AI & ML Course
Quiz 1(Feb 5, 2024)

Time: 30 minutes

Instructions

- Answer all questions
- See upload instructions in the form

Question:	1	2	3	Total
Points:	15	5	0	20
Score:				

Read Carefully: In the following we will use the following notations. (X, Y) be a random instance drawn from a distribution $\mathcal{P}$ where $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$.

$$h : \mathcal{X} \subseteq \mathbb{R}^d \to \mathcal{Y}$$

will denote a classifier. For Binary classification we will use the following notation,

$$\mathcal{Y} = \{-1, 1\}, P_1(x) = P(X = x|Y = 1), P_2(x) = P(X = x|Y = -1), p_1 = P(Y = 1), p_2 = P(Y = -1).$$

For multicategory classification we will use the following notation

$$\mathcal{Y} = \{1, \ldots, K\}, P_i(x) = P(X = x|Y = i), p_i = P(Y = i), i \in \{1, \ldots, K\}.$$

$\mathbf{I}$ will denote the identity matrix,dimension will be clear from the context. $N(\mathbf{x}|\mu, C)$ is as defined in the class. For any $\mathbf{x} \in \mathbb{R}^d$, $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^d x_i^2}$

1. Consider the problem of Binary Classification. We are given the following information

$$\mathcal{X} \subseteq \mathbb{R}^2,\ P_1(\mathbf{x}) = N(\mathbf{x}|\mu_1, \sigma^2\mathbf{I}),\ P_2(\mathbf{x}) = N(\mathbf{x}|\mu_2, k\sigma^2\mathbf{I}),\ \mu_1 = [a, a]^\top, \mu_2 = -\frac{1}{\sqrt{3}}\mu_1, p_1 = \frac{1}{k+1}$$

The values of σ^2, k, a are unknown but it is given that they are all positive. It is claimed that the Bayes Classifer for this problem reduces to

$$h(\mathbf{x}) = \text{ sign } (f(\mathbf{x})),\ f(\mathbf{x}) = \|\mathbf{x}\|^2 + \mathbf{w}^\top\mathbf{x}$$

The value of $\|\mathbf{w}\| = \sqrt{2} + \sqrt{6}$. Based on the information provided, answer the following

Solution: The Bayes classifier is

$$h(\mathbf{x}) = \text{ sign } (f(\mathbf{x}))\ f(\mathbf{x}) = \log\left(\frac{N(\mathbf{x}|\mu_1, \sigma^2)}{N(\mathbf{x}|\mu_2, k\sigma^2)}\frac{p_1}{p_2}\right)$$

$$f(\mathbf{x}) = -\frac{1}{2\sigma^2}\|\mathbf{x} - \mu_1\|^2 + \frac{1}{2k\sigma^2}\|\mathbf{x} - \mu_2\|^2 - \frac{1}{2}\log\frac{(2\pi\sigma^2)^2}{(2\pi k\sigma^2)^2} - \log k$$

where we have used $\frac{p_1}{p_2} = \frac{p_1}{1-p_1} = \frac{1}{k}$. Thus

$$f(\mathbf{x}) = \frac{1}{2\sigma^2}\left(\frac{1}{k} - 1\right)\|\mathbf{x}\|^2 + \frac{1}{\sigma^2}(\mu_1 - \frac{1}{k}\mu_2)^\top\mathbf{x} + \frac{1}{\sigma^2}\left(\frac{1}{k}\|\mu_2\|^2 - \|\mu_1\|^2\right) + \log k - \log k$$

Equating terms:

$$\frac{1}{k}\|\mu_2\|^2 - \|\mu_1\|^2 = 0 \implies \quad k = \frac{\|\mu_2\|^2}{\|\mu_1\|^2} = \frac{1}{3}$$

$$\frac{1}{2\sigma^2}\left(\frac{1}{k} - 1\right) = 1, \quad \implies \sigma^2 = 1$$

Since $\sigma^2 = 1$,

$$\mathbf{w} = \mu_1 - \frac{1}{k}\mu_2 = \mu_1 + 3\frac{1}{\sqrt{3}}\mu_1 = (1 + \sqrt{3})\mu_1$$

Since a is positive,

$$\|\mathbf{w}\| = (1 + \sqrt{3})\sqrt{2}a = \sqrt{2} + \sqrt{6}$$

and hence $a = 1$.

It is to noted that the above derivation holds for

$$f(\mathbf{x}) = LLR(\mathbf{x}) = \log\left(\frac{N(\mathbf{x}|\mu_1, \sigma^2)}{N(\mathbf{x}|\mu_2, k\sigma^2)}\frac{p_1}{p_2}\right)$$

As per the problem,

$$LLR(\mathbf{x}) = \frac{1}{\sigma^2}\left(\frac{1}{2}\left(\frac{1}{k} - 1\right)\|\mathbf{x}\|^2 + (\mu_1 - \frac{1}{k}\mu_2)^\top \mathbf{x}\right)$$

However $f(\mathbf{x}) = cLLR(\mathbf{x})$ is also a Bayes classifier for any $c > 0$. Thus any choice of $\sigma^2 > 0$ will satisfy all the conditions of the problem and hence it is not unique.

