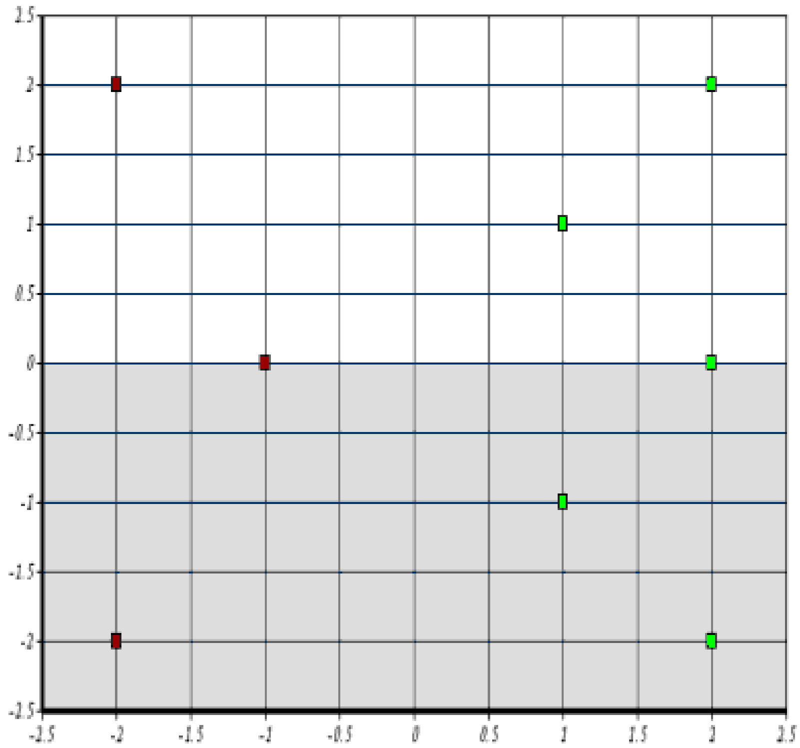
Ejercicioº 1. Queremos aprender un margen estricto lineal SVM de los puntos de la Figura, donde los puntos verdes tienen etiqueta positiva y los rojos etiquetas negativas. Encuéntrelo mediante un algoritmo de máquina de soporte lineal? ¿Qué se consigue mediante el entrenamiento mediante un clasificador de regresión logística con el conjunto de datos dado?



Solución: Para resolver el problema mediante el algoritmo de SVM tenemos que encontrar los valores de **w** y b. Hay que hallar dos hiperplanos con el margen más grande, el primero pasa a través de los puntos (1,-1) y (1,1), el segundo hiperplano pasa a través del punto (-1,0) De esta manera las ecuaciones son:

$(-1,0).\mathbf{w}^{\mathbf{T}} + b=-1$ $(1,-1).\mathbf{w}^{\mathbf{T}} + b=1$ $(1,1).\mathbf{w}^{\mathbf{T}} + b=1$

Que se transforman en

$-w_1+b=-1$ (1); $w_1-w_2+b=1$ (2)

$w_1+w_2+b=1$ (3)

Cuya solución es b=0; $w_1=1$y $w_2=0$ y la ecuación del hiperplano separador es

$g(\mathbf{x})=x_1$ y la frontera es $g(\mathbf{x})=0$; esto es $x_1=0$ (ver gráfica)

Ejercicio 2.- Dados dos patrones de entrenamiento en R^1 x_1=0,y_1=1 x_2=1,y_2=-1. Determinar el hiperplano separador ?

Ahora los datos son x_1= 1, x_2= 0 con y = $[+1,-1]^T$. Tenemos que encontrar los coeficientes w y b$\in R^1$, de forma tal que se verifique

$$\begin{array}{ll} \min_{\mathbf{w}} & \frac{1}{2}w^2 \\ s.a. & 1(w\times 1+\mathrm{b}) \geq 1 \quad (1) \\ & -1(\mathrm{w}\times 0+\mathrm{b}) \geq 1, \quad (2) \end{array} \qquad \begin{array}{ll} \min_{\mathbf{w}} & \frac{1}{2}w^2 \\ s.a. & 1(w\times 1+\mathrm{b}) \geq 1, \quad \mathrm{w} \geq 1\text{-b} \quad (1) \\ & -1(\mathrm{w}\times 0+\mathrm{b}) \geq 1, \quad \text{-b} \geq 1, \quad (2) \end{array}$$

