Ejercicios Tema 6

Percepción

Curso 2021/2022

1. Dada la siguiente matriz Gramm

$$\mathbf{K} = \begin{bmatrix} 1 & 1/3 & 1/3 & 1/3 & 1/5 \\ 1/3 & 1 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1 & 1/5 & 1/3 \\ 1/3 & 1/3 & 1/5 & 1 & 1/3 \\ 1/5 & 1/3 & 1/3 & 1/3 & 1 \end{bmatrix}$$

calculada a partir del conjunto de aprendizaje $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, -1), (\mathbf{x}_3, +1), (\mathbf{x}_4, -1), (\mathbf{x}_5, +1)\}$. Se pide calcular los α_n con el algoritmo Kernel Perceptron empleando el orden dado de las muestras para el recorrido de las mismas.

2. Se define la función kernel $K(\mathbf{x}, \mathbf{y}) = \frac{1}{d(\mathbf{x},\mathbf{y})+1}$ siendo $d(\mathbf{x}, \mathbf{y})$ la distancia de Hamming $\left(d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{D} 1 - \delta(x_i, y_i)\right)$, $\delta(a, b) = 1$ si $a = b$, $\delta(a, b) = 0$ si $a \neq b$:

   a) Calcula la matriz Gramm $\mathbf{K}$ para el siguiente conjunto de aprendizaje:

   $$X = \{((1, 1), +1), ((2, 2), -1), ((1, 3), +1), ((3, 1), -1), ((3, 3), +1)\}$$

   b) ¿Es K un kernel?

3. Sea la función Kernel $K(\mathbf{x}, \mathbf{y}) = \frac{1}{d(\mathbf{x},\mathbf{y})+1}$ siendo $d(\mathbf{x}, \mathbf{y})$ la distancia de Hamming $\left(d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{D} 1 - \delta(x_i, y_i)\right)$. Sea el conjunto de entrenamiento $\mathbf{x}_1 = (0, 0)$, $\mathbf{x}_2 = (0, 1)$, $\mathbf{x}_3 = (1, 0)$, $\mathbf{x}_4 = (1, 1)$, siendo $\mathbf{x}_1, \mathbf{x}_4 \in +1$ y $\mathbf{x}_2, \mathbf{x}_3 \in -1$.

   a) Representa gráficamente el conjunto de entrenamiento y di si es o no linealmente separable

   b) Aplica una iteración de Kernel Perceptron partiendo de $\alpha = (0, 0, 0, 0)$, indicando el valor final de α

   c) ¿Cuál es el error de clasificación de la función $g(\mathbf{x})$ resultante del apartado anterior sobre el conjunto de entrenamiento?

4. Sea la función Kernel $K(\mathbf{x}, \mathbf{y}) = \exp(\mathbf{x}^t \cdot \mathbf{y})$ y sea el conjunto de entrenamiento $\mathbf{x}_1 = (1, 0)$, $\mathbf{x}_2 = (-1, 0)$, $\mathbf{x}_3 = (-1, 1)$, $\mathbf{x}_4 = (0, -1)$, siendo $\mathbf{x}_1, \mathbf{x}_3 \in -1$ y $\mathbf{x}_2, \mathbf{x}_4 \in +1$.

   a) Aplicar una iteración de Kernel Perceptron partiendo de $\alpha = (0, 0, 0, 0)$, indicando el valor final de α

   b) ¿Tiene sentido aplicar Kernels en este conjunto de entrenamiento?

5. Dada la función kernel $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 1)^2$ y el conjunto de aprendizaje en $\mathbb{R}^2$ $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, -1), (\mathbf{x}_3, +1), (\mathbf{x}_4, -1)\}$, siendo $\mathbf{x}_1 = (0, 2)$, $\mathbf{x}_2 = (-1, 0)$, $\mathbf{x}_3 = (1, 1)$, $\mathbf{x}_4 = (0, 1)$, se pide:

   a) Obtener la matriz Gramm $\mathbf{K}$ asociada a X

   b) Realizar una iteración del algoritmo Kernel Perceptron partiendo del conjunto de pesos $\alpha = (0, 0, 0, 0)$

   c) Clasificar las muestras $\mathbf{y}_1 = (1, 2)$ y $\mathbf{y}_2 = (0, 0)$ de acuerdo a los pesos obtenidos en el apartado previo

   d) ¿Crees que en este problema de clasificación era necesario emplear kernels? Razona la respuesta

6. Dada la función kernel $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 3)^2$ y el conjunto de aprendizaje en $\mathbb{R}^2$ $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, +1), (\mathbf{x}_3, -1), (\mathbf{x}_4, -1), (\mathbf{x}_5, +1)\}$, siendo $\mathbf{x}_1 = (1, -1)$, $\mathbf{x}_2 = (0, 1)$, $\mathbf{x}_3 = (1, 1)$, $\mathbf{x}_4 = (2, -1)$, $\mathbf{x}_5 = (1, 0)$, se pide:

   a) Obtener la matriz Gramm $\mathbf{K}$ asociada a X

   b) Realizar una iteración del algoritmo Kernel Perceptron partiendo del conjunto de pesos $\alpha = (0, 0, 0, 0, 0)$

   c) Clasificar la muestra $\mathbf{y} = [-2\ 1]$ de acuerdo a los pesos obtenidos en el apartado previo

