Answer Key for Practice Quiz - 1

CS 591Q/791V - Pattern Recogniton
Instructor: Dr. Arun Ross
Posted on: Mar 8, 2010

1. a. Bayes Risk: [pp. 24-25] Consider the problem of designing a pattern classification system where the goal is to assign a d-dimensional feature vector $\boldsymbol{x}$ to one of c classes, $\omega_1, \ldots \omega_c$, resulting in one of a actions, $\alpha_1, \ldots \alpha_a$. Denote as $\lambda(\alpha_i|\omega_j)$ the cost of taking action α_i when the true class of the input pattern $\boldsymbol{x}$ is ω_j. Then the conditional risk of taking action α_i upon encountering feature vector $\boldsymbol{x}$ is

$$R(\alpha_i|\boldsymbol{x}) = \sum_{j=1}^{c} \lambda(\alpha_i|\omega_j) P(\omega_j|\boldsymbol{x}),$$

for $i = 1, \ldots a$. Thus, the decision policy would be to select that action α_i for which $R(\alpha_i|\boldsymbol{x})$ is minimum. The resulting minimum overall risk is called the Bayes Risk.

b. Minimum Distance Classifier: [p. 39] Given a feature vector $\boldsymbol{x}$ and a set of mean vectors $\boldsymbol{\mu_1}, \ldots \boldsymbol{\mu_c}$ corresponding to c classes, $\omega_1, \ldots \omega_c$, the minimum distance classifier assigns x to that class whose mean vector is located at the smallest Euclidean distance from $\boldsymbol{x}$. Thus,

$$\text{Assign } \boldsymbol{x} \text{ to class } \omega_{j^*}, \text{ where } j^* = \arg \min_{j=1,\ldots c} ||\boldsymbol{x} - \boldsymbol{\mu_j}||.$$

2. [see Figure 2.13 in p. 41 to view a similar problem]

The decision boundary can be computed as:

$$\begin{aligned}
&P(\omega_1|x) = P(\omega_2|x) \\
&\Rightarrow p(x|\omega_1|) = p(x|\omega_2), \qquad \text{(since } P(\omega_1) = P(\omega_2)\text{)} \\
&\Rightarrow \frac{1}{\sqrt{2\pi 1}} exp\left(-\frac{1}{2}x^2\right) = \frac{1}{\sqrt{2\pi 4}} exp\left(-\frac{1}{2}\frac{(x-\frac{1}{2})^2}{4}\right)
\end{aligned}$$

Applying ln on both sides and multiplying by -1, we have:

$$\begin{aligned}
&\frac{x^2}{2} = \ln 2 + \frac{(x-\frac{1}{2})^2}{8} \\
&\Rightarrow \frac{x^2}{2} = \ln 2 + \frac{1}{8}\left(x^2 - x + \frac{1}{4}\right) \\
&\Rightarrow 4x^2 = 8\ln 2 + (x^2 - x + \frac{1}{4}) \\
&\Rightarrow 3x^2 + x - 5.79 = 0 \\
&\Rightarrow x = -1.56 \quad \text{or} \quad x = 1.23
\end{aligned}$$

Thus the decision rule will be,

$$\begin{cases} x \in \omega_1, & \text{if } -1.56 < x < 1.23, \\ x \in \omega_2, & \text{otherwise.} \end{cases}$$

3. [see section 2.7]

Note that $\omega_{max}(\boldsymbol{x})$ denotes the class that $\boldsymbol{x}$ is assigned to by the Bayes classifier. It could be one of $\omega_1, \ldots \omega_c$ depending on the feature vector $\boldsymbol{x}$. So, $P(correct|\boldsymbol{x}) = P(\omega_{max}|\boldsymbol{x})$. Thus,

$$
\begin{aligned}
& P(correct|\boldsymbol{x}) = P(\omega_{max}|\boldsymbol{x}) \\
& \Rightarrow P(correct) = \int P(\omega_{max}|\boldsymbol{x})p(\boldsymbol{x})dx \\
& \Rightarrow P(error) = 1 - \int P(\omega_{max}|\boldsymbol{x})p(\boldsymbol{x})dx. \square
\end{aligned}
$$

4. [see section 2.5.1]

The entropy of a distribution $p(x)$ is given by

$$H(p(x)) = -\int p(x)\ln[p(x)]\ dx,$$

where the integral is in the interval $[-\infty, \infty]$. In this case,

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

and

$$ln[p(x)] = -\ln[\sqrt{(2\pi\sigma^2)}] - \frac{x^2}{2\sigma^2} = -(\ln[\sqrt{(2\pi\sigma^2)}] - \frac{x^2}{2\sigma^2}).$$

Thus,

$$
\begin{aligned}
H(p(x)) &= -\int p(x)\ln[p(x)]dx \\
&= \int p(x)(\ln[\sqrt{(2\pi\sigma^2)}] + \frac{x^2}{2\sigma^2})dx \\
&= \int p(x)\ln[\sqrt{(2\pi\sigma^2)}]dx + \int p(x)\frac{x^2}{2\sigma^2}dx \\
&= \ln[\sqrt{(2\pi\sigma^2)}] \int p(x)dx + \frac{1}{2\sigma^2}\int x^2 p(x)dx \\
&= \ln[\sqrt{(2\pi\sigma^2)}] + \frac{1}{2\sigma^2}\sigma^2 \\
& \text{(since } \int p(x))dx = 1, \text{ and } \int x^2 p(x)dx \text{ is the formula for the variance of a 0-mean gaussian distribution)} \\
&= \ln[\sqrt{(2\pi\sigma^2)}] + \frac{1}{2} \\
&= \ln[\sqrt{(2\pi\sigma^2)}] + \frac{1}{2}\ln(e) \\
&= \ln[\sqrt{(2e\pi\sigma^2)}]. \square \\
& \text{(since } \frac{1}{2}\ln(e) = \ln(\sqrt{e}))
\end{aligned}
$$