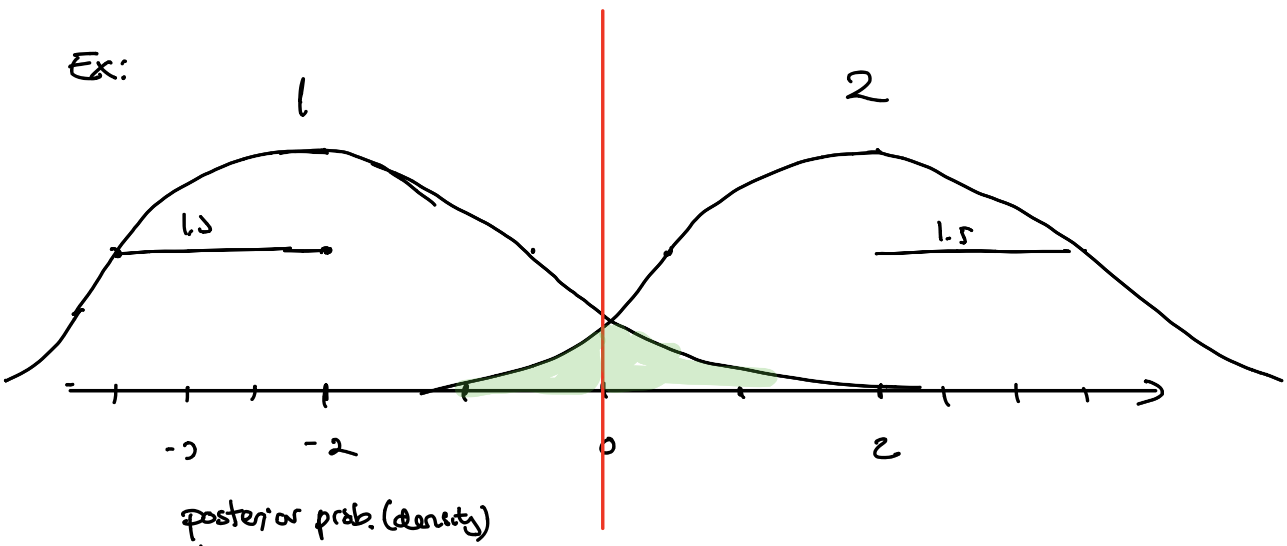
M4: Classification

29.01.2018

<u>Bayes classifier:</u> At observation $X = x_0$ classify to the class k where $P(Y=k \mid X = x_0)$ is largest, for all $k = 1, \dots, K$

Ex:



posterior prob. (density)

$$P(Y=1 \mid X = x) > P(Y=2 \mid X = x) \rightarrow \text{classify to } 1$$

$$\frac{P(Y=1 \cap X = x)}{P(X=x)} > \frac{P(Y=2 \cap X=x)}{P(X=x)} \quad \text{def of cond. prob.}$$

$$\frac{P(X=x \mid Y=1) \cdot P(Y=1)}{\cancel{P(X=x)}} > \frac{P(X=x \mid Y=2) \cdot P(Y=2)}{\cancel{P(X=x)}}$$

$$P(X=x) = \underbrace{P(X=x \mid Y=1)}_{f_1(x)} \cdot \underbrace{P(Y=1)}_{\pi_1} + \underbrace{P(X=x \mid Y=2)}_{f_2(x)} \underbrace{P(Y=2)}_{\pi_2}$$

class density $\rightarrow f_1(x)$

π_1 ← prior

$f_1(x)\cdot\pi_1 > f_2(x)\cdot\pi_2$ border for $f_1(x) = f_2(x)$

$\downarrow$

Here: $\pi_1=\pi_2=\frac{1}{2} \Rightarrow$ classify to 1 if $f_1(x) > f_2(x)$

$$\cancel{\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}}\, e^{-\frac{1}{2\sigma^2}(x-\mu_1)^2} = \cancel{\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}}\, e^{-\frac{1}{2\sigma^2}(x-\mu_2)^2}$$