7. Dada la función kernel
$$K(\mathbf{x},\mathbf{y}) = (|\mathbf{x}|^2 + |\mathbf{y}|^2)(\mathbf{x}\cdot\mathbf{y} + 2)$$
y el conjunto de aprendizaje en $\mathbb{R}^2$
$$X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, +1), (\mathbf{x}_3, -1), (\mathbf{x}_4, +1), (\mathbf{x}_5, -1)\}$$
siendo
$$\mathbf{x}_1 = (-1, 1), \mathbf{x}_2 = (0, 1), \mathbf{x}_3 = (-1, 0), \mathbf{x}_4 = (-1, -1), \mathbf{x}_5 = (0, -1)$$
se pide:

   a) Obtener la matriz Gramm $\mathbf{K}$ asociada a X
   b) Realizar una iteración completa del algoritmo Kernel Perceptron (de $\mathbf{x}_1$ a $\mathbf{x}_5$) partiendo del conjunto de pesos $\alpha = (0, 0, 0, 0, 0)$
   c) Clasificar la muestra $\mathbf{y} = (1, 1)$ de acuerdo a los pesos obtenidos en el apartado previo

8. Dada la función kernel $K(\mathbf{x},\mathbf{y}) = (\mathbf{x}\cdot\mathbf{y} + 2)^2$ y el conjunto de aprendizaje en $\mathbb{R}^3$ $X = \{(\mathbf{x}_1, -1), (\mathbf{x}_2, +1), (\mathbf{x}_3, -1), (\mathbf{x}_4, +1), (\mathbf{x}_5, +1)\}$, siendo $\mathbf{x}_1 = (0, 0, -1)$, $\mathbf{x}_2 = (1, 0, 0)$, $\mathbf{x}_3 = (0, 0, 1)$, $\mathbf{x}_4 = (0, 1, 0)$, $\mathbf{x}_5 = (0, -1, 0)$, se pide:

   a) Obtener la matriz Gramm $\mathbf{K}$ asociada a X
   b) Realizar iteraciones hasta la convergencia del algoritmo Kernel Perceptron partiendo del conjunto de pesos $\alpha = (1, 1, 1, 1, 1)$
   c) Clasificar la muestra $\mathbf{y} = (-1, 0, 0)$ de acuerdo a los pesos obtenidos en el apartado previo

9. Sea la siguiente función kernel, $K(\mathbf{x},\mathbf{y}) = (\mathbf{x}\cdot\mathbf{y}+1)^2$ y sea el siguiente conjunto de aprendizaje $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, +1), (\mathbf{x}_3, +1), (\mathbf{x}_4, -1), (\mathbf{x}_5, -1), (\mathbf{x}_6, -1)\}$ con:
$$\mathbf{x}_1 = \begin{bmatrix}0\\0\end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix}1\\0\end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix}0\\1\end{bmatrix} \quad \mathbf{x}_4 = \begin{bmatrix}-1\\-1\end{bmatrix} \quad \mathbf{x}_5 = \begin{bmatrix}1\\1\end{bmatrix} \quad \mathbf{x}_6 = \begin{bmatrix}2\\1\end{bmatrix}$$
Se pide:

   a) Obtén la matriz Gramm $\mathbf{K}$ de las muestras de entrenamiento
   b) Realiza una iteración (desde $\mathbf{x}_1$ hasta $\mathbf{x}_6$) del algoritmo Kernel Perceptron
   c) Clasifica la muestra de test $\mathbf{x} = (-1, 1)$ con el valor de los pesos α obtenido en el apartado anterior

10. Sea la siguiente función kernel, $K(\mathbf{x},\mathbf{y}) = (\mathbf{x}\cdot\mathbf{y}+1)^2$ y sea el siguiente conjunto de aprendizaje $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, +1), (\mathbf{x}_3, +1), (\mathbf{x}_4, -1), (\mathbf{x}_5, -1), (\mathbf{x}_6, -1)\}$ con:
$$\mathbf{x}_1 = \begin{bmatrix}1\\1\end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix}1\\0\end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix}0\\1\end{bmatrix} \quad \mathbf{x}_4 = \begin{bmatrix}2\\1\end{bmatrix} \quad \mathbf{x}_5 = \begin{bmatrix}1\\2\end{bmatrix} \quad \mathbf{x}_6 = \begin{bmatrix}0\\0\end{bmatrix}$$
Se pide:

   a) Obtén la matriz Gramm $\mathbf{K}$ de las muestras de entrenamiento
   b) Realiza una iteración (desde $\mathbf{x}_1$ hasta $\mathbf{x}_6$) del algoritmo Kernel Perceptron
   c) Clasifica la muestra de test $\mathbf{x} = (0, 2)$ con el valor de los pesos α obtenido en el apartado anterior

11. Tenemos la muestra de entrenamiento $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, -1), (\mathbf{x}_3, +1), (\mathbf{x}_4, -1)\}$, y la matriz Gramm:
$$\mathbf{K} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{2} & 1 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 \end{pmatrix}$$
Obtén el valor de las α_n que daría el algoritmo Kernel Perceptron para este conjunto de datos.

12. Sea la siguiente función kernel, $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 1)^d$, con $d = 2$. Sea el siguiente conjunto de aprendizaje $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, +1), (\mathbf{x}_3, +1), (\mathbf{x}_4, -1), (\mathbf{x}_5, -1)\}$, con: $\mathbf{x}_1 = (1, 1)$, $\mathbf{x}_2 = (1, 0)$, $\mathbf{x}_3 = (0, 1)$, $\mathbf{x}_4 = (2, 1)$, $\mathbf{x}_5 = (1, 2)$. Se tienen los siguientes pesos α_n: $\alpha_1 = \alpha_2 = \alpha_3 = 1$, $\alpha_4 = \alpha_5 = 0{,}5$. Clasifica la muestra de test $\mathbf{y} = (0, 2)$.

