01-02-22

Assignment 1

Q1. To find the decision boundary that minimizes the error in the given cases.

We can minimize the probability of error by the likelihood ratio.

If $\dfrac{P(x|\omega_1)}{P(x|\omega_2)} > \dfrac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}} \dfrac{P(\omega_2)}{P(\omega_1)}$

then decide ω_1 else decide ω_2.

(i) Zero-one loss

$\Rightarrow \lambda_{11} = 0 \ ; \ \lambda_{12} = 1 \ ; \ \lambda_{21} = 1 \ ; \ \lambda_{22} = 0$

$$RHS = \frac{1-0}{1-0} \frac{3}{4} \cdot 4 = \cancel{4} \ 3$$

$$LHS = \frac{P(x|\omega_1)}{P(x|\omega_2)} = \frac{1}{\sqrt{2\pi}} \frac{e^{-1/2(x-2)^2}}{e^{-1/2(x-5)^2}} \sqrt{2\pi}$$

$$= e^{-1/2(x-2)^2 + 1/2(x-5)^2}$$

$$= e^{-1/2[(x-2)^2-(x-5)^2]}$$

$$= e^{-1/2[(x-2+x-5)(x-2-x+5)]}$$

$$= e^{-1/2(2x-7)(3)}$$

Now, $\cancel{e^{-3/2(2x-7)} > 4}$ $\qquad e^{-3/2(2x-7)} > 3$

$\Rightarrow \cancel{-\frac{3}{2}(2x-7) > \ln 4}$ $\qquad -\frac{3}{2}(2x-7) > \ln(3)$

$\Rightarrow \cancel{-3(2x-7) > 2.77}$ $\qquad \Rightarrow -\frac{3}{2}(2x-7) > 1.0986$

$\Rightarrow \cancel{(2x-7) < -0.92}$

$\Rightarrow \cancel{x < \frac{7-0.92}{2}}$ $\qquad \Rightarrow 2x-7 < \frac{1.0986 \times 2}{-3}$

$\Rightarrow$

DATE
PAGE

~~$\Rightarrow x < 3.038$~~ $\Rightarrow x < 3.1338$

Ans. The decision boundary is as follows:
If $\boxed{x < 3.1338 \text{ decide } \omega_1}$
else decide ω_2. Decision boundary = $\boxed{3.1338}$

(ii) Given: $\lambda_{12} = 2$; $\lambda_{21} = 3$; $\lambda_{11} = 0$; $\lambda_{22} = 0$

$$RHS = \frac{2-0}{3-0} \, 3 = \frac{2}{3} \cdot 3 = 2$$

$$LHS \Rightarrow e^{-3/2(2x-7)} > 2$$

$$\Rightarrow \frac{-3}{2}(2x-7) > \ln 2$$

$$\Rightarrow 2x - 7 < -0.462$$

$$\Rightarrow x < 3.2689$$

Ans. The decision boundary is as follows:
If $\boxed{x < 3.2689 \text{ decide } \omega_1}$
else decide ω_2.
Decision boundary = $\boxed{3.2689}$

In tasks like cancer prediction, predicting no cancer when the patient already has cancer is more dangerous than predicting cancer for a no-cancer patient. In zero-one loss, the penalties for false positives is same as false negatives.

Hence, zero-one loss for such tasks should not be preferred considering ~~it~~ serious life implications.

DATE
PAGE

Q2. Given $Y = A^T X + B$

$$A = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$$

$$B = 5$$

$$Y = A^T X + B$$

$$E[y] = E[A^T X + B]$$

$\because$ $E[\cdot]$ is a linear operator { $E[ax+b] = aE[x]+b$ if a and b are scalars }

$$\Rightarrow E[y] = A^T E[x] + B$$

$$= \begin{bmatrix} 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -5 \\ 6 \end{bmatrix} + 5$$

$$= 27 + 5$$

$$= 32$$

Ans. $\boxed{32}$

04-02-2022

DATE
PAGE

Q3 (A) Given: $P(x|\omega_i) = \frac{1}{\pi b} \cdot \frac{1}{1+\left(\frac{x-a_i}{b}\right)^2}$, $i = 1, 2$

If $\frac{P(x|\omega_1)}{P(x|\omega_2)} > \frac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}} \frac{P(\omega_2)}{P(\omega_1)}$

~~Also, P(ω~~ then decide ω_1 else decide ω_2.

$\because \quad P(\omega_1) = P(\omega_2)$

$\Rightarrow$ If $\frac{P(x|\omega_1)}{P(x|\omega_2)} > \frac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}}$

then decide ω_1 else decide ω_2.

