Assignment 1

2020 年 9 月 28 日

说明

- 作业用中文撰写，鼓励使用LaTex.
- 文档按“学号_姓名.pdf”命名提交.
- 本次作业截止时间为2020年10月11日，请到课程网站及时提交。

Question 1

In many pattern classification problems, one has the option either to assign the pattern to one of c classes, or to reject it as being unrecognizable. If the cost of rejection is not too high, rejection may be a desirable action. Let

$$\lambda(\alpha_i|w_j) = \begin{cases} 0, & i = j \quad i, j = 1, \ldots, c \\ \lambda_r, & i = c + 1 \\ \lambda_s, & otherwise \end{cases}$$

where λ_r is the loss incurred for choosing the $(c + 1)$-th action, rejection, and λ_s is the loss incurred for making a substitution error. Show that the minimum risk is obtained if we decide w_i if $P(w_i|x) \geq P(w_j|x)$ for all j and if $P(w_i|x) \geq 1 - \frac{\lambda_r}{\lambda_s}$, and reject otherwise. What happens if $\lambda_r = 0$? And what happens if $\lambda_r > \lambda_s$?

Question 2

Let $p(x|w_i) \sim \mathcal{N}(\mu_i, \sigma^2)$ for a two-category one-dimensional problem with $p(w_1) = p(w_2) = 0.5$.

(a) Show that the minimum probability of error is given by

$$P_e = \frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-\frac{\mu^2}{2}} d\mu,$$

where $a = \frac{|\mu_1 - \mu_2|}{2\sigma}$

(b) Use the inequality

$$P_e \leq \frac{1}{\sqrt{2\pi} a} e^{-\frac{a^2}{2}}$$

to show that P_e goes to zero as $\frac{|\mu_1 - \mu_2|}{\sigma}$ goes to infinity.

Question 3

To classify a feature vector $x \in \mathcal{R}^d$ in a task of c classes, we assume that for each class, the prior is same and the class conditional probability density is a Gaussian distribution.

(a) Write the mathematical form of the conditional probability density function

(b) Write the discriminant function of minimum error rate in the following two cases: (a) class covariance matrices are unequal; (b) class covariance matrices are same.

(c) For the quadratic discriminant function based on Gaussian probability density, it becomes incalculable when the covariance matrix is singular. Name two ways to overcome the singularity.

Question 4

Suppose we have two normal distributions with the same covariance but different means: $\mathcal{N}(\mu_1, \Sigma)$ and $\mathcal{N}(\mu_2, \Sigma)$. In terms of their prior probabilities $P(w_1)$ and $P(w_2)$, state the condition that Bayes decision boundary does not pass between the two means.

Question 5

Maximum likelihood methods apply to estimate of prior probability as well. Let samples be drawn by successive, independent selections of a state of nature w_i with unknown probability $P(w_i)$. Let $z_{ik} = 1$ if the state of nature for the k-th sample is w_i and $z_{ik} = 0$ otherwise.

(a) Show that

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

(b) Show that the maximum likelihood estimate for $P(w_i)$ is

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

Question 6

Let the sample mean $\hat{\mu}_n$ and the sample covariance matrix C_n for a set of n d-dimensional samples $x_1, \ldots, x_n$ be defined by

$$\hat{\mu}_n = \frac{1}{n} \sum_{k=1}^{n} x_k, \qquad C_n = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \hat{\mu}_n)(x_k - \hat{\mu}_n)^T.$$

(a) Show that alternative, recursive techniques for calculating $\hat{\mu}_n$ and C_n based on the successive addition of new samples x_{n+1} can be derived using the recursion relations

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

and

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

(b) Discuss the computational complexity of finding $\hat{\mu}_n$ and C_n by the recursive methods.