$$\frac{1}{2\sigma^2}(x-\mu_1)^2 = \frac{1}{2\sigma^2}(x-\mu_2)^2$$

$$(x-\mu_1)^2 = (x-\mu_2)^2$$

$$\cancel{x^2} - 2\mu_1 x + \mu_1^2 = \cancel{x^2} - 2\mu_2 x + \mu_2^2$$

$$\mu_1^2 - \mu_2^2 = 2(\mu_1-\mu_2)x$$

$$x = \frac{\mu_1^2-\mu_2^2}{2(\mu_1-\mu_2)} = \boxed{\frac{\mu_1+\mu_2}{2}}$$

$(\mu_1-\mu_2)(\mu_1+\mu_2)$

Here: $\mu_1 = -\mu_2 \Rightarrow x=0$ boundary. Makes sense

Bayes error: $\frac{1}{2}P(X>0 \mid Y=1) + \frac{1}{2}(X<0 \mid Y=2)$

$$= \frac{1}{2}\, 2\int_0^\infty f_1(x)\,dx = \dots = 009 \rightarrow 9\%$$

If we make rule where error rate $< 9\% \Rightarrow$ something's <u>wrong</u>!

LDA ($p=1$) for $K=2$

$$\delta_k(x) = x \cdot \frac{\mu_k}{\sigma^2} - \frac{\mu_k^2}{2\sigma^2} + \log(\pi_k)$$

$$\delta_1(x) = \delta_2(x)$$

$$x \cdot \frac{\mu_1}{\sigma^2} - \frac{\mu_1^2}{2\sigma^2} + \log \pi_1 = x \frac{\mu_2}{\sigma^2} - \frac{\mu_2^2}{2\sigma^2} + \log \pi_2$$

$$x\left(\frac{\mu_1}{\sigma^2} - \frac{\mu_2}{\sigma^2} \right) = \frac{\mu_1^2}{2\sigma^2} - \frac{\mu_2^2}{2\sigma^2} + \log \pi_2 - \log \pi_1$$

$$x = \frac{\frac{1}{2\sigma^2}\left(\mu_1^2 - \mu_2^2\right) + \log \pi_2 - \log \pi_1}{\frac{\mu_1 - \mu_2}{\sigma^2}}$$

$$= \frac{1}{2}\left(\mu_1 + \mu_2\right) + \frac{\left(\log \pi_2 - \log \pi_1\right)\sigma^2}{\mu_1 - \mu_2}$$

Why is $P(Y=k \mid X=x) = \frac{e^{\delta}}{\sum e^{\delta}}$?

$$P(Y=k \mid X=x) = \frac{\pi_k \cdot \cancel{\frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma^{-1}(x-\mu_k)}}{\sum_{l=1}^{K} \pi_l \cancel{\frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}}} e^{-\frac{1}{2}(x-\mu_l)^T \Sigma^{-1}(x-\mu_l)}}$$

$$= \frac{\pi_k \, e^{-\frac{1}{2}(x-\mu_k)^T \Sigma^{-1}(x-\mu_k)}}{\pi_1 \cdot e + \pi_2 e + \cdots + \pi_k e} = \frac{e^{\ln \pi_k - \frac{1}{2}(\,)\Sigma^{-}\cdots}}{e^{\ln \pi + \cdots} + e + \cdots}$$

$$\frac{1}{2}(x-\mu_k)^T \Sigma^{-1}(x-\mu_k) = \frac{1}{2}x^T\Sigma^{-1}x$$

$$- \frac{1}{2} x^T \Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^T \Sigma^{-1} x + \frac{1}{2}\mu_k^T \Sigma^{-1}\mu_k$$

$$= \frac{1}{2} x^T \Sigma^{-1} x - x^T \Sigma^{-1}\mu_k + \frac{1}{2}\mu_k^T \Sigma^{-1}\mu_k$$

$\to$ exponent: $\quad \ln \pi_k - \frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^T\Sigma^{-1}\mu_k$

$\delta_k(x) =$ green

$$P(Y=k \mid X=x) = \frac{e^{\delta_k(x)} \cdot e^{-\frac{1}{2}x^T\Sigma^{-1}x}}{e^{\delta_1(x)} e^{-\frac{1}{2}x^T\Sigma^{-1}x} + e^{\delta_2(x)} \cdot e^{-\frac{1}{2}\cdots} + \cdots}$$

$$= \frac{e^{\delta_k(x)}}{\sum e^{\delta_l(x)}}$$ ok.

QDA :

$$(\underbrace{p \times p}_{\text{all}} - \underbrace{p}_{\text{diag}})/2 + \underbrace{p}_{\text{diag}}$$

$$= \frac{1}{2}p^2 - \frac{1}{2}p + p = \frac{1}{2}p^2 + \frac{1}{2}p = \underline{p(p+1)/2}$$

5