13. Se tiene el siguiente conjunto de datos en $\mathbb{R}^5$ para tres clases distintas:

	Muestra					Clase
x_1	1	-1	-5	1	3	1
x_2	-1	0	3	0	-2	1
x_3	3	1	5	1	3	2
x_4	2	2	2	-1	-2	2
x_5	-4	-2	-3	1	-3	3
x_6	-1	0	4	-2	4	3

Se ha realizado el proceso para obtener la matriz de transformación para LDA para el máximo número de dimensiones posible:

	$\mathrm{W}_{\mathrm{LDA}}$				
w_1	0,85	0,20	0,04	0,37	$-0{,}31$
w_2	$-0{,}38$	0,82	$-0{,}11$	0,39	0,15

Se pide:

a) Realizar la proyección de las muestras $x_1, \ldots, x_6$ mediante LDA a $\mathbb{R}^2$

b) Representar gráficamente la proyección

c) ¿Sería razonable aplicar proyección LDA a una única dimensión? Razona la respuesta

14. Se dispone del siguiente conjunto de datos sobre $\mathbb{R}^5$ para un problema de tres clases:

	Muestra					Clase
x_1	1	1	1	1	1	1
x_2	2	1	-1	0	1	1
x_3	-1	2	-1	1	0	2
x_4	1	-1	1	2	-1	2
x_5	0	-1	-1	0	1	3
x_6	0	1	-2	2	1	3

Se han realizado los procesos para obtener proyecciones LDA al espacio $\mathbb{R}^2$, dando los siguientes resultados:

$$W_{lda} = \begin{bmatrix} 0{,}6 & -0{,}1 \\ 0{,}4 & -0{,}4 \\ 0{,}4 & -0{,}2 \\ -0{,}2 & 0{,}3 \\ 0{,}5 & 0{,}9 \end{bmatrix}$$

Se pide:

a) Obtener la proyección de los datos mediante el uso de LDA

b) Realizar una representación gráfica de la proyección

c) ¿Aplicar LDA a una única dimensión daría buenos resultados? Razona tu respuesta

15. Se tiene el siguiente conjunto de datos para dos clases en $\mathbb{R}^2$:

Muestra			Clase
x_1	1,50	1,76	A
x_2	$-1{,}11$	$-0{,}01$	A
x_3	5,16	2,85	A
x_4	3,26	$-4{,}79$	B
x_5	$-8{,}81$	0,19	B

Se pide:

a) Calcular las matrices S_b y S_w que se emplearían para calcular la proyección LDA a partir de las muestras de $\mathbb{R}^2$ dadas

b) Si se quiere emplear un clasificador lineal, ¿bastaría con usar los datos originales o sería necesario aplicar LDA? Razona la respuesta

16. Se dispone de un conjunto de muestras en $\mathbb{R}^3$ clasificadas en cuatro clases:

n	1	2	3	4	5	6	7	8
x_1	4	4	-2	2	-2	2	-4	-4
x_2	4	4	2	-2	2	-2	-4	-4
x_3	1	-1	-1	1	1	-1	-1	1
c	A	B	D	A	C	B	D	C

Por otra parte se ha calculado LDA, obteniéndose los siguientes vectores de proyección ordenados por valor propio generalizado de mayor (w_1) a menor (w_3):

	$\mathrm{W}_{\mathrm{LDA}}$		
w_1	0	0	1
w_2	1	0	0
w_3	0	1	0

Se pide:

a) Calcula la proyección de las muestras mediante LDA a $\mathbb{R}^2$

b) ¿Consideras que la proyección LDA es adecuada para minimizar el error de clasificación?

Soluciones

1. En convergencia $\alpha_1 = \alpha_3 = \alpha_4 = \alpha_5 = 1$ y $\alpha_2 = 2$.

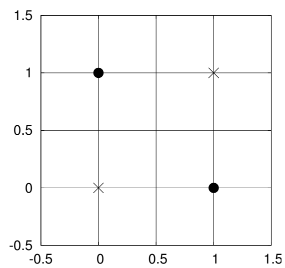
2. La matriz Gramm es:

$$\mathbf{K} = \begin{bmatrix} 1 & 1/3 & 1/2 & 1/2 & 1/3 \\ 1/3 & 1 & 1/3 & 1/3 & 1/3 \\ 1/2 & 1/3 & 1 & 1/3 & 1/2 \\ 1/2 & 1/3 & 1/3 & 1 & 1/2 \\ 1/3 & 1/3 & 1/2 & 1/2 & 1 \end{bmatrix}$$

Se puede demostrar que K es un kernel verificando que la matriz Gramm genérica para dos pares de muestras cualesquiera es semidefinida positiva.

