Pattern Recognition
Reconhecimento de Padrões

2017/2018

Exame Normal <u>18 June 2018</u> Duration: 2h30

Name: **Number:**   **Practical Class:**

Question	pts	Results	Graded by:
1)	20		
2)	20		
3)	20		
4)	25		
5)	15		

Graded by:

Question 1 - Minimum Distance Classifiers (MDC)

☐ **20 pts**

(a) Describe how a MDC works, in which consists its training, and given new patterns how they are labeled.

(b) Given the data available in Figure 1:

- Explain why a Euclidean MDC is unable to classify data with zero error? Demonstrate graphically.

- Develop a Mahalanobis MDC for the data. To which class the classifier will assign the pattern $\mathbf{x} = [-2 \ \ 0]^T$?

- Compute the decision boundary $d_{12}(\mathbf{x})$ and the separation hyperplane for the Mahalanobis MDC. Are the hyperplane what you expect? Justify.
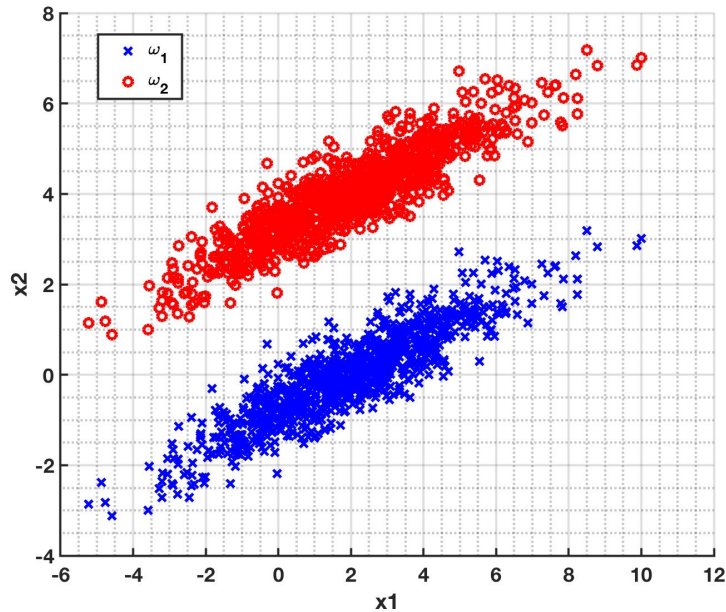

Figure 1: Binary classification problem described by features x1 and x2

Calculus support:

Mean vectors: $\mu_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ $\mu_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

Covariance matrices: $\mathbf{C} = \mathbf{C}_1 = \mathbf{C}_2 = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$

Inverse covariance matrix: $\mathbf{C}^{-1} = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$

Your answer 1):

Cont. your answer to 1):

Cont. your answer to 1):

Question 2 - Hierarchical Clustering

a) b) Hierarchical clustering algorithms outputs a dendrogram. Describe what is a dendrogram?

b) Consider the patterns P1, P2, P3, P4 and P5, and the inter-pattern distance represented in the following matrix:

	P1	P2	P3	P4
P1	0	1	4	5
P2	-	0	2	6
P3	-	-	0	3
P4	-	-	-	0

Develop the dendrogram considering the **Ward's** approach (Note: Consider the Lance-Williams formula).

Your answer to 2):

Cont. your answer to 2):

