

Memoria Práctica 1. SVM

Marcos Esteve Casademunt, Enric Bonet Cortés

Noviembre 2018

Índice

1. trSep y tr	1
2. Conjunto no separable	4
3. Corpus de dos clases: SPAM	6
4. Corpus multi-clase: USPS	7

Índice de figuras

1. Conjunto separable	2
2. Conjunto no separable	3
3. Tolerancia del margen para $C = 1$	4
4. Tolerancia del margen para $C = 10$	5
5. Tolerancia del margen para $C = 100$	5
6. Tolerancia del margen para $C = 1000$	6
7. Ejemplo de muestras de la clase 0 a 9	8
8. Ejemplo de muestras de la clase 0 a 9 que son vectores soporte	9

1. trSep y tr

Disponemos de un conjunto de datos separables(trSep.dat y trSepLabels.dat) y un conjunto de datos no separables(tr.dat y trLabels.dat). Una vez cargados los datos, entrenamos un SVM con kernel lineal y $C = 1000$. **res = svmtrain(trLabels, tr, '-t 0 -c 1000')**

Una vez entrenada el modelo procedemos a obtener los siguientes parámetros:

- Multiplicadores de Lagrange α : res.sv_coef (multiplicado por la etiqueta -1 o 1)
- Vectores de soporte: res.SVs

- Vector de pesos $\theta : res.sv_coef^t * res.SVs$
- Umbral θ_0 : res.rho
- Margen $1/(\theta * \theta^t)$

En cuanto a los parámetros de la recta de separación:

Ecuación de la recta:

$$y = -\frac{\theta_1}{\theta_2}x - \frac{\theta_0}{\theta_2} \quad (1)$$

Ecuación de las rectas que definen los márgenes:

$$y = -\frac{\theta_1}{\theta_2}x - \frac{\theta_0 \pm 1}{\theta_2} \quad (2)$$

Una vez calculadas las ecuaciones de la recta y los respectivos márgenes pasamos a representar las gráficas para los dos conjuntos de datos. Se marcará con un + los vectores soporte

