Aprendizagem

Instituto Superior Técnico

setembro de 2023

Homework 2 - Report

Joana Pimenta (103730), Rodrigo Laia (102674)

Pen and Paper

1. (a) y_1, y_2, y_3, y_4 and y_5 independent $\implies p(y_1, y_2, y_3, y_4, y_5) = p(y_1, y_2) \times p(y_3, y_4) \times p(y_5)$

   Fórmulas utilizadas:

$$P(y_6 = H|\vec{x}) = \frac{P(\vec{x}|y_6 = H)}{P(\vec{x})} \tag{1}$$

$$P(\vec{x}|\mu, \sigma^2) = \frac{1}{(2\pi)^{m/2}\sqrt{|\Sigma|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x}-\vec{\mu})} \tag{2}$$

$$\vec{\mu} = \begin{bmatrix} E(y_1) \\ E(y_2) \end{bmatrix} \tag{3}$$

$$\Sigma = \begin{bmatrix} cov(y_1, y_2) & cov(y_1, y_1) \\ cov(y_2, y_1) & cov(y_2, y_2) \end{bmatrix} \tag{4}$$

$$|\Sigma| = cov(y_1, y_2) \cdot cov(y_2, y_1) - cov(y_1, y_1) \cdot cov(y_2, y_2) \tag{5}$$

$$\Sigma^{-1} = \frac{1}{|\Sigma|} \cdot \begin{bmatrix} cov(y_2, y_2) & -cov(y_1, y_2) \\ -cov(y_2, y_1) & cov(y_1, y_1) \end{bmatrix} \tag{6}$$

   Parâmetros das gaussianas multivariadas:

   Classe A:

$$n\vec{\mu}_A = \begin{bmatrix} 0.24 \\ 0.52 \end{bmatrix}$$

$$\Sigma_A = \begin{bmatrix} 0.004267 & -0.0064 \\ -0.0064 & 0.02240 \end{bmatrix}$$

$$|\Sigma|_A = 5.4613 \cdot 10^{-5}$$

$$\Sigma_A^{-1} = \begin{bmatrix} 410.1563 & -117.1875 \\ -117.1875 & 78.125 \end{bmatrix}$$

$$P(\vec{x}|A) = N(\vec{x}|\mu_A, \Sigma_A) = \frac{1}{(2\pi)^{m/2}\sqrt{|\Sigma_A|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu}_A)^T \cdot \Sigma_A^{-1} \cdot (\vec{x}-\vec{\mu}_A)}$$

Classe B:

$$\vec{\mu}_B = \begin{bmatrix} 0.5925 \\ 0.3275 \end{bmatrix}$$

$$\Sigma_B = \begin{bmatrix} 0.01717 & -0.00732 \\ -0.00732 & 0.02362 \end{bmatrix}$$

$$|\Sigma|_B = 3.519 \cdot 10^{-4}$$

$$\Sigma_B^{-1} = \begin{bmatrix} 67.1101 & 20.7954 \\ 20.7954 & 48.7831 \end{bmatrix}$$

$$P(\vec{x}|B) = N(\vec{x}|\mu_B, \Sigma_B) = \frac{1}{(2\pi)^{m/2}\sqrt{|\Sigma_B|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu}_B)^T \cdot \Sigma_B^{-1} \cdot (\vec{x}-\vec{\mu}_B)}$$

Probabilidades para $\{y_3, y_4\}$ condicionadas a A e B :

Classe A:

	$y_3 = 0$	$y_3 = 1$
$y_4 = 0$	P=0	P=1/3
$y_4 = 1$	P=1/3	P=1/3

Tabela 1: Probabilidades para y_3, y_4 condicionadas a A

Classe B:

	$y_3 = 0$	$y_3 = 1$
$y_4 = 0$	P=1/2	P=1/4
$y_4 = 1$	P=1/4	P=0

Tabela 2: Probabilidades para y_3, y_4 condicionadas a B

Probabilidades para $\{y_5\}$ condicionadas a A e B :

Classe A:

$$P(y_5 = 0|A) = 1/3$$

$$P(y_5 = 1|A) = 1/3$$

$$P(y_5 = 2|A) = 1/3$$

Classe B:

$$P(y_5 = 0|A) = 1/4$$

$$P(y_5 = 1|A) = 1/2$$

$$P(y_5 = 2|A) = 1/4$$

Priors:

$$P(A) = \frac{3}{7}$$

$$P(B) = \frac{4}{7}$$

(b) Uma vez que o denominador é o mesmo para todas para saber qual a classe mais provável, basta comparar os numeradores das probabilidades.

