

Lista de exercícios para a disciplina de Reconhecimento de Padrões
(actualizada em 27 de Janeiro de 2011)

1. Considere um hipotético problema de classificação com duas classes (ω_1 e ω_2) e observações descritas por uma única variável. As funções de densidade de probabilidade das observações nas classes são dadas por

$$p(x/\omega_1) = \begin{cases} 2e^{-2x} & , x \geq 0 \\ 0 & , x < 0 \end{cases} \quad \text{e} \quad p(x/\omega_2) = \begin{cases} 4e^{-4x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

e as probabilidades *a priori* das classes verificam $P(\omega_1) = 2P(\omega_2)$. A função de custo do problema é dada por $\lambda_{11} = \lambda_{22} = 0$, $\lambda_{12} = 2$ e $\lambda_{21} = 1$.

(a) Mostre que as regiões de decisão para a regra de decisão de erro mínimo deste problema são definidas por $\mathcal{R}_1 = \{x : x \in (\frac{\log 2}{2}, +\infty)\}$ e $\mathcal{R}_2 = \{x : x \in [0, \frac{\log 2}{2}]\}$.

(b) Indique a expressão da função de decisão $\alpha(x)$ e determine o risco global da regra de decisão. Como interpreta esse valor?

2. Considere um problema de classificação com duas classes (ω_1 , representada por 1, e ω_2 , representada por 2) e observações descritas por duas variáveis (x e y). Suponha que dispõe de uma amostra de treino com 50 observações recolhida aleatoriamente nas duas classes, representada pela matriz amostra cujas 6 primeiras linhas são indicadas abaixo.

	x	y	classe
[1,]	-1.3434878	-2.3930602	2
[2,]	0.6866480	3.3167074	1
[3,]	3.2808657	5.1351276	1
[4,]	-0.1228949	-0.8906718	2
[5,]	1.7382977	-3.9843433	2
[6,]	-0.3933782	3.1319467	2

Considere os seguintes resultados de alguns comandos da aplicação R:

```
> amostral=amostra[amostra[,3]==1, ]
```

```
> dim(amostral)
```

```
[1] 20 3
```

```
> amostra2=amostra[amostra[,3]==2, ]
```

```
> apply(amostral,2,mean)
```

```
      x      y  classe
2.265094 4.438462 1.000000
```

```
> apply(amostra2,2,mean)
```

```

          x          y      classe
0.03012504 -0.42521895  2.00000000

> cov(amostra1[, -3])
          x          y
x 1.0429166 0.6915359
y 0.6915359 1.0564407

> solve(cov(amostra1[, -3]))
          x          y
x 1.694215 -1.109017
y -1.109017 1.672526

> cov(amostra2[, -3])
          x          y
x 3.2912602 -0.1834484
y -0.1834484 4.3438848

> solve(cov(amostra2[, -3]))
          x          y
x 0.30455187 0.01286166
y 0.01286166 0.23075185

```

- (a) Indique as estimativas de máxima verosimilhança das probabilidades *a priori* de $P(\omega_1)$, $P(\omega_2)$, dos vectores média para ω_1 e para ω_2 , e das matrizes de variância-covariância para ω_1 e para ω_2 .
- (b) Considere a regra de decisão de erro mínimo de Bayes, sob o pressuposto que as probabilidades *a priori* das classes são as estimadas na alínea anterior e que as distribuições das observações nas classes são normais multivariadas com parâmetros iguais aos estimados na alínea anterior.
- Indique as expressões das funções discriminantes correspondentes a essa regra de decisão.
 - Indique uma equação para a fronteira de decisão entre as regiões de decisão \mathcal{R}_1 e \mathcal{R}_2 .
 - Mostre, efectuando os cálculos, que a regra de decisão afecta as duas novas observações (12.265094, 14.438462) e (0.265094, 2.438462) à classe ω_2 .
 - Conclua, com base no resultado da alínea anterior e no output, que uma das regiões de decisão não é convexa e indique qual. Interprete esse resultado à luz dos dados.
3. Considere um problema de classificação com duas classes (ω_1 , representada por 1, e ω_2 , representada por 0) e observações descritas por duas variáveis (x_1 e x_2). Suponha que se treinou uma rede neuronal do tipo “feed-forward”, cuja função de activação é uma função logística com expressão $f(u) = \frac{1}{1+e^{-u}}$, com

2 vértices de input x_1 e x_2 , um vértice de output g , e uma camada intermédia com 3 vértices a, b, c , e que os pesos da rede obtidos são os seguintes.

aresta	peso
$x \rightarrow a$	$w_{x_1 a} = 1$
$x \rightarrow b$	$w_{x_1 b} = 1$
$x \rightarrow c$	$w_{x_1 c} = 1$
$y \rightarrow a$	$w_{x_2 a} = 2$
$y \rightarrow b$	$w_{x_2 b} = 1$
$y \rightarrow c$	$w_{x_2 c} = 1$
$a \rightarrow g$	$w_{ag} = 2$
$b \rightarrow g$	$w_{bg} = -1$
$c \rightarrow g$	$w_{cg} = -1$

- (a) Considerando que o vértice g representa a função discriminante e que a regra de decisão é “decide ω_1 se $g(x_1, x_2) \geq \frac{1}{2}$ e decide ω_2 caso contrário”, determine a afectação da observação $(0, 0)$ dada pela rede neuronal.
- (b) Mostre que a função de activação verifica $f'(u) = f(u) [1 - f(u)]$.
- (c) Suponha que se obtém uma nova observação de classe ω_1 para a qual $x = 0$ e $y = 1$. Qual a direcção de variação dos pesos w_{ag} , w_{bg} e w_{cg} da rede que causa o maior decréscimo do erro de treino $J(w) = 1 - g_w(0, 1)$ sobre essa observação?