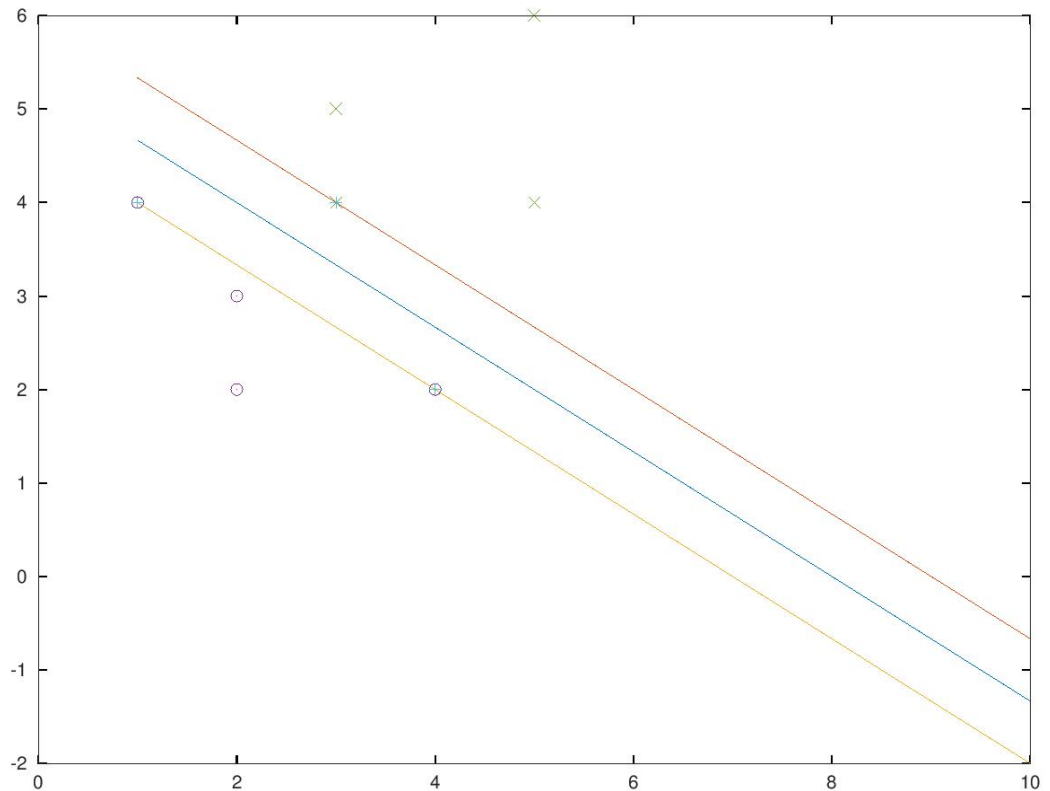


Figura 1: Conjunto separable

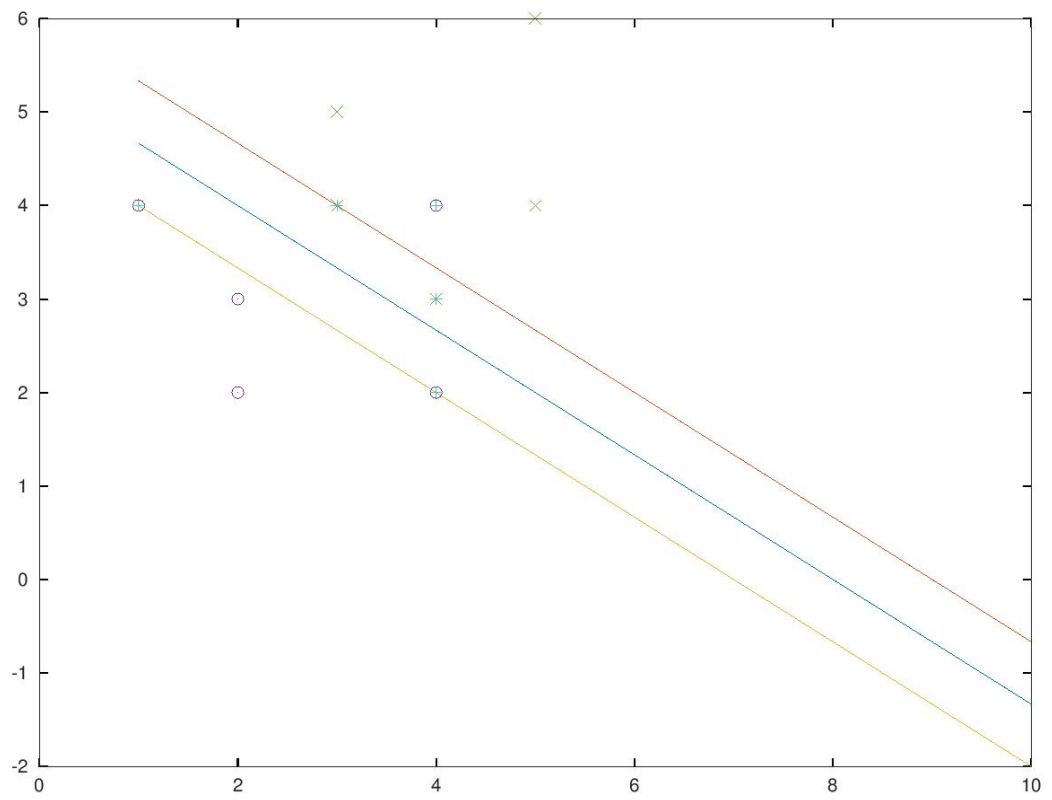


Figura 2: Conjunto no separable

Como se puede observar, en el conjunto no separable existen algunos vectores soporte erróneos. Los cuales los estudiaremos en la siguiente sección.