3. (Examen Recuperación Julio 2017)

   a) Como se puede observar en la figura el conjunto de entrenamiento no es linealmente separable.



   b) La matriz de Gramm sobre el conjunto de entrenamiento para la función Kernel definida en el enunciado es

$$\mathbf{K} = \begin{array}{ccccc} & (0,0) & (0,1) & (1,0) & (1,1) \\ (0,0) & 1 & 1/2 & 1/2 & 1/3 \\ (0,1) & 1/2 & 1 & 1/2 & 1/2 \\ (1,0) & 1/2 & 1/2 & 1 & 1/2 \\ (1,1) & 1/3 & 1/2 & 1/2 & 1 \end{array}$$

La traza del algoritmo Kernel Perceptron es la siguiente:

	$g(\mathbf{x})$	$g(\mathbf{x}_n)$	c_n	error	α
$\mathbf{x}_1$	0	0	$+1$	Sí	$(1,0,0,0)$
$\mathbf{x}_2$	$K(\mathbf{x}_1,\mathbf{x}) + 1$	$3/2$	-1	Sí	$(1,1,0,0)$
$\mathbf{x}_3$	$K(\mathbf{x}_1,\mathbf{x}) + 1 - K(\mathbf{x}_2,\mathbf{x}) - 1$	0	-1	Sí	$(1,1,1,0)$
$\mathbf{x}_4$	$K(\mathbf{x}_1,\mathbf{x}) + 1 - K(\mathbf{x}_2,\mathbf{x}) - 1 - K(\mathbf{x}_3,\mathbf{x}) - 1$	$-5/6$	$+1$	Sí	$(1,1,1,1)$

   c) Considerando el vector de pesos $\alpha = (1,1,1,1)$ obtenidos, la función discriminante es:

$$\begin{aligned} g(\mathbf{x}) &= K(\mathbf{x}_1,\mathbf{x}) + 1 - K(\mathbf{x}_2,\mathbf{x}) - 1 - K(\mathbf{x}_3,\mathbf{x}) - 1 + K(\mathbf{x}_4,\mathbf{x}) + 1 \\ &= K(\mathbf{x}_1,\mathbf{x}) - K(\mathbf{x}_2,\mathbf{x}) - K(\mathbf{x}_3,\mathbf{x}) + K(\mathbf{x}_4,\mathbf{x}) \end{aligned}$$

	$g(\mathbf{x}_n)$					c_i	error
$\mathbf{x}_1$	1	-1/2	-1/2	$+1/3 =$	1/3	$+1$	No
$\mathbf{x}_2$	1/2	-1	-1/2	$+1/2 =$	-1/2	-1	No
$\mathbf{x}_3$	1/2	-1/2	-1	$+1/2 =$	-1/2	-1	No
$\mathbf{x}_4$	1/3	-1/2	-1/2	$+1 =$	1/3	$+1$	No

Por tanto, el error de clasificación es cero y la función Kernel empleada proyecta el conjunto de entrenamiento a un espacio donde son linealmente separables.

4. (Examen Junio 2017)

   a) $g(\mathbf{x}_1) = 0$, $c_1 g(\mathbf{x}_1) = 0 \leq 0$, $\alpha = (1, 0, 0, 0)$
   $g(\mathbf{x}_2) = 1 \cdot (-1) \exp(\mathbf{x}_1^t \cdot \mathbf{x}_2) + 1 \cdot (-1) = -\exp(-1) - 1$, $c_2 g(\mathbf{x}_2) = -(\exp(-1) + 1) \leq 0$, $\alpha = (1, 1, 0, 0)$
   $g(\mathbf{x}_3) = 1 \cdot (-1) \exp(\mathbf{x}_1^t \cdot \mathbf{x}_3) + 1 \cdot (+1) \exp(\mathbf{x}_2^t \cdot \mathbf{x}_3) + 1 \cdot (-1) + 1 \cdot (+1) = -\exp(-1) + \exp(1)$, $c_3 g(\mathbf{x}_3) = \exp(-1) - \exp(1) \leq 0$, $\alpha = (1, 1, 1, 0)$
   $g(\mathbf{x}_4) = 1 \cdot (-1) \exp(\mathbf{x}_1^t \cdot \mathbf{x}_4) + 1 \cdot (+1) \exp(\mathbf{x}_2^t \cdot \mathbf{x}_4) + 1 \cdot (-1) \exp(\mathbf{x}_3^t \cdot \mathbf{x}_4) + 1 \cdot (-1) + 1 \cdot (+1) + 1 \cdot (-1) = -\exp(0) + \exp(0) - \exp(-1) - 1 = -(\exp(-1) + 1)$, $c_4 g(\mathbf{x}_4) = -(\exp(-1) + 1) \leq 0$, $\alpha = (1, 1, 1, 1)$
   Resultado final: $\alpha = (1, 1, 1, 1)$

   b) No sería necesario, pues las muestras ya son linealmente separables.

5. (Examen Recuperación Junio 2016)

   a) Matriz Gramm:

$$\mathbf{K} = \begin{vmatrix} 25 & 1 & 9 & 9 \\ 1 & 4 & 0 & 1 \\ 9 & 0 & 9 & 4 \\ 9 & 1 & 4 & 4 \end{vmatrix}$$

   b) Iteraciones algoritmo Kernel Perceptron

```
muestra 1, clase 1, fdl=0.000000  0.000000 --> error
muestra 2, clase -1, fdl=2.000000  -2.000000 --> error
muestra 3, clase 1, fdl=9.000000  9.000000 --> ok
muestra 4, clase -1, fdl=8.000000  -8.000000 --> error

alfas:
1.000000 1.000000 0.000000 1.000000
```

   c) Clasificación de $\mathbf{y}_1$:
   $g(\mathbf{y}_1) = 25 - 0 + 0 - 9 + 1 - 1 + 0 - 1 = 15$ por lo tanto $\mathbf{y}_1$ pertenece a la clase $+1$
   Clasificación de $\mathbf{y}_2$:
   $g(\mathbf{y}_2) = 1 - 1 + 0 - 1 + 1 - 1 + 0 - 1 = -2$ por lo tanto $\mathbf{y}_2$ pertenece a la clase -1

   d) No, no es necesario puesto que las muestras de entrenamiento en el espacio original ya son linealmente separables

