TEMA 5 – RECONOCIMIENTO DE OBJETOS

5.3.2.- Clasificadores basados en distribuciones normales

••

EJERCICIOS

Observación:

- Los ejercicios 1, 2 y 3 deben resolverse con la única ayuda de una calculadora. Para verificar que los resultados son correctos, además, deben ser también implementados en Matlab.
- Los ejercicios 4 y 5 deben resolverse utilizando Matlab.

••

EJERCICIO 1

A partir de los datos de la tabla, construir un Clasificador de Mínima Distancia Euclídea:

CLASE 1

PATRON	1	2	3	4	5	6	7	8	9	10
x_1	1	2	2	2	2	3	3	4	5	1
x_2	3	1	2	3	4	2	3	3	2	2

CLASE 2

PATRON	1	2	3	4	5	6	7	8	9	10
x_1	4	5	5	4	6	6	6	7	4	8
x_2	5	5	6	7	5	6	7	6	6	7

- Funciones de decisión de cada clase y función discriminante entre clases, expresadas en función de una instancia X genérica definida por x1 y x2.
- Establecer la regla de decisión del clasificador.
- En el espacio de características definido por x1 y x2, representar el conjunto de instancias de cada clase y la frontera lineal de separación entre clases del clasificador. ¿Qué significado tienen los puntos de esta frontera lineal?

EJERCICIO 2

Diseñar un Clasificador de Mínima Distancia Mahalanobis suponiendo 3 clases de patrones, cada uno de ellos representados por 2 características, y los siguientes datos:

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \; ; \; P(C_1) = P(C_2) = P(C_3)$$

$$Vectores\ Promedio\ de\ cada\ clase\text{:}\;\; M^1 = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \; ; \; M^2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix} \; ; \quad M^3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$Matrices\ Covariaza\ de\ cada\ clase\text{:}\; C^1 = C^2 = C^3 = C = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/4 \end{bmatrix} ; \; C^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

El diseño del clasificador implica la obtención de las siguientes funciones de x1 y x2:

- Funciones de decisión de cada clase

- Funciones discriminantes entre las muestras de las clases dos a dos.

- Establecer la regla de decisión del clasificador.

EJERCICIO 3

Teniendo en cuenta la muestra de la tabla, diseñar un Clasificador de Mínima Distancia Mahalanobis suponiendo que las dos clases tienen la misma matriz de covarianzas (para su estimación considerar conjuntamente los patrones de ambas clases):

CLASE 1

PATRON	1	2	3	4
x_1	2	3	3	4
x_2	1	2	3	2

CLASE 2

PATRON	1	2	3
x_1	6	5	7
x_2	1	2	3

- Funciones de decisión de cada clase y función discriminante entre clases, expresadas en función de una instancia X genérica definida por x1 y x2.

- Establecer la regla de decisión del clasificador.

- En el espacio de características definido por x1 y x2, representar el conjunto de instancias de cada clase y la frontera lineal de separación entre clases del clasificador. ¿Qué significado tienen los puntos de esta frontera lineal?

EJERCICIO 4

Para realizar un sistema diagnóstico de una determinada enfermedad, se han obtenido datos de 5 biomarcadores diferentes sobre personas sanas y personas afectadas por la enfermedad. Para ello se han realizado 2 experimentos, el primero con los dos primeros biomarcadores y el segundo con los tres restantes. Los resultados de estos experimentos se facilitan en el archivo comprimido *datos_ejercicio4.zip*.

a) Representa para cada experimento los datos de los diferentes biomarcadores, distinguiendo en la representación aquellas muestras que se corresponden a personas sanas de aquellas tomadas a personas afectadas por la enfermedad.

b) Diseña un clasificador lineal que permita predecir la enfermedad en un paciente a partir de los biomarcadores utilizados en los experimentos 1 y 2. Evalúa la precisión del clasificador en el conjunto de muestras disponibles de cada experimento.

c) En el caso del experimento 2, diseña también un clasificador cuadrático basado en mínima distancia de Mahalanobis.

Observación:

La clasificación lineal basada en mínima distancia de Mahalanobis asume que las clases presentan la misma matriz de covarianzas calculada de la siguiente forma:

$$C = \frac{N_1 C_1 + N_2 C_2}{N_1 + N_2}, donde\ N_i\ y\ C_i\ es\ el\ número\ de\ datos\ y\ la\ matriz\ de\ covarianzas\ de\ la\ clase\ i$$

En el caso cuadrático, la función discriminante se obtiene como la diferencia de las funciones de decisión definidas para cada clase, asumiendo que cada clase tiene su propia matriz de covarianzas.

EJERCICIO 5

En el archivo comprimido *datos_ejercicio5.zip* se facilitan tres conjuntos de datos X-Y:

- datos_MDE_2dimensiones.mat
- datos_MDM_2dimensiones.mat
- datos_MDM_3dimensiones.mat

Para cada conjunto de datos X-Y:

1.- Divide el conjunto completo en dos subconjuntos: entrenamiento (70% de los datos seleccionados de forma aleatoria) y test (30% de los datos restantes). Para ello, utiliza el siguiente código:

```
numDatos = size(X,1);
porcentajeTrain = 0.7;
numDatosTrain = round(porcentajeTrain*numDatos);
numerosMuestrasTrain = randsample(numDatos,numDatosTrain);
numerosMuestrasTest =
find(not(ismember(1:numDatos,numerosMuestrasTrain)));

% Conjunto de Train
XTrain = X(numerosMuestrasTrain,:);
YTrain = Y(numerosMuestrasTrain);
% Conjunto de Test
XTest = X(numerosMuestrasTest,:);
YTest = Y(numerosMuestrasTest);
```

2.- En una misma ventana tipo figure, representa en dos gráficas independientes las muestras de entrenamiento y de test en el espacio de características.

3.- Utilizando el conjunto de entrenamiento, diseña un clasificador MDE o MDM, según indica el nombre del fichero de datos *.mat. Para el caso del clasificador MDM, se debe diseñar su formulación lineal y cuadrática.

4.- Incorpora a la representación del punto 2, la frontera lineal que utiliza el clasificador diseñado para particionar el espacio de características.

5.- Evalúa la precisión del clasificador en el conjunto de muestras de test. En el caso de MDM, evalúa la versión cuadrática y lineal del clasificador.