$$\begin{aligned}
P(A|\vec{x}_8) &= \frac{P(\vec{x}_8|A) \cdot P(A)}{P(\vec{x}_8)} \\
&= \frac{P(y_1 = 0.38, y_2 = 0.52|A) \cdot P(y_3 = 0, y_4 = 1|A) \cdot P(y_5 = 0|A) \cdot P(A)}{P(\vec{x}_8)} \\
&= \frac{\frac{3}{7} \cdot 0.3868 \cdot \frac{1}{3} \cdot \frac{1}{3}}{P(\vec{x}_8)} \\
&= \frac{0.018}{P(\vec{x}_8)}
\end{aligned}$$

$$\begin{aligned}
P(B|\vec{x}_8) &= \frac{P(\vec{x}_8|B) \cdot P(B)}{P(\vec{x}_8)} \\
&= \frac{P(y_1 = 0.38, y_2 = 0.52|B) \cdot P(y_3 = 0, y_4 = 1|B) \cdot P(y_5 = 0|B) \cdot P(B)}{P(\vec{x}_8)} \\
&= \frac{\frac{4}{7} \cdot 1.7678 \cdot \frac{1}{4} \cdot \frac{1}{4}}{P(\vec{x}_8)} \\
&= \frac{0.063}{P(\vec{x}_8)}
\end{aligned}$$

y2	0.36	0.48	0.72	0.11	0.39	0.28	0.53	0.52	0.59
new y2	0	0	1	0	0	0	1	1	1

Como $P(A|\vec{x}_8) < P(B|\vec{x}_8)$, então $\vec{x}_8$ é classificado como B.

$$\begin{aligned} P(A|\vec{x}_9) &= \frac{P(\vec{x}_9|A) \cdot P(A)}{P(\vec{x}_9)} \\ &= \frac{P(y_1 = 0.42, y_2 = 0.59|A) \cdot P(y_3 = 0, y_4 = 1|A) \cdot P(y_5 = 0|A) \cdot P(A)}{P(\vec{x}_9)} \\ &= \frac{\frac{3}{7} \cdot 0.1013 \cdot \frac{1}{3} \cdot \frac{1}{3}}{P(\vec{x}_9)} \\ &= \frac{0.0048}{P(\vec{x}_9)} \end{aligned}$$

$$\begin{aligned} P(B|\vec{x}_8) &= \frac{P(\vec{x}_8|B) \cdot P(B)}{P(\vec{x}_8)} \\ &= \frac{P(y_1 = 0.42, y_2 = 0.59|B) \cdot P(y_3 = 0, y_4 = 1|B) \cdot P(y_5 = 1|B) \cdot P(B)}{P(\vec{x}_8)} \\ &= \frac{\frac{4}{7} \cdot 1.4927 \cdot \frac{1}{4} \cdot \frac{1}{2}}{P(\vec{x}_8)} \\ &= \frac{0.1066}{P(\vec{x}_8)} \end{aligned}$$

Como $P(A|\vec{x}_9) < P(B|\vec{x}_9)$, então $\vec{x}_9$ é classificado como B.

(c) Assumindo o critério de Maximum Likelihood, para classificar uma observação apenas interessam as probabilidades $P(\vec{x}|A)$ e $P(\vec{x}|B)$:

$$h = argmax(P())$$

Considerando diferentes thresholds para as probabilidades é possível maximizar a accuracy do nosso classificador.

$$P(\vec{x}_8|A) = P(y_1 = 0.38, y_2 = 0.52|A) \cdot P(y_3 = 0, y_4 = 1|A) \cdot P(y_5 = 0|A) = 0.043$$

2. Para discretizar a variável y_2, considerando equal-width, é necessário dividir o intervalo $[0, 1]$ em 2 partes iguais. Assim, os intervalos são: $[0, 0.5]$, $[0.5, 1]$.

Para cada observação, y_2 pode assumir os valores 0 ou 1, consoante o intervalo em que se encontra o seu valor.

(a) Porque não há shuffling:

Programming - Código Python e Resultados Obtidos

1.