6. (Examen Abril 2016)

   a) Matriz Gramm:

$$\mathbf{K} = \begin{vmatrix} 25 & 4 & 9 & 36 & 16 \\ 4 & 16 & 16 & 4 & 9 \\ 9 & 16 & 25 & 16 & 16 \\ 36 & 4 & 16 & 64 & 25 \\ 16 & 9 & 16 & 25 & 16 \end{vmatrix}$$

   b) Iteraciones algoritmo Kernel Perceptron

```
muestra 1, clase 1, fdl=0.000000  0.000000 --> error
muestra 2, clase 1, fdl=5.000000  5.000000 --> ok
muestra 3, clase -1, fdl=10.000000  -10.000000 --> error
muestra 4, clase -1, fdl=20.000000  -20.000000 --> error
muestra 5, clase 1, fdl=-26.000000  -26.000000 --> error

alfas:
1.000000 0.000000 1.000000 1.000000 1.000000
```

   c) Clasificación de $\mathbf{y}$: $g(\mathbf{y}) = -7$ por lo tanto $\mathbf{y}$ pertenece a la clase -1

7. (Examen Recuperación Junio 2015)

   a) $\mathbf{K} = \begin{vmatrix} 16 & 9 & 9 & 8 & 3 \\ 9 & 6 & 4 & 3 & 2 \\ 9 & 4 & 6 & 9 & 4 \\ 8 & 3 & 9 & 16 & 9 \\ 3 & 2 & 4 & 9 & 6 \end{vmatrix}$

   b) $\mathbf{x}_1$: $g(\mathbf{x}_1) = 0$, $c_1 g(\mathbf{x}_1) = 0 \leq 0$, $\alpha_1 = 1$, $\alpha = (1, 0, 0, 0, 0)$
   $\mathbf{x}_2$: $g(\mathbf{x}_2) = 10$, $c_2 g(\mathbf{x}_2) = 10 > 0$
   $\mathbf{x}_3$: $g(\mathbf{x}_3) = 10$, $c_3 g(\mathbf{x}_3) = -10 \leq 0$, $\alpha_3 = 1$, $\alpha = (1, 0, 1, 0, 0)$
   $\mathbf{x}_4$: $g(\mathbf{x}_4) = -1$, $c_4 g(\mathbf{x}_4) = -1 \leq 0$, $\alpha_4 = 2$, $\alpha = (1, 0, 1, 1, 0)$
   $\mathbf{x}_5$: $g(\mathbf{x}_5) = 9$, $c_5 g(\mathbf{x}_5) = -9 \leq 0$, $\alpha_5 = 1$, $\alpha = (1, 0, 1, 1, 1)$

   c) $K(\mathbf{x}_1, \mathbf{y}) = 8$, $K(\mathbf{x}_2, \mathbf{y}) = 9$, $K(\mathbf{x}_3, \mathbf{y}) = 3$, $K(\mathbf{x}_4, \mathbf{y}) = 0$, $K(\mathbf{x}_5, \mathbf{y}) = 3$
   Por tanto, $g(\mathbf{y}) = 1{\cdot}1{\cdot}8+0{\cdot}1{\cdot}9+1{\cdot}(-1){\cdot}3+1{\cdot}1{\cdot}0+1{\cdot}(-1){\cdot}3+1{\cdot}1+0{\cdot}1+1{\cdot}(-1)+1{\cdot}1+1{\cdot}(-1) = 8-3-3 = 2$, $c(\mathbf{y}) = +1$

8. (Examen Marzo 2015)

   a) $\mathbf{K} = \begin{vmatrix} 9 & 4 & 1 & 4 & 4 \\ 4 & 9 & 4 & 4 & 4 \\ 1 & 4 & 9 & 4 & 4 \\ 4 & 4 & 4 & 9 & 1 \\ 4 & 4 & 4 & 1 & 9 \end{vmatrix}$

   b) $\mathbf{x}_1$: $g(\mathbf{x}_1) = 3$, $c_1 g(\mathbf{x}_1) = -3 \leq 0$, $\alpha_1 = 2$, $\alpha = (2, 1, 1, 1, 1)$
   $\mathbf{x}_2$: $g(\mathbf{x}_2) = 5$, $c_2 g(\mathbf{x}_2) = 5 > 0$
   $\mathbf{x}_3$: $g(\mathbf{x}_3) = 1$, $c_3 g(\mathbf{x}_3) = -1 \leq 0$, $\alpha_3 = 2$, $\alpha = (2, 1, 2, 1, 1)$
   $\mathbf{x}_4$: $g(\mathbf{x}_4) = -3$, $c_4 g(\mathbf{x}_4) = -3 \leq 0$, $\alpha_4 = 2$, $\alpha = (2, 1, 2, 2, 1)$
   $\mathbf{x}_5$: $g(\mathbf{x}_5) = -1$, $c_5 g(\mathbf{x}_5) = -1 \leq 0$, $\alpha_5 = 2$, $\alpha = (2, 1, 2, 2, 2)$