Question 3 - Multivariate Normal Bayesian Classification

☐ **20 pts** Consider the conditional probability density functions in Figure 2 that are generated by a multivariate Gaussian process. The mean vectors for both classes are: $\mu_1 = [-1 \quad -1]^T$ and $\mu_2 = [1 \quad 1]^T$. The covariance matrix is the identity matrix for both classes. Consider also that patterns from ω_1 and ω_2 appears in equal prevalence.
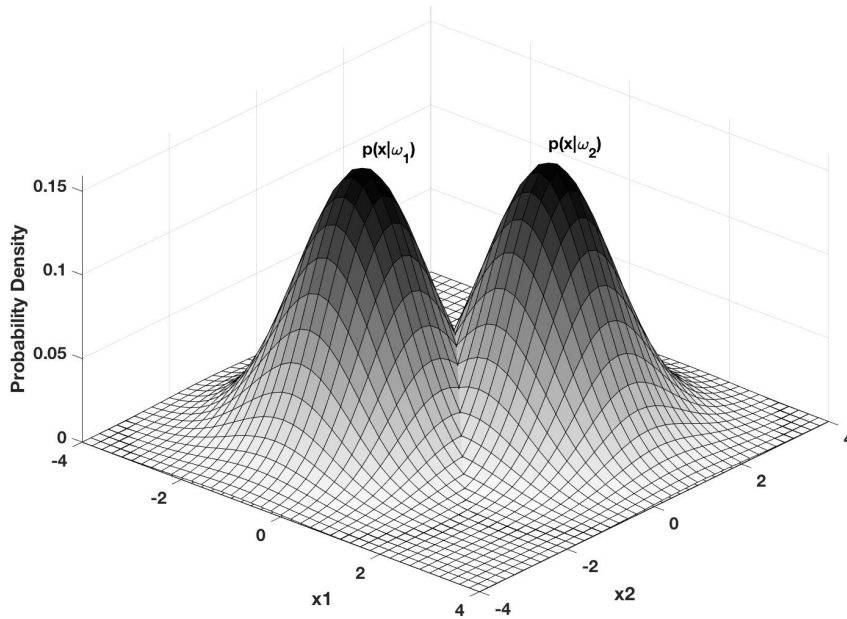

Figure 2: Conditional probability density functions defined in a space spanned by features x1 and x2.

a) What is the problem dimensionality? Justify.

b) Starting with the Bayes rule develop the Likelihood Ratio Test (LRT).

c) Develop the generic decision function of the form $d_{12}(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + w_0$. Specify all the steps.

d) Specify $d_{12}(\mathbf{x})$ for the concrete values of this exercise. What is the relation between this classifier and an Euclidean MDC? Justify.

Your answer to 3):

Cont. your answer to 3):

Cont. your answer to 3):

Question 4 - Multiclass Support Vector Machines

☐ **25 pts**

Consider a three class linear SVM classification strategy based on the One-vs-One strategy, implemented by considering the error-correcting output codes (ECOC) approach. The parameters $\mathbf{w}$ and b for each one of the classifiers are:

Classifier	Problem	$\mathbf{w}$	b
SVM1	ω_1 (+) vs ω_2 (-)	$[-1 \quad 1]^T$	1
SVM2	ω_1 (+) vs ω_3 (-)	$[-1 \quad -1]^T$	7
SVM3	ω_2 (+) vs ω_3 (-)	$[0 \quad -1]^T$	2

(a) Compute the separation margin for SVM1.

(b) What is the coding design matrix? If training One-vs-All was selected what should be the coding design matrix?

(c) Is $\mathbf{x} = [6 \quad 1]^T$ a support vector? If yes, what is its class? Justify your answers.

(d) What is the accuracy if the following testing patterns $\mathbf{x} = [x1 \quad x2]^T$ are considered?

Note: Consider the Hamming distance in the decoding phase.

$x1$	2	3	5
$x2$	0	0	5
True class(ω_k)	1	2	3

Your answer 4):

Cont. your answer to 4):

Cont. your answer to 4):

Question 5 - kNN probability density function (*pdf*) estimation

☐ **15 pts**

Write a function in your language of choice to perform **univariate** non-parametric probability density function (*pdf*) estimate by using the kNN approach. The function should receive the univariate data vector and the number of neighbors (k). The function should return a vector with the *pdf* estimate.

Consider that you have already a function that given the univariate data vector and a testing point, it computes the indexes and distances of the closest k points to the testing point in the data vector.

Your answer to 5):

Cont. your answer to 5):