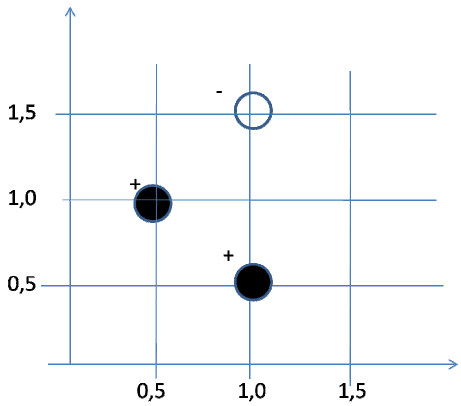
Ejercicio 1.- (2 puntos) Considere los tres vectores bidimensionales linealmente separables de la siguiente figura. Encuentre el SVM lineal que separa de manera óptima las clases al maximizar el margen.



Solución

Todos los puntos son vectores soporte el hiperplano de margen H^+ es la línea que pasa por los dos puntos positivos. El hiperplano de margen H^- es la recta que pasa por el punto negativo y es paralela a H^+. La función de decisión es la recta que está entre H^+ y H^-. Esta recta tiene por ecuación $x_2 = -x_1+2$

Ejercicio 2.- (2 puntos) 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 $(0,2)^T$, $(1,0)^T$, $(0,-1)^T$ y $(-1,0)^T$ y la función de transformación del espacio de características de entrada.

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

a) **(2 puntos)** Calcule los vectores soporte y la ecuación del hiperplano de separación.

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 todos quedan igual menos el (0,2) que se transforma en el (4,6) . De esta forma, los vectores soporte podrían ser el (4,6) y el (4,7). 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*
4	6	-1
4	7	$+1$

$$w_1 x_1 + w_2 x_2 + w_0 = -1 \qquad\qquad w_1 x_1 + w_2 x_2 + w_0 = 1$$
$$4w_1 + 6w_2 + w_0 = -1 \qquad\qquad 4w_1 + 7w_2 + w_0 = 1$$

de esta dos ecuaciones obtenemos que $w_2=2$ y que $4w_1+w_0=-13$, pero tenemos dos ecuaciones con tres incógnitas, así que tomamos como vector soporte positivo también el (10,7). Si ponemos las tres ecuaciones tenemos

$$w_2 = 2$$
$$4w_1 + w_0 = -13$$
$$10w_1 + 7w_2 + w_0 = 1$$

de donde despejando tenemos $w_0= -13$, $w_1=0$ y $w_2=2$, y la ecuación del hiperplano es

$H: 2x_2 - 13 = 0;\ x_2 = 13/2 = 6.5$. Esta solución no es la mejor pero si aproximada

Sol 2.- Si consideramos solo los vectores soporte (4,6) como negativo y (4,7) como positivo, 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 $\Phi(s_1) = (4,6)^{\mathrm{T}}$ y $\Phi(s_2) = (4,7)^{\mathrm{T}}$, por lo que

$$\max\ \alpha_1 + \alpha_2 - \frac{1}{2}\left(52\alpha_1^2 - 58\alpha_1\alpha_2 - 58\alpha_2\alpha_1 + 65\alpha_2^2\right)$$

$$s.a.\quad -\alpha_1 + \alpha_2 = 0,\ \alpha_1 \geq 0,\ \alpha_2 \geq 0$$

derivando con respecto a α_1, α_2 y λ, tenemos

$$\begin{cases} 1 - 52\alpha_1 + 58\alpha_2 + \lambda = 0 \\ 1 + 58\alpha_1 - 65\alpha_2 - \lambda = 0 \\ \qquad \alpha_1 - \alpha_2 = 0 \end{cases}$$

de donde $\alpha_1 = \alpha_2 = 2$ y $\lambda = -13$

$$\hat{\mathbf{w}} = \sum_{i \in Sop} \hat{\alpha}_i y_i \Phi(s_i) = (2)(-1)\begin{pmatrix}4\\6\end{pmatrix} + (2)(1)\begin{pmatrix}4\\7\end{pmatrix} = \begin{pmatrix}0\\2\end{pmatrix}$$

$$\hat{\mathrm{w}}_0 = -1 - (\Phi(s_1)^T\hat{\mathbf{w}}) = -1 - (4,6)\begin{pmatrix}0\\2\end{pmatrix} = -1 - 12 = -13$$