De la ecuación (2) tenemos −b ≥ 1. y sustituyendo en (1), w ≥ 2. Así, para todo par (w, b) que satisfaga las ecuaciones (1) y (2), w ≥ 2. Como queremos minimizar $(1/2)w^2$ tenemos que el valor más pequeño es w=2, por lo que la solución óptima es (w, b) = (2,−1). El hiperplano separador es entonces g(x)=w×x+b=0, y sustituyendo tenemos 2x − 1 = 0, x=1/2 que es el punto medio entre x_1 y x_2

Ejercicio 3.- Dados los datos etiquetados como positivos definidos como vectores traspuestos $(2,5, 0)^T$, $(0, 2,5)^T$, $(-2,5,0)^T$, $(0, -2,5)^T$ y los datos etiquetados como negativos $(0,1)^T$, $(1,0)^T$, $(0,-1)^T$ y $(-1,0)^T$ y la función de transformación del espacio de características de entrada.

$$\Phi(x_1,x_2)^T = \begin{cases} (4-x_2+|3x_1-x_2|,\ 4-x_1+|3x_1-x_2|)^T & \text{si } \sqrt{x_1^2+x_2^2}>2 \\ (x_1,x_2)^T & \text{en otro caso} \end{cases}$$

a) Calcule los vectores soporte y la ecuación del hiperplano de separación.
b) Utilice el algoritmo SVM para clasificar el patrón de coordenadas $(2,2,1)^T$ ¿Qué conclusiones podemos sacar?

Solución.- La transformación del vector (2,5, 0) es el vector (11,5, 9), la del (0, 2,5) es el vector (4, 6,5), la del vector (-2,5, 0) se transforma en el vector (11,5, 14) y el vector (0, -2,5) en el vector (9, 6,5); mientras que los negativos quedan igual. De esta forma, los vectores soporte podrían ser el (0,1) y el (4, 6,5), puesto que la distancia euclídea entre ellos **es menor** que la distancia entre (1,0) y (4, 6,5) dado que

$$d((1,0),(4,\ 6,5)) = \sqrt{51,25};\ \text{mientras que d((0,1),(4, 6,5))=}\sqrt{46,25} \qquad .$$

De esta manera, con dos vectores soporte tenemos que trabajar en el espacio dual, dado que en el primal tendremos dos ecuaciones con tres incógnitas, a no ser que elijamos también como vector soporte el (1,0).

$$\begin{array}{l} H_0 : \mathbf{w.x} + w_0 = 1 \\ H_1 : \mathbf{w.x} + w_0 = -1 \end{array} \qquad \text{los vectores soporte son} \quad \begin{array}{ccc} x_1 & x_2 & Clase \\ 0 & 1 & -1 \\ 4 & 6,5 & +1 \end{array}$$

$$w_1x_1 + w_2x_2 + w_0 = -1$$
$$0w_1 + 1w_2 + w_0 = -1$$
$$w_2 = -1 - w_0$$

para el primer vector soporte y para el segundo

$$w_1x_1 + w_2x_2 + w_0 = 1$$
$$4w_1 + 6{,}5(-1 - w_0) + w_0 = 1$$
$$4w_1 = 5{,}5w_0 + 7{,}5$$

pero tenemos dos ecuaciones con tres incógnitas, así que tomamos como vector soporte negativo también el (1,0) y aunque no es la mejor solución, nos sirve para tener un hiperplano aproximado

$$w_1x_1 + w_2x_2 + w_0 = -1$$
$$1w_1 + 0w_2 + w_0 = -1$$
$$w_1 = -1 - w_0$$

y si ponemos las tres ecuaciones

$$w_1 = -1 - w_0$$
$$w_2 = -1 - w_0$$
$$4w_1 = 5{,}5w_0 + 7{,}5$$

de donde despejando tenemos w_0= -1,21, w_1= 0,21 y w_2= 0,21, y la ecuación del hiperplano es

$H:\ (0{,}21,\ 0{,}21)\cdot \mathbf{x}^T - 1{,}21 = 0{,}21x_1 + 0{,}21x_2 - 1{,}21 = 0;\ \mathrm{x}_2 = -\mathrm{x}_1 + 5{,}76$ **, esta solución no es la mejor pero si aproximada**

El vector (2,2,1) se transforma en el (6,6) por lo que si le aplicamos la ecuación del hiperplano tenemos que el signo es positivo, 1,31, y estaría en la clase positiva, por lo que el clasificador al menos para este patrón es adecuado; aunque esta en el margen.