For zero-one loss $\lambda_{12} = 1, \lambda_{22} = 0, \lambda_{21} = 1, \lambda_{11} = 0$

The likelihood ratio becomes

If $P(x|\omega_1) > P(x|\omega_2)$
decide ω_1 else ω_2.

$$\Rightarrow \cancel{\frac{1}{\pi b}} \cdot \frac{1}{1+\left(\frac{x-a_1}{b}\right)^2} > \cancel{\frac{1}{\pi b}} \cdot \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}$$

$$\Rightarrow \cancel{1}+\left(\frac{x-a_2}{b}\right)^2 > \cancel{1}+\left(\frac{x-a_1}{b}\right)^2$$

$$\Rightarrow \left(\frac{x-a_2}{b}\right)^2 - \left(\frac{x-a_1}{b}\right)^2 > 0$$

$$\Rightarrow \left(\frac{x-a_2+x-a_1}{b}\right)\left(\frac{x-a_2-x+a_1}{b}\right) > 0$$

$$\Rightarrow \left(\frac{2x-(a_1+a_2)}{b}\right)\left(\frac{a_1-a_2}{b}\right) > 0$$

In cauchy pdf $b>0$

Let $a_1 > a_2$

$\Rightarrow 2x > a_1 + a_2$

$\Rightarrow x > \frac{a_1+a_2}{2}$

If $a_1 < a_2$

$\Rightarrow 2x < a_1 + a_2$

$\Rightarrow x < \frac{a_1+a_2}{2}$

Thus, the decision boundary is $\frac{a_1+a_2}{2}$.

(B) $\because p(\omega_1) = p(\omega_2)$

$\Rightarrow p(\omega_1) = p(\omega_2) = 0.5$

$$p(\omega_1|x) = \frac{p(x|\omega_1)\,p(\omega_1)}{p(x)}$$

$$p(x) = \sum_{j=1}^{2} p(x|\omega_j)\,p(\omega_j)$$

$$= p(x|\omega_1)\,p(\omega_1) + p(x|\omega_2)\,p(\omega_2)$$

$$= 0.5\left(p(x|\omega_1) + p(x|\omega_2)\right)$$

$$= 0.5\left(\frac{1}{\pi b}\cdot\frac{1}{1+(x-3)^2} + \frac{1}{\pi b}\cdot\frac{1}{1+(x-5)^2}\right)$$

$$\Rightarrow p(\omega_1|x) = \frac{\cancel{0.5}\,\cancel{\frac{1}{\pi b}} \cdot \frac{1}{1+(x-3)^2}}{\cancel{0.5}\,\cancel{\frac{1}{\pi b}} \cdot \left(\frac{1}{1+(x-3)^2} + \frac{1}{1+(x-5)^2}\right)}$$

$$\Rightarrow p(\omega_1|x) = \frac{\frac{1}{1+(x-3)^2}}{\frac{1}{1+(x-3)^2} + \frac{1}{1+(x-5)^2}}$$

(C) $P(\text{error}) = \int P(\text{error}|x)\, P(x)\, dx$

$$= \int_{-\infty}^{\infty} \min[P(\omega_1|x), P(\omega_2|x)]\, P(x)\, dx$$

$$= \int_{-\infty}^{\infty} \min[P(x|\omega_1)P(\omega_1), P(x|\omega_2)P(\omega_2)]\, dx$$

Now, for the ~~decision boundary~~ point where $P(x|\omega_1)$ and $P(x|\omega_2)$ intersect

$$\Rightarrow \cancel{\frac{1}{\pi b}} \frac{1}{1+(x-3)^2} = \cancel{\frac{1}{\pi b}} \frac{1}{1+(x-5)^2}$$

$$\Rightarrow (x-3)^2 = (x-5)^2$$

$$\Rightarrow (x-3-x+5)(x-3+x-5) = 0$$

$$\Rightarrow 2x = 8$$

$$\Rightarrow \boxed{x = 4}$$

For $x < 4$

$P(x|\omega_1)$ let $x = 0$: $= \frac{1}{\pi}\frac{1}{10} = \frac{1}{10\pi}$

$P(x|\omega_2)$ " $: = \frac{1}{\pi}\frac{1}{26} = \frac{1}{26\pi}$

$\Rightarrow \quad P(x|\omega_1) > P(x|\omega_2) \quad \forall\, x \in (-\infty, 4)$

$\Rightarrow \quad P(x|\omega_1) < P(x|\omega_2) \quad \forall\, x \in (4, \infty)$

$4 = \frac{a_1+a_2}{2}$ ← decision boundary

$$\Rightarrow P(\text{error}) = 0.5\int_{-\infty}^{4} \frac{1}{\pi} \cdot \frac{1}{(1+(x-5)^2)}\,dx + 0.5\int_{4 = \frac{a_1+a_2}{2}}^{\infty} \frac{1}{\pi} \cdot \frac{1}{1+(x-3)^2}\,dx$$

$$= \frac{1}{2\pi}\left[\int_{-\infty}^{4} \frac{1}{1+(x-5)^2}\,dx + \int_{4}^{\infty} \frac{1}{1+(x-3)^2}\,dx\right]$$