y la ecuación del hiperplano separador es

$0x_1 + 2x_2 - 13 = 0$, o lo que es igual

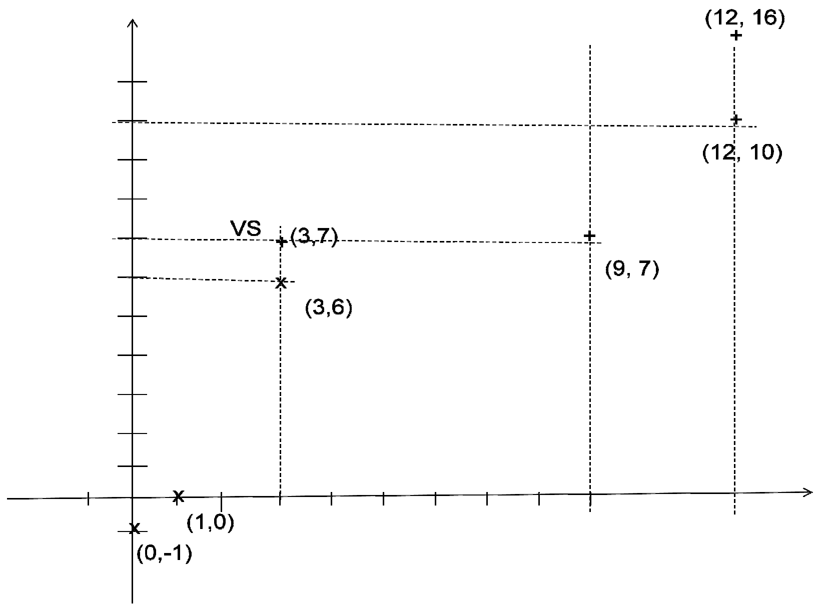
$x_2 = 13/2 = 6.5$

Ejercicio 1.- (3puntos) 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 $(0,2)^T$, $(1,0)^T$, $(0,-1)^T$ y $(-2,0)^T$ y la función de transformación del espacio de características de entrada.

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

a) **(2,5 puntos)** Calcule los vectores soporte y la ecuación del hiperplano de separación.
b) **(0,5 puntos)** Utilice el algoritmo SVM para clasificar el patrón de coordenadas $(-1,0)^T$ ¿Qué conclusiones podemos sacar?

Sol 1.- La transformación del vector (3,0) es el vector (12,10), la del (0,3) es el vector (3,7), la del vector (-3,0) se transforma en el vector (12,16) y el vector (0,-3) en el vector (9,7); mientras que los negativos todos quedan igual menos el (0,2) que se transforma en el (3,6) y el (-2,0) que se transforma en el (9,12). De esta forma, los vectores soporte podrían ser el (3,6) y el (3,7). De esta manera, con estos vectores soporte tenemos.



x_1	x_2	$Clase$
3	6	-1
3	7	$+1$

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

$$w_1x_1 + w_2x_2 + w_0 = -1 \qquad w_1x_1 + w_2x_2 + w_0 = 1$$
$$3w_1 + 6w_2 + w_0 = -1 \qquad 3w_1 + 7w_2 + w_0 = 1$$

de esta dos ecuaciones obtenemos que $w_2=2$ y que $3w_1+w_0= -13$, pero tenemos dos ecuaciones con tres incógnitas, así que tomamos como vector soporte positivo también el (9,7). Si ponemos las tres ecuaciones tenemos

$$\begin{aligned} w_2 &= 2 \\ 3w_1 + w_0 &= -13 \\ 9w_1 + 7w_2 + w_0 &= 1 \end{aligned}$$

de donde despejando tenemos $w_0= -13$, $w_1=0$ y $w_2=2$, y la ecuación del hiperplano es

$H : 2x_2 - 13 = 0;\ x_2 = 13/2=6.5$. Esta solución hace que el patrón (-2,0) de la clase negativa no esté bien clasificado en el conjunto de entrenamiento.

Sol 2.- Si consideramos solo los vectores soporte (3,6) como negativo y (3,7) como positivo, tenemos en el dual

$$\max \alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2(\Phi(s_1)^T.\Phi(s_1)) - \alpha_1\alpha_2(\Phi(s_1)^T.\Phi(s_2)) - \alpha_2\alpha_1(\Phi(s_2)^T.\Phi(s_1)) + \alpha_2^2(\Phi(s_2)^T.\Phi(s_2))\right)$$

$$s.a. \quad -\alpha_1 + \alpha_2 = 0,\ \alpha_1 \ge 0,\ \alpha_2 \ge 0$$

,

donde $\Phi(s_1) = (3,6)^T$ y $\Phi(s_2) = (3,7)^T$, por lo que

$$\max\ \alpha_1 + \alpha_2