Pattern Recognition (Fall 2026) August 27, 2026

Assignment 1: Decision Theory and Bayes Error

Wei Chen
2026013579
Department of Computer Science
wchen@example.edu

Problems

Problem: 无标号问题

在一个 10 类的模式识别问题中，有 3 类单独满足多类情况 1，其余的类别满足多类情况 2. 问该模式识别问题所需判别函数的最少数目是多少？

Answer: 每类满足多类情况 1 需要 1 个判别函数，共需 3 个；其余 7 类满足多类情况 2，两两之间需要判别函数，共需 $\binom{7}{2} = 21$ 个. 因此所需判别函数的最少数目为 $3 + 21 = 24$ 个.

Problem 1

Consider the following decision rule for a two-category one-dimensional problem: Decide ω_1 if $x > \theta$; otherwise decide ω_2. (a) Show the probability of error for this rule is given by

$$P(\text{error}) = P(\omega_1) \int_{-\infty}^{\theta} p(x|\omega_1)\,\mathrm{d}x + P(\omega_2) \int_{\theta}^{\infty} p(x|\omega_2)\,\mathrm{d}x.$$

(b) By differentiating, show that a necessary condition to minimize $P(\text{error})$ is that satisfy

$$p(\theta|\omega_1)P(\omega_1) = p(\theta|\omega_2)P(\omega_2).$$

(c) Does this equation define θ uniquely?

Answer: 考虑分类错误率的条件概率：

$$P(\text{error}|x) = \begin{cases} p(\omega_1|x), & \text{if we decide } \omega_2, \\ p(\omega_2|x), & \text{otherwise.} \end{cases} \tag{P1.1}$$

根据式 (P1.1) 对 x 进行积分，可得：

$$\begin{aligned} P(\text{error}) &= \int_{-\infty}^{\infty} P(\text{error}|x)p(x)\mathrm{d}x && \text{(P1.2)} \\ &= P(x \le \theta, x \text{ is } \omega_1) + P(x > \theta, x \text{ is } \omega_2) && \text{(P1.3)} \\ &= p(x \le \theta|\omega_1)P(\omega_1) + p(x > \theta|\omega_2)P(\omega_2) && \text{(P1.4)} \\ &= P(\omega_1) \int_{-\infty}^{\theta} p(x|\omega_1)\,\mathrm{d}x + P(\omega_2) \int_{\theta}^{\infty} p(x|\omega_2)\,\mathrm{d}x. && \text{(P1.5)} \end{aligned}$$

对 (c)：不一定唯一，该方程仅给出一个必要条件（驻点条件），若 $p(x|\omega_1)P(\omega_1)$ 与 $p(x|\omega_2)P(\omega_2)$ 的图像有多个交点，则方程可能有多个解.

Computer Exercises

Several of the computer exercises will rely on the following data.

此处还可以插入一些说明.

Computer Exercise 1

Illustrate the fact that the average of a large number of independent random variables will approximate a Gaussian by the following: (a) Write a program to generate n random integers from a uniform distribution $U(x_l, x_u)$. (b) Generate and plot a histogram of the accumulation of 10^4 points sampled as described above. (c) Calculate the mean and standard deviation of your histogram, and plot it.

Answer: 答案写在此处,如代码 1 所示.

```python
import random

def sample_uniform_sum(n, x_l, x_u, trials=10000):
    return [sum(random.uniform(x_l, x_u) for _ in range(n))
            for _ in range(trials)]
```

Listing 1: 代码测试

为了得到 $p(x|\omega) \sim \mathcal{N}(0,1)$:

$$p(x|\omega_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \tag{C1.1}$$

对式 (C1.1),随着 n 增大,样本均值的分布将根据中心极限定理趋近于高斯分布.