2. Conjunto no separable

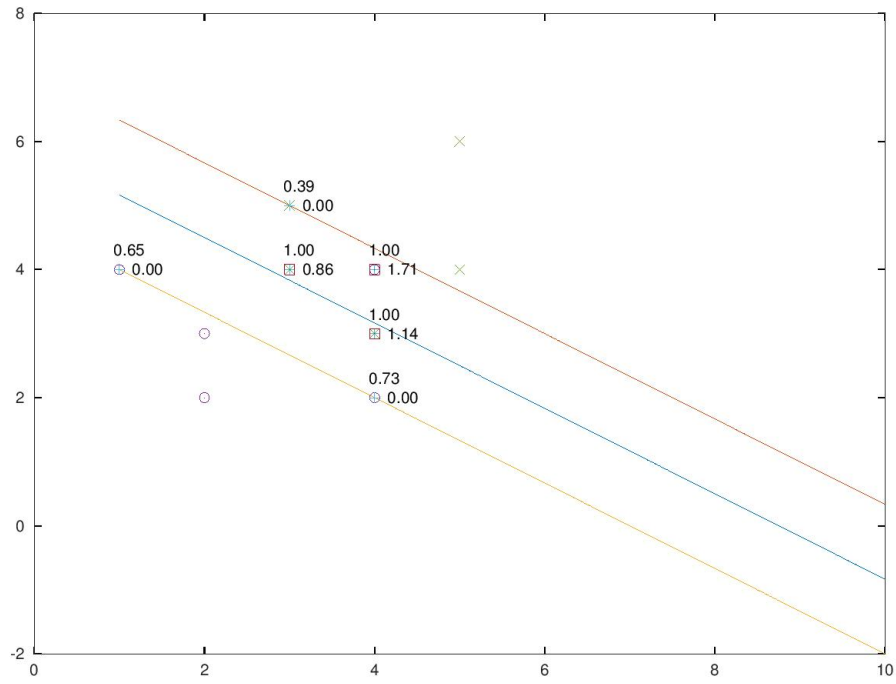
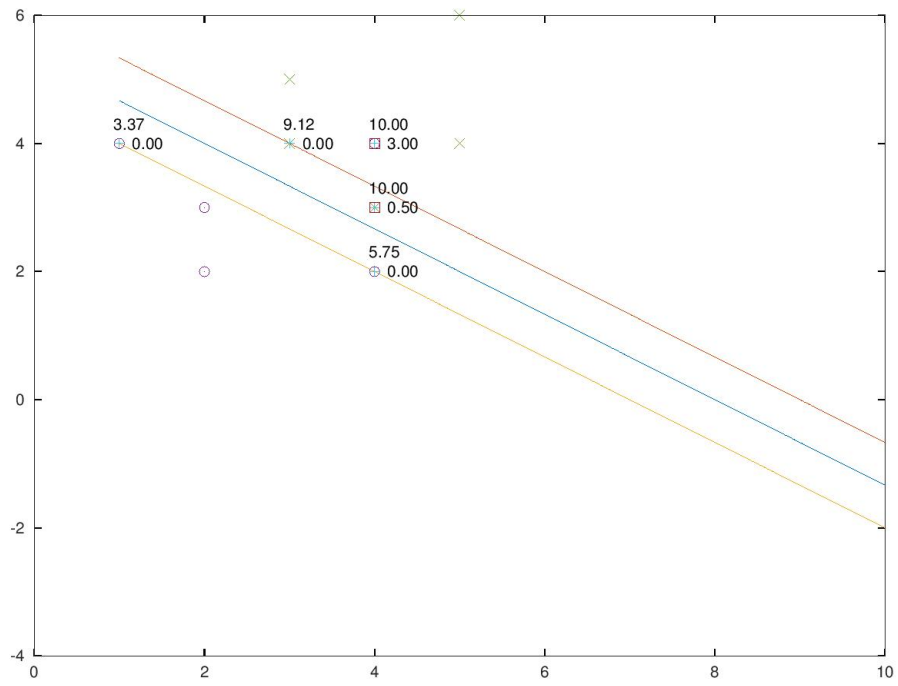