$\cancel{=} \ \cancel{\frac{1}{2\pi}}$ Let $x-5 = u$ and $x-3 = v$

$\Rightarrow dx = du$ and $dx = dv$

$$= \frac{1}{2\pi}\left[\int_{-\infty}^{-1} \frac{1}{1+u^2}\,du + \int_{1}^{\infty} \frac{1}{1+v^2}\,dv\right]$$

$$= \frac{1}{2\pi}\left[\left(\tan^{-1} u\right)_{-\infty}^{-1} + \left(\tan^{-1} v\right)_{1}^{\infty}\right]$$

$$\cancel{= \frac{1}{2\pi}\left(-\frac{\pi}{2} + \frac{\pi}{4} + \right.}$$

$$= \frac{1}{2\pi}\left[\cancel{-\frac{\pi}{4}} + \frac{\pi}{2} + \cancel{\frac{\pi}{2}} - \cancel{\frac{\pi}{4}}\right]$$

$$= \frac{1}{4}$$

$$= \boxed{0.25}$$

Ans. The overall error rate $\boxed{0.25}$

4 (a) $f(a) = \theta^{a}(1-\theta)^{1-a}$

$f(b) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{b-m}{\sigma}\right)^2}$

$x = [a \quad b]$ a 2-d vector

The covariance of x $\begin{bmatrix} \theta(1-\theta) & 0 \\ 0 & \sigma^2 \end{bmatrix}$

~~Now, cov of~~ x = $\cancel{\begin{matrix} \sigma_{x1}^2 & \sigma_1\sigma_2 \\ \sigma_2\sigma_1 & \sigma_2^2 \end{matrix}}$

Now, cov of x $\begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1x_2}^2 \\ \sigma_{x_1x_2}^2 & \sigma_{x_2}^2 \end{bmatrix}$

where, $\sigma^2_{x_i}$ the the variance of x_i in $\vec{x}$.
$\sigma^2_{x_ix_j}$ is the covariance of x_i and x_j in $\vec{x}$.

$\because \sigma^2_{x_1x_2} = 0$

$\Rightarrow$ x_1 and x_2 are independent (assumed as sir told in class)

$\because x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$

$f(x) = f(a,b) = f(a)\,f(b)$

Ans. $\Rightarrow \boxed{f(x) = \frac{\theta^{a}(1-\theta)^{1-a}}{\sigma\sqrt{2\pi}} \cdot e^{-1/2\left(\frac{b-m}{\sigma}\right)^2}}$

(b) $$p(x) = \frac{\theta^{a}(1-\theta)^{1-a}}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b-m}{\sigma}\right)^2}$$

$q(x)$ = N iid distribution of $P(x)$.

$$q(x) = \prod_{i=1}^{N} p(x_i)$$

$$= \prod_{i=1}^{N} \frac{\theta^{a_i}(1-\theta)^{1-a_i}}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b_i-m}{\sigma}\right)^2}$$

$$\ln q(x) = \sum_{i=1}^{N} \ln\left[\frac{\theta^{a_i}(1-\theta)^{1-a_i}}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b_i-m}{\sigma}\right)^2}\right]$$

$$= \sum_{i=1}^{N}\left[-\ln[\sigma\sqrt{2\pi}] + \ln\theta^{a_i} + \ln(1-\theta)^{1-a_i} - \frac{1}{2}\left(\frac{b_i-m}{\sigma}\right)^2\right]$$

Now, differentiating $\ln q(x)$ wrt θ

$$\Rightarrow \frac{1}{q(x)} \frac{dq(x)}{dx} \cdot \frac{dx}{d\theta} = \sum_{i=1}^{N}\left[\frac{a_i}{\theta} - \frac{(1-a_i)}{1-\theta}\right]$$

Putting $\frac{dq(x)}{d\theta} = 0$ to maximize $q(x)$

$$\Rightarrow \left\{\sum_{i=1}^{N}\left[\frac{a_i}{\theta} - \frac{(1-a_i)}{1-\theta}\right]\right\} \frac{q(x)}{dx} = 0$$

$$\Rightarrow \sum_{i=1}^{N} \frac{a_i}{\theta} = \sum_{i=1}^{N} \frac{1-a_i}{1-\theta}$$

Let $\sum_{i=1}^{N} a_i = Y$

$$\Rightarrow \frac{Y}{\theta} = \frac{N-Y}{1-\theta}$$

$$Y(1-\theta) = \theta(N-Y)$$

$$\Rightarrow Y - \cancel{Y\theta} = N\theta - \cancel{\theta Y}$$

$$\Rightarrow \theta = \frac{Y}{N}$$

Ans. $\Rightarrow \boxed{\theta = \frac{1}{N} \cdot \sum_{i=1}^{N} a_i}$