(a) (5 points) The value of k is

○ $\frac{4}{3}$ ○ 1 √ $\frac{1}{3}$ ○ 0.75

(b) (5 points) The value of σ^2 is

○ $\frac{4}{3}$ √ 1 ○ 0.75 ○ 0.6

(c) (5 points) The value of a is

○ 1.5 ○ 3 ○ $\frac{2}{3}$ √ 1

2. (5 points) For the above problem, instead of the Bayes classifier we wish to determine a linear classifier of the form

$$h(\mathbf{x}) = \text{ sign } (\mathbf{w}^\top \mathbf{x}).$$

where $\mathbf{w}$ is determined through Fisher Discriminant. The value of $\hat{\mathbf{w}} = \frac{\mathbf{w}}{\|\mathbf{w}\|}$ is

○ $[\frac{\sqrt{2}}{\sqrt{3}}, \frac{1}{\sqrt{3}}]^\top$ ○ $[\frac{1}{\sqrt{3}}, \frac{\sqrt{2}}{\sqrt{3}}]^\top$ ○ $[0.6, 0.8]^\top$ √ $[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^\top$

Solution:

Fisher solution is

$$\mathbf{w} = (k\sigma^2\mathbf{I} + \sigma^2\mathbf{I})^{-1}(\mu_1 - \mu_2)$$

Using

$$\mu_1 - \mu_2 = (1 + \frac{1}{\sqrt{3}})[a, a]^\top$$

we note that both coordinates of $\mathbf{w}$ are equal and hence $\hat{\mathbf{w}} = \frac{1}{\sqrt{2}}[1, 1]^\top$.

3. A hospital wants to design a Genome based therapy for an unknown Disease `DISEASE`. A set of 3 genes were identified to be candidates for therapy. Since the underlying biology is highly uncertain, the first step in the therapy is to understand if the identified genes can predict if a person is healthy or diseased.

Let $X \in \{0, 1\}^3$ be a random vector denoting the expression levels of the three genes where X_i corresponds to the expression level of the ith gene. The random variable Y takes value 1 for `DISEASE` and -1 for Healthy. Let $\hat{Y}(X)$ be the prediction on a random instance of X. The loss incurred is $l(\hat{Y}(X), Y)$ where Y is the right label. It is given that

$l(1, -1) = 1, l(-1, 1) = 3, l(-1, -1) = l(1, 1) = 0$. Consider the predictor

$$\hat{Y}(X) = \begin{cases} 1 & log\ \frac{P(X=x|Y=1)}{P(X=x|Y=-1)} \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

It is assumed that

$$P(X = x|Y = y) = \Pi_{i=1}^{3} P(X_i = x_i|Y = y)$$

Furthermore,

$$P(X_i = 1|Y = 1) = q_i^{(1)}, P(X_i = 1|Y = -1) = q_i^{(2)}, i \in \{1, 2, 3\}$$

It is given that $q^{(1)} = [\frac{1}{3}, \frac{1}{2}, \frac{3}{4}]^\top$ and $q^{(2)} = [\frac{1}{3}, \frac{1}{3}, \frac{1}{4}]^\top$ For what values of $\mathbf{w}, p_1, p_2$ does

$$\hat{Y}(\mathbf{x}) = sign(\mathbf{w}^\top \mathbf{x})$$

is the Bayes classifier.

Solution:

$$\hat{Y}(\mathbf{x}) = sign(f(\mathbf{x})), \quad f(\mathbf{x}) = \sum_{i=1}^{3} \log \frac{P(X_i = x_i|Y = 1)}{P(X_i = x_i|Y = -1)}$$

$$f(\mathbf{x}) = \sum_{i=1}^{3} \left(x_i \log \frac{q_i^{(1)}}{q_i^{(2)}} + (1 - x_i) \log \frac{1 - q_i^{(1)}}{1 - q_i^{(2)}} \right)$$

$$f(\mathbf{x}) = \sum_{i=1}^{3} \left(x_i \left(\log \frac{q_i^{(1)}}{q_i^{(2)}} \frac{(1 - q_i^{(2)})}{(1 - q_i^{(1)})} \right) + \log \left(\frac{(1 - q_i^{(1)})}{(1 - q_i^{(2)})} \right) \right)$$

$$f(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + \log \left(\frac{(1 - q_i^{(1)})}{(1 - q_i^{(2)})} \right)$$

$$\mathbf{w}_1 = 0, \mathbf{w}_2 = \log \frac{\frac{1}{2}}{\frac{1}{3}} \frac{(1 - \frac{1}{3})}{(1 - \frac{1}{2})} = \log 2, \ \mathbf{w}_3 = \log \frac{\frac{3}{4}(1 - \frac{1}{4})}{\frac{1}{4}(1 - \frac{3}{4})} = \log 9$$

The Bayes classifier for the problem is

$$h_B(\mathbf{x}) = sign(f_b(\mathbf{x})), \quad f_b(\mathbf{x}) = log \frac{P(X = x|Y = 1)}{P(X = x|Y = -1)} (\frac{p_2 l(1, -1)}{p_1 l(-1, 1)})$$

$$f_b(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + \sum_{i=1}^{3} \log \frac{1 - q_i^{(1)}}{1 - q_i^{(2)}}$$

$$\sum_{i=1}^{3} \log \frac{1 - q_i^{(1)}}{1 - q_i^{(2)}} + \log \frac{p_2}{3p_1} = \log 3/4 + \log 1/3 + \log \frac{p_2}{3p_1} = 0$$

$$\sum_{i=1}^{3} \log \frac{1 - q_i^{(1)}}{1 - q_i^{(2)}} + \log \frac{p_2}{3p_1} = \log \frac{p_2}{12p_1} = 0$$

$$p_1 = \frac{1}{13}, \quad p_2 = \frac{12}{13}$$