Sol 2.- Si consideramos solo los vectores soporte (0,1) y (4, 6,5) en el espacio de características tenemos en el dual

$$\max \alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2(\Phi(s_1)^T.\Phi(s_1)) - \alpha_1\alpha_2(\Phi(s_1)^T.\Phi(s_2)) - \alpha_2\alpha_1(\Phi(s_2)^T.\Phi(s_1)) + \alpha_2^2(\Phi(s_2)^T.\Phi(s_2))\right)$$
$$s.a.\quad -\alpha_1 + \alpha_2 = 0,\ \alpha_1 \geq 0,\ \alpha_2 \geq 0$$

,

donde $\Phi(s_1) = (0{,}1)^{\mathrm{T}}$ y $\Phi(s_2) = (4,\ 6{,}5)^{\mathrm{T}}$, por lo que

$$\max\ \alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2 - 6{,}5\alpha_1\alpha_2 - 6{,}5\alpha_2\alpha_1 + 58{,}25\alpha_2^2\right)$$
$$s.a.\quad -\alpha_1 + \alpha_2 = 0,\ \alpha_1 \geq 0,\ \alpha_2 \geq 0$$

derivando $\mathrm{L}=\alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2 - 6{,}5\alpha_1\alpha_2 - 6{,}5\alpha_2\alpha_1 + 58{,}25\alpha_2^2\right) - \lambda(-\alpha_1 + \alpha_2)$

con respecto a $\alpha_1,\ \alpha_2$ y λ, tenemos

$$\begin{cases} 1-\alpha_1+6,5\alpha_2+\lambda=0 \\ 1+6,5\alpha_1-58,25\alpha_2-\lambda=0 \\ \alpha_1-\alpha_2=0 \end{cases}$$

$$\textit{de donde } \alpha_1=\frac{2}{46,25},\ \alpha_2=\frac{2}{46,25},\ \text{y } \lambda=-\frac{103,5-46,25}{46,25}$$

$$\hat{\mathbf{w}}=\sum_{i\in Sop}\hat{\alpha}_i y_i \Phi(s_i)=(2/46,25)(-1)\begin{pmatrix}0\\1\end{pmatrix}+(2/46,25)(1)\begin{pmatrix}4\\6,5\end{pmatrix}=\begin{pmatrix}8/46,25\\11/46,25\end{pmatrix}$$

$$\hat{\mathrm{w}}_0=-1-(\Phi(s_1)^T\hat{\mathbf{w}})=-1-(0,1)\begin{pmatrix}8/46,25\\11/46,25\end{pmatrix}=-57,25/46,25$$

y la ecuación del hiperplano separador es

$(8/46,25)\mathrm{x}_1+(11/46,25)\mathrm{x}_2-57,25/46,25=0$, o lo que es igual

$8x_1+11x_2-57,25=0,$ de donde $\mathrm{x}_2=-0,73\ \mathrm{x}_1+5,2$

El vector (2,2,1) se transforma en el (6,6) por lo que si le aplicamos la ecuación del hiperplano tenemos que el signo es positivo, 56,75, y estaría en la clase positiva, por lo que el clasificador al menos para este patrón es adecuado; aunque esta en el margen.

Ejercicio 4.- Dados los datos etiquetados como positivos definidos como vectores traspuestos $(3,0)^T$, $(0,3)^T$, $(-3,0)^T$, $(0,-3)^T$ y los datos etiquetados como negativos $(0,1)^T$, $(1,0)^T$, $(0,-1)^T$ y $(-1,0)^T$ y la función de transformación del espacio de características de entrada.

$$\Phi(x_1,x_2)^T=\begin{cases}(4-x_2+|3x_1-x_2|,\ 4-x_1+|3x_1-x_2|)^T & \text{si } \sqrt{x_1^2+x_2^2}>2\\ (x_1,x_2)^T & \text{en otro caso}\end{cases}$$

a) Calcule los vectores soporte y la ecuación del hiperplano de separación.

b) Utilice el algoritmo SVM para clasificar el patrón de coordenadas $(1,1)^T$ con etiqueta negativa ¿Qué conclusiones podemos sacar?