Figura 3: Tolerancia del margen para $C = 1$



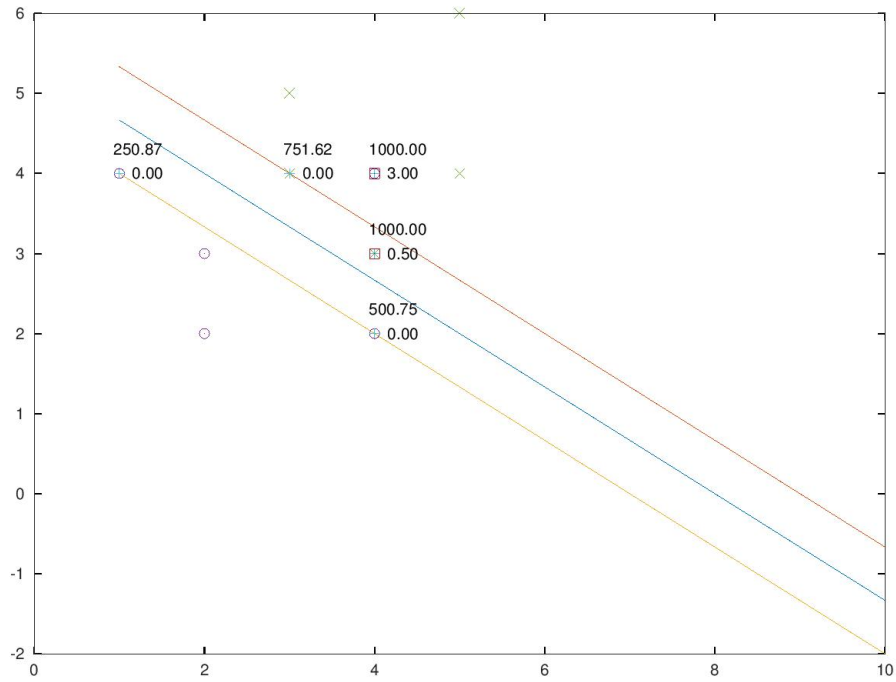


Figura 6: Tolerancia del margen para $C = 1000$

3. Corpus de dos clases: SPAM

Se nos proporciona un corpus de dos clases dividido en train y test. Este corpus está formado por 3220 muestras de aprendizaje y 1381 de test. Analizaremos la combinación de C y Kernel con el fin de obtener el mínimo error.

c	precisión	intervalo(95 %)
0.01	1	0
0.1	0.999	0.001
1	0.997	0.002
10	0.997	0.002
100	0.997	0.002
1000	0.997	0.002

Cuadro 1: Kernel Lineal

c	precisión	intervalo(95 %)
0.01	0.987	0.005
0.1	0.990	0.005
1	0.992	0.004
10	0.996	0.003
100	0.997	0.002
1000	0.997	0.002

Cuadro 2: Kernel polinomial de grado 3

c	precisión	intervalo(95 %)
0.01	0.606	0.025
0.1	0.737	0.023
1	0.867	0.017
10	0.908	0.015
100	0.907	0.015
1000	0.907	0.015

Cuadro 3: Kernel radial

En este Corpus como podemos observar los mejores resultados se obtienen con el Kernel Lineal llegando hasta una precisión de 1 si $c = 0.01$. Además podemos observar que un aumento en C contribuye a una mejora en los resultados obtenidos si utilizamos un Kernel radial. Lo mismo ocurre con el Kernel polinomial.

