

Relación 6 de problemas

1. Se dispone de 7 observaciones bidimensionales (x_1, x_2) correspondientes a dos grupos indicados por la variable y :

Obs.	x_1	x_2	y
1	3	4	-1
2	2	2	-1
3	4	4	-1
4	1	4	-1
5	2	1	1
6	4	3	1
7	4	1	1

Responde a las siguientes preguntas (es posible averiguar las respuestas sin usar el ordenador y posteriormente comprobar con **R** si son correctas):

- (a) Representa gráficamente las observaciones. ¿Es una muestra linealmente separable?
- (b) ¿Cuál es la ecuación del hiperplano que maximiza el margen? ¿Cuál es el valor del margen para este hiperplano óptimo?
- (c) ¿Cuáles son los vectores soporte para el hiperplano óptimo?
- (d) ¿Cómo cambia el hiperplano óptimo si en la observación 7 se cambia el valor 4 de x_1 por el valor 10?

2. El *data frame* `datos.SVM` de **R** contiene una muestra de entrenamiento $(x_1, y_1), \dots, (x_n, y_n)$ de $n = 400$ observaciones independientes, donde $x_i \in \mathbb{R}^2$ es un vector bidimensional de variables e $y_i \in \{-1, 1\}$ es una variable que indica el grupo al que corresponde la observación x_i . Hay 200 observaciones de cada grupo. Los datos se pueden bajar en [este enlace](#) (guarda el fichero en el directorio de trabajo y usa posteriormente el comando `load('datos.SVM.RData')` para cargarlos en **R**).

- (a) Calcula la regla de clasificación lineal basada en máquinas de vectores soporte para $C = 10$. Representa gráficamente los datos y la frontera que separa los grupos de acuerdo con esta regla.
- (b) Considera los valores $C = 10, 50, 100, 200$. Estima el porcentaje de acierto en la clasificación para cada valor usando validación cruzada (dividiendo la muestra en cuatro partes y usando cada una de ellas como muestra de test). Completa la siguiente tabla con los resultados obtenidos:

C	10	50	100	200
% acierto estimado				

- (c) Fija $C = 10$. Estima el porcentaje de acierto en la clasificación (el mismo método de validación cruzada) cuando se cambia el núcleo lineal por otros tres núcleos y completa la tabla siguiente:

Núcleo	$(x'y + 1)^2$	$\exp(-\ x - y\ ^2/2)$	$\exp(-\ x - y\ ^2/20)$
% acierto estimado			

Consulta la ayuda de **R** para saber cómo seleccionar estos núcleos.