Sol 1.- La transformación del vector (3,0) es el vector (13,10), la del (0,3) es el vector (4,7), la del vector (-3,0) se transforma en el vector (13, 16) y el vector (0,-3) en el vector (10,7); mientras que los negativos quedan igual. De esta forma, los vectores soporte podrían ser el (1,0) y el (4,7), puesto que la distancia euclídea entre ellos **es menor**, 7,61. De esta manera, con estos vectores soporte tenemos.

$$H_0:\mathbf{w.x}+w_0=1$$
$$H_1:\mathbf{w.x}+w_0=-1$$

x_1	x_2	*Clase*
1	0	−1
4	7	+1

$$w_1x_1 + w_2x_2 + w_0 = -1$$
$$1w_1 + 0w_2 + w_0 = -1$$
$$w_1 = -1 - w_0$$

o también

$$w_1x_1 + w_2x_2 + w_0 = 1$$
$$4(-1 - w_0) + 7w_2 + w_0 = 1$$
$$-4 - 4w_0 + 7w_2 + w_0 = 1$$
$$7w_2 = 3w_0 + 5$$

pero tenemos dos ecuaciones con tres incógnitas, así que tomamos como vector soporte negativo también el (0,1) y aunque no es la mejor solución, nos sirve para tener un hiperplano aproximado

$$w_1x_1 + w_2x_2 + w_0 = -1$$
$$0w_1 + 1w_2 + w_0 = -1$$
$$w_2 = -1 - w_0$$

y si ponemos las tres ecuaciones

$$w_1 = -1 - w_0$$
$$w_2 = -1 - w_0$$
$$7w_2 = 3w_0 + 5$$

de donde despejando tenemos w_0=-1,2, w_1=0,2 y w_2=0,2, y la ecuación del hiperplano es

$H: \ (2,2)\cdot \mathbf{x}^T - 3 = 0,2x_1 + 0,2x_2 - 1,2 = 0;\ x_2 = -x_1 + 6$, esta solución no es la mejor pero si aproximada

El vector (1,1) con etiqueta negativa se transforma en el (1,1) por lo que si le aplicamos la ecuación del hiperplano tenemos que el signo es negativo y estaría en la clase negativa, por lo que el clasificador al menos para este patrón es adecuado

Sol 2.- Si consideramos solo los vectores soporte (1,0) y (4,7) tenemos en el dual

$$\max \alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2(\Phi(s_1)^T.\Phi(s_1)) - \alpha_1\alpha_2(\Phi(s_1)^T.\Phi(s_2)) - \alpha_2\alpha_1(\Phi(s_2)^T.\Phi(s_1)) + \alpha_2^2(\Phi(s_2)^T.\Phi(s_2))\right)$$
$$s.a.\quad -\alpha_1 + \alpha_2 = 0,\ \alpha_1 \geq 0,\ \alpha_2 \geq 0$$

,

donde $\Phi(s_1) = (1,0)^{\mathrm{T}}$ y $\Phi(s_2) = (4,7)^{\mathrm{T}}$, por lo que

$$\max\ \alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2 - 4\alpha_1\alpha_2 - 4\alpha_2\alpha_1 + 65\alpha_2^2\right)$$
$$s.a.\quad -\alpha_1 + \alpha_2 = 0,\ \alpha_1 \geq 0,\ \alpha_2 \geq 0$$

derivando con respecto a α_1, α_2 y λ, tenemos

$$\begin{cases} 1 - \alpha_1 + 4\alpha_2 + \lambda = 0 \\ 1 + 4\alpha_1 - 65\alpha_2 - \lambda = 0 \\ \alpha_1 - \alpha_2 = 0 \end{cases}$$

$$\textit{de donde } \alpha_1 = \alpha_2 = 2/58, \text{ y } \lambda = -\frac{64}{58} = -\frac{32}{29}$$

$$\hat{\mathbf{w}} = \sum_{i \in Sop} \hat{\alpha}_i y_i \Phi(s_i) = (2/58)(-1)\begin{pmatrix}1\\0\end{pmatrix} + (2/58)(1)\begin{pmatrix}4\\7\end{pmatrix} = \begin{pmatrix}3/29\\7/29\end{pmatrix}$$

$$\hat{w}_0 = -1 - (\Phi(s_1)^T \hat{\mathbf{w}}) = -1 - (1,0)\begin{pmatrix} 3/29 \\ 7/29 \end{pmatrix} = -32/29$$

y la ecuación del hiperplano separador es

$3x_1 + 7x_2 - 32 = 0$, o lo que es igual

$$x_2 = -(3/7)x_1 + (32/7)$$

el vector de coordenadas (1,1) tiene valor asociado 3+7-32= -22, negativo luego pertenece a la clase negativa y por tanto el clasificador, al menos para este patrón, es adecuado.