4. Corpus multi-clase: USPS

Se nos proporciona un corpus dividido en train y test. Este corpus está formado por 7291 muestras de aprendizaje y 2007 muestras de test. Analizaremos la combinación de C y Kernel con el fin de obtener el mínimo error.

c	precisión	intervalo(95 %)
0.01	0.931	0.011
0.1	0.928	0.011
1	0.926	0.011
10	0.925	0.011
100	0.925	0.011
1000	0.925	0.011

Cuadro 4: Kernel Lineal

c	precisión	intervalo(95 %)
0.01	0.556	0.021
0.1	0.854	0.015
1	0.928	0.011
10	0.938	0.010
100	0.941	0.010
1000	0.940	0.010

Cuadro 5: Kernel polinomial de grado 3

c	precisión	intervalo(95 %)
0.01	0.746	0.019
0.1	0.914	0.010
1	0.942	0.010
10	0.950	0.009
100	0.949	0.009
1000	0.949	0.009

Cuadro 6: Kernel radial

Como podemos observar la mejor precisión se obtiene utilizando un kernel radial con $C = 10$. Además, en el kernel polinomial podemos observar como un incremento de C contribuye a mejorar la precisión del modelo.

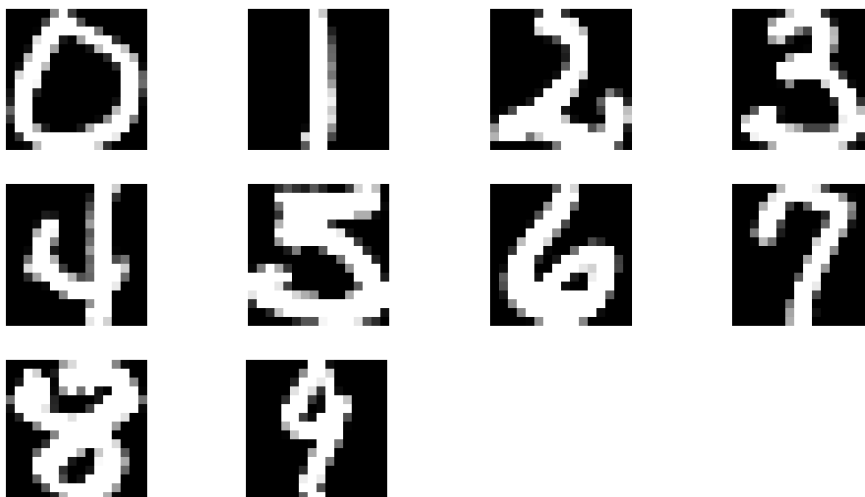


Figura 7: Ejemplo de muestras de la clase 0 a 9

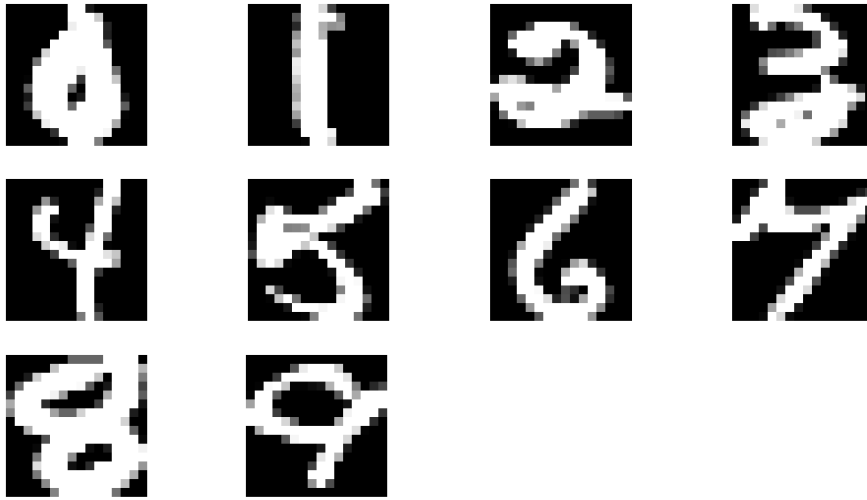


Figura 8: Ejemplo de muestras de la clase 0 a 9 que son vectores soporte

Las imágenes que son vectores soporte son aquellas que están en los márgenes y por tanto son imágenes extremas, que pueden confundirse con otras clases. Por ejemplo el 7 podría confundirse con un 4 o el 3 con un 2