   $\mathbf{x}_1$: $g(\mathbf{x}_1) = 1$, $c_1 g(\mathbf{x}_1) = -1 \leq 0$, $\alpha_1 = 3$, $\alpha = (3, 1, 2, 2, 2)$
   $\mathbf{x}_2$: $g(\mathbf{x}_2) = 5$, $c_2 g(\mathbf{x}_2) = 5 > 0$
   $\mathbf{x}_3$: $g(\mathbf{x}_3) = -1$, $c_3 g(\mathbf{x}_3) = 1 > 0$
   $\mathbf{x}_4$: $g(\mathbf{x}_4) = 4$, $c_4 g(\mathbf{x}_4) = 4 > 0$
   $\mathbf{x}_5$: $g(\mathbf{x}_5) = 4$, $c_5 g(\mathbf{x}_5) = 4 > 0$

   $\mathbf{x}_1$: $g(\mathbf{x}_1) = -9$, $c_1 g(\mathbf{x}_1) = 9 > 0$

   c) $K(\mathbf{x}_1, \mathbf{y}) = 4$, $K(\mathbf{x}_2, \mathbf{y}) = 1$, $K(\mathbf{x}_3, \mathbf{y}) = 4$, $K(\mathbf{x}_4, \mathbf{y}) = 4$, $K(\mathbf{x}_5, \mathbf{y}) = 4$
   Por tanto, $g(\mathbf{y}) = -3$, $c(\mathbf{y}) = -1$

9. (Examen Recuperación Junio 2014)

   a) $\mathbf{K} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 4 & 1 & 0 & 4 & 9 \\ 1 & 1 & 4 & 0 & 4 & 4 \\ 1 & 0 & 0 & 9 & 1 & 4 \\ 1 & 4 & 4 & 1 & 9 & 16 \\ 1 & 9 & 4 & 4 & 16 & 36 \end{bmatrix}$

   b) $\mathbf{x}_1 \rightarrow$ error $\Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0)$
   $\mathbf{x}_4 \rightarrow$ error $\Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0)$
   $\mathbf{x}_5 \rightarrow$ error $\Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0)$

   Final: $\alpha = (1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0)$

   c) $g(\mathbf{x}) = -2 \Rightarrow c(\mathbf{x}) = -1$

10. (Examen Abril 2014)

   a) $\mathbf{K} = \begin{bmatrix} 9 & 4 & 4 & 16 & 16 & 1 \\ 4 & 4 & 1 & 9 & 4 & 1 \\ 4 & 1 & 4 & 4 & 9 & 1 \\ 16 & 9 & 4 & 36 & 25 & 1 \\ 16 & 4 & 9 & 25 & 36 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$

   b) $\mathbf{x}_1 \to \text{error} \Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0)$
   $\mathbf{x}_4 \to \text{error} \Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0)$
   $\mathbf{x}_6 \to \text{error} \Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1)$

   Final: $\alpha = (1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1)$

   c) $g(\mathbf{x}) = 9 + 1 - 9 - 1 - 1 - 1 = -2 \Rightarrow c(\mathbf{x}) = -1$

11. (Examen Recuperación Junio 2013)

   Partimos de $\alpha = \{0, 0, 0, 0\}$, usando la función a optimizar $g(\mathbf{x}) = \sum_{n=1}^{4} \alpha_n \cdot c_n \cdot K(\mathbf{x}, \mathbf{x}_n) + \sum_{n=1}^{4} \alpha_n \cdot c_n$.

   1^a iteración

   - $g(\mathbf{x}_1) = 0 \to c_1 g(\mathbf{x}_1) = 0 \leq 0 \to \alpha_1 = 1 \to \alpha = \{1, 0, 0, 0\}$
   - $g(\mathbf{x}_2) = K(\mathbf{x}_2, \mathbf{x}_1) + \alpha_1 = \frac{1}{2} + 1 = \frac{3}{2} \to c_2 g(\mathbf{x}_2) = -\frac{3}{2} \leq 0 \to \alpha_2 = 1 \to \alpha = \{1, 1, 0, 0\}$
   - $g(\mathbf{x}_3) = K(\mathbf{x}_3, \mathbf{x}_1) - K(\mathbf{x}_3, \mathbf{x}_2) + \alpha_1 - \alpha_2 = \frac{1}{3} - \frac{1}{2} + 1 - 1 = -\frac{1}{6} \to c_3 g(\mathbf{x}_3) = -\frac{1}{6} \leq 0 \to \alpha_3 = 1 \to \alpha = \{1, 1, 1, 0\}$
   - $g(\mathbf{x}_4) = K(\mathbf{x}_4, \mathbf{x}_1) - K(\mathbf{x}_4, \mathbf{x}_2) + K(\mathbf{x}_4, \mathbf{x}_3) + \alpha_1 - \alpha_2 + \alpha_3 = \frac{2}{3} - \frac{1}{3} + \frac{1}{3} + 1 - 1 + 1 = \frac{5}{3} \to c_4 g(\mathbf{x}_4) = -\frac{5}{3} \leq 0 \to \alpha_4 = 1 \to \alpha = \{1, 1, 1, 1\}$

   2^a iteración

   - $g(\mathbf{x}_1) = K(\mathbf{x}_1, \mathbf{x}_1) - K(\mathbf{x}_1, \mathbf{x}_2) + K(\mathbf{x}_1, \mathbf{x}_3) - K(\mathbf{x}_1, \mathbf{x}_4) + \alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{2}{3} + 1 - 1 + 1 - 1 = \frac{1}{6} \to c_1 g(\mathbf{x}_1) = \frac{1}{6} > 0 \to \alpha = \{1, 1, 1, 1\}$
   - $g(\mathbf{x}_2) = K(\mathbf{x}_2, \mathbf{x}_1) - K(\mathbf{x}_2, \mathbf{x}_2) + K(\mathbf{x}_2, \mathbf{x}_3) - K(\mathbf{x}_2, \mathbf{x}_4) + \alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 = \frac{1}{2} - 1 + \frac{1}{2} - \frac{1}{3} + 1 - 1 + 1 - 1 = -\frac{1}{3} \to c_2 g(\mathbf{x}_2) = \frac{1}{3} > 0 \to \alpha = \{1, 1, 1, 1\}$
   - $g(\mathbf{x}_3) = K(\mathbf{x}_3, \mathbf{x}_1) - K(\mathbf{x}_3, \mathbf{x}_2) + K(\mathbf{x}_3, \mathbf{x}_3) - K(\mathbf{x}_3, \mathbf{x}_4) + \alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 = \frac{1}{3} - \frac{1}{2} + 1 - \frac{1}{3} + 1 - 1 + 1 - 1 = \frac{1}{2} \to c_3 g(\mathbf{x}_3) = \frac{1}{2} > 0 \to \alpha = \{1, 1, 1, 1\}$
   - $g(\mathbf{x}_4) = K(\mathbf{x}_4, \mathbf{x}_1) - K(\mathbf{x}_4, \mathbf{x}_2) + K(\mathbf{x}_4, \mathbf{x}_3) - K(\mathbf{x}_4, \mathbf{x}_4) + \alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 = \frac{2}{3} - \frac{1}{3} + \frac{1}{3} - 1 + 1 - 1 + 1 - 1 = -\frac{1}{3} \to c_4 g(\mathbf{x}_4) = \frac{1}{3} > 0 \to \alpha = \{1, 1, 1, 1\}$

   Al no haber cambios en esta segunda iteración, se tiene finalmente que $\alpha = \{1, 1, 1, 1\}$.

12. (Examen Junio 2013)

$$g(\mathbf{y}) = \sum_{n=1}^{n} \alpha_n c_n K(\mathbf{y}, \mathbf{x}_n) + \sum_{n=1}^{n} \alpha_n c_n =$$

$$1 \cdot (+1) \cdot K(\mathbf{y}, \mathbf{x}_1) + 1 \cdot (+1) \cdot K(\mathbf{y}, \mathbf{x}_2) + 1 \cdot (+1) \cdot K(\mathbf{y}, \mathbf{x}_3) + (0{,}5) \cdot (-1) \cdot K(\mathbf{y}, \mathbf{x}_4) + (0{,}5) \cdot (-1) \cdot K(\mathbf{y}, \mathbf{x}_5)$$

$$+1 \cdot (+1) + 1 \cdot (+1) + 1 \cdot (+1) + (0{,}5) \cdot (-1) + (0{,}5) \cdot (-1) =$$

$$(+1) \cdot 9 + (+1) \cdot 1 + (+1) \cdot 9 + (-0{,}5) \cdot 9 + (-0{,}5) \cdot 25 + (+1) + (+1) + (+1) + (-0{,}5) + (-0{,}5) = 19 - 17 + 2 = 4$$

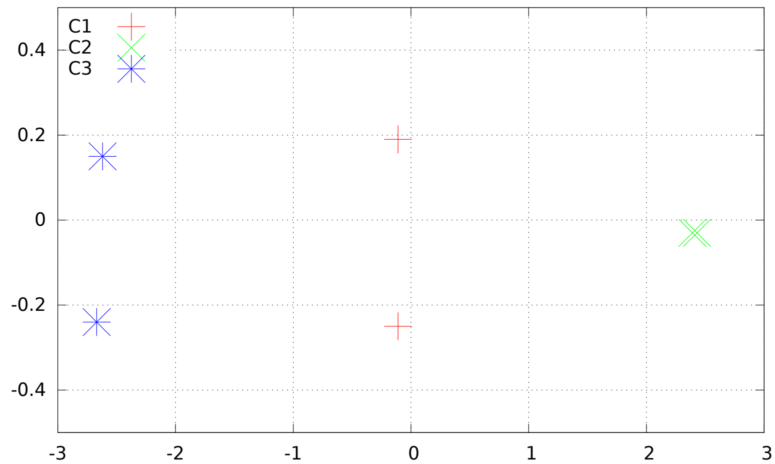
   Por tanto, al ser $g(\mathbf{y}) > 0$, se clasifica en $+1$.

13. (Examen Abril 2017)

   a) El resultado de la proyección sería:

x'_1	$-0{,}11$	$0{,}19$
x'_2	$-0{,}11$	$-0{,}25$
x'_3	$2{,}39$	$-0{,}03$
x'_4	$2{,}43$	$-0{,}03$
x'_5	$-2{,}62$	$0{,}15$
x'_6	$-2{,}67$	$-0{,}24$

b) La representación gráfica sería:



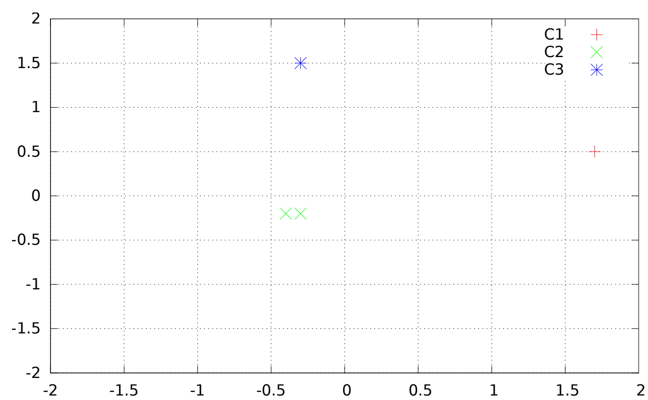
c) Sí, es razonable, ya que en ese caso se tomaría sólo la primera componente del resultado de la proyección LDA, en la cual sigue habiendo una clara separabilidad entre las muestras de las distintas clases

14. (Examen Abril 2016)

a) Proyección:

	Muestra		Clase
$x_1 W_{lda}$	$1{,}7$	$0{,}5$	1
$x_2 W_{lda}$	$1{,}7$	$0{,}5$	1
$x_3 W_{lda}$	$-0{,}4$	$-0{,}2$	2
$x_4 W_{lda}$	$-0{,}3$	$-0{,}2$	2
$x_5 W_{lda}$	$-0{,}3$	$1{,}5$	3
$x_6 W_{lda}$	$-0{,}3$	$1{,}5$	3

b)



c) No, pues la proyección de la primera dimensión haría indistinguibles los elementos de las clases 2 y 3

15. (Examen Marzo 2018)

a) Para calcular S_b es preciso calcular la media global de los datos proyectados $\bar{\mathbf{x}}^t = (0, 0)$ y la media por clase:

$$\bar{\mathbf{x}}_A = \begin{pmatrix} 1{,}85 \\ 1{,}53 \end{pmatrix} \qquad \bar{\mathbf{x}}_B = \begin{pmatrix} -2{,}78 \\ -2{,}30 \end{pmatrix}$$

A partir de aquí, como $S_b = \sum_{c=1}^{C} N_c(\bar{\mathbf{x}}_c - \bar{\mathbf{x}}) \cdot (\bar{\mathbf{x}}_c - \bar{\mathbf{x}})^t$, tendremos:

$$S_b = 3 \cdot \begin{pmatrix} 1{,}85 \\ 1{,}53 \end{pmatrix} \cdot (1{,}85, 1{,}53) + 2 \cdot \begin{pmatrix} -2{,}78 \\ -2{,}30 \end{pmatrix} \cdot (-2{,}78 - 2{,}30) = \begin{pmatrix} 25{,}613 & 21{,}233 \\ 21{,}233 & 17{,}603 \end{pmatrix}$$

Para calcular S_w es preciso calcular las matrices de covarianzas de los datos proyectados (x'_i), que son las siguientes:

$$\Sigma_A = \frac{1}{3}\left((x'_1 - \bar{\mathbf{x}}_A) \cdot (x'_1 - \bar{\mathbf{x}}_A)^t + (x'_2 - \bar{\mathbf{x}}_A) \cdot (x'_2 - \bar{\mathbf{x}}_A)^t + (x'_3 - \bar{\mathbf{x}}_A) \cdot (x'_3 - \bar{\mathbf{x}}_A)^t\right) =$$

$$\frac{1}{3}\left(\begin{pmatrix} -0{,}35 \\ 0{,}23 \end{pmatrix} \cdot (-0{,}35, 0{,}23) + \begin{pmatrix} -2{,}96 \\ -1{,}54 \end{pmatrix} \cdot (-2{,}96, -1{,}54) + \begin{pmatrix} 3{,}31 \\ 1{,}32 \end{pmatrix} \cdot (3{,}31, 1{,}32)\right) = \begin{pmatrix} 6{,}61 & 2{,}95 \\ 2{,}95 & 1{,}39 \end{pmatrix}$$

$$\Sigma_B = \frac{1}{2}\left((x'_4 - \bar{\mathbf{x}}_B) \cdot (x'_4 - \bar{\mathbf{x}}_B)^t + (x'_5 - \bar{\mathbf{x}}_B) \cdot (x'_5 - \bar{\mathbf{x}}_B)^t\right) =$$

$$\frac{1}{2}\left(\begin{pmatrix} 6{,}04 \\ -2{,}49 \end{pmatrix} \cdot (6{,}04, -2{,}49) + \begin{pmatrix} -6{,}04 \\ 2{,}49 \end{pmatrix} \cdot (-6{,}04, 2{,}49)\right) = \begin{pmatrix} 36{,}42 & -15{,}03 \\ -15{,}03 & 6{,}20 \end{pmatrix}$$

Por tanto:

$$S_w = \Sigma_A + \Sigma_B = \begin{pmatrix} 43{,}03 & -12{,}08 \\ -12{,}08 & 7{,}59 \end{pmatrix}$$

b) No sería necesario proyectar a LDA; si tomamos los datos originales, se puede definir una frontera de decisión lineal entre ambas clases empleando, por ejemplo, $x_1 + x_2 = -1{,}5$ como frontera; de esta forma, aquellas muestras cuyas componentes sumen más de -1.5 quedan para la clase A y las que suman menos quedan para la clase B.

16. (Examen Recuperación P1 Junio 2018)

a) Proyectamos sobre los dos vectores LDA

n	1	2	3	4	5	6	7	8
x_1	1	-1	-1	1	1	-1	-1	1
x_2	4	4	-2	2	-2	2	-4	-4
c	A	B	D	A	C	B	D	C

b) La proyección LDA separa linealmente los datos de diferentes clases, y por tanto es adecuada.