$W_i = -\frac{1}{2}\Sigma_i^{-1} \quad w_i = \Sigma_i^{-1}\mu_i$$

$$w_{i0} = -\frac{1}{2}\mu_i^t \Sigma_i^{-1}\mu_i - \frac{1}{2}\ln|\Sigma_i| + \ln P(w_i)$$

(b) $\Sigma_i = \Sigma$ $\quad g_i(x) = -\frac{1}{2}(x-\mu_i)^t \Sigma^{-1}(x-\mu_i) + \ln P(w_i)$

$$g_i(x) = w_i^t x + w_{i0}$$

$$w_i^t = \Sigma^{-1}\mu_i$$

$$w_{i0} = -\frac{1}{2}\mu_i^t \Sigma^{-1}\mu_i + \ln P(w_i)$$

(c) 求矩阵的伪逆、特征降维、扩大非相关样本等.

Question 4

判别面 $x_0 = \frac{1}{2}(\mu_1+\mu_2) - \frac{\ln[P(w_1)/P(w_2)]}{(\mu_1-\mu_2)^t \Sigma^{-1}(\mu_1-\mu_2)}(\mu_1-\mu_2)$ 假设 $\mu_1 < \mu_2$

当 $x_0 < \mu_1$ 时 = $\ln[P(w_1)/P(w_2)] > -\frac{1}{2}(\mu_1-\mu_2)^2\Sigma^{-1}$

当 $x_0 > \mu_2$ 时 = $\ln[P(w_1)/P(w_2)] < \frac{1}{2}(\mu_2-\mu_2)^2\Sigma^{-1}$

Question 5

依题 $z_{ik} = \begin{cases} 1 & \text{第k个样本的自然状态为} w_i \\ 0 & \text{其他} \end{cases}$

(a) $P_r[z_{ik}=1|P(w_i)] = P(w_i)$ $\quad$ 统一为 $P[z_{ik}|P(w_i)] = [P(w_i)]^{z_{ik}}[1-P(w_i)]^{1-z_{ik}}$

$P_r[z_{ik}=0|P(w_i)] = 1-P(w_i)$

连续独立抽取 n 个后：

$$P[z_{i1},z_{i2},\cdots,z_{in}|P(w_i)] = \prod_{k=1}^{n} P[z_{ik}|P(w_i)] = \prod_{k=1}^{n}[P(w_i)]^{z_{ik}}[1-P(w_i)]^{1-z_{ik}}$$

(b) $P(w_i)$ 的对数似然函数为

$$l(P(w_i)) = \ln\{P[z_{i1},\cdots,z_{in}|P(w_i)]\} = \ln\{\prod_{k=1}^{n}[P(w_i)]^{z_{ik}}[1-P(w_i)]^{1-z_{ik}}\}$$

$$= \sum_{k=1}^{n}\{z_{ik}\ln[P(w_i)] + (1-z_{ik})\ln[1-P(w_i)]\}$$

求最大似然估计 则 $\nabla_{P(w_i)} l(P(w_i)) = \frac{1}{P(w_i)}\sum_{k=1}^{n} z_{ik} - \frac{1}{1-P(w_i)}\sum_{k=1}^{n}(1-z_{ik}) = 0$

2020E8017782032
蒲尧

即 $\frac{1}{\hat{P}(w_i)}\sum_{k=1}^{n} z_{ik} = \frac{1}{1-\hat{P}(w_i)}\sum_{k=1}^{n}(1-z_{ik})$

$$(1-\hat{P}(w_i))\sum_{k=1}^{n} z_{ik} = n\hat{P}(w_i) - \hat{P}(w_i)\sum_{k=1}^{n} z_{ik}$$

$$\hat{P}(w_i) = \frac{1}{n}\sum_{k=1}^{n} z_{ik}$$

Question 6.

(a) $$\hat{\mu}_{n+1} = \frac{1}{n+1}\sum_{k=1}^{n+1} x_k = \frac{1}{n+1}\left(\sum_{k=1}^{n} x_k + x_{n+1}\right) = \frac{n}{n+1}\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) + \frac{1}{n+1}x_{n+1}$$

$$= \frac{n+1-1}{n+1}\hat{\mu}_n + \frac{1}{n+1}x_{n+1} = \hat{\mu}_n + \frac{1}{n+1}(x_{n+1}-\hat{\mu}_n)$$

$$C_{n+1} = \frac{1}{n}\sum_{k=1}^{n+1}(x_k-\hat{\mu}_{n+1})(x_k-\hat{\mu}_{n+1})^T$$

$$= \frac{1}{n}\sum_{k=1}^{n}(x_k-\hat{\mu}_{n+1})(x_k-\hat{\mu}_{n+1})^T + \frac{1}{n}\left[x_{n+1}-\hat{\mu}_n-\frac{1}{n+1}(x_{n+1}-\hat{\mu}_n)\right]\left[x_{n+1}-\hat{\mu}_n-\frac{1}{n+1}(x_{n+1}-\hat{\mu}_n)\right]^T$$

$$= \frac{1}{n}\sum_{k=1}^{n}\left[x_k-\hat{\mu}_n-\frac{1}{n+1}(x_{n+1}-\hat{\mu}_n)\right]\left[x_k-\hat{\mu}_n-\frac{1}{n+1}(x_{n+1}-\hat{\mu}_n)\right]^T$$

$$+\frac{n}{(n+1)^2}(x_{n+1}-\hat{\mu}_n)(x_{n+1}-\hat{\mu}_n)^T$$

$$= \frac{1}{n}\left[\sum_{k=1}^{n}(x_k-\hat{\mu}_n)(x_k-\hat{\mu}_n)^T - \frac{1}{n+1}\underbrace{\sum_{k=1}^{n}(x_k-\mu_n)}_{=0}(x_{n+1}-\hat{\mu}_n)^T - \frac{1}{n+1}(x_{n+1}-\hat{\mu}_n)\underbrace{\sum_{k=1}^{n}(x_k-\hat{\mu}_n)^T}_{=0}\right.$$

$$\left.+\frac{1}{(n+1)^2}\sum_{k=1}^{n}(x_{n+1}-\hat{\mu}_n)(x_{n+1}-\hat{\mu}_n)^T\right] + \frac{n}{(n+1)^2}(x_{n+1}-\hat{\mu}_n)(x_{n+1}-\hat{\mu}_n)^T$$

$$= \frac{n-1}{n}\cdot\frac{1}{n-1}\sum_{k=1}^{n}(x_k-\hat{\mu}_n)(x_k-\hat{\mu}_n)^T + \left[\frac{1}{n}\cdot n\cdot\frac{1}{(n+1)^2}+\frac{n}{(n+1)^2}\right](x_{n+1}-\hat{\mu}_n)(x_{n+1}-\hat{\mu}_n)^T$$

$$= \frac{n-1}{n}C_n + \frac{1}{n+1}(x_{n+1}-\hat{\mu}_n)(x_{n+1}-\hat{\mu}_n)^T$$

(b) 对于 $\hat{\mu}_n$ 复杂度 $O(dn)$；对 C_n 复杂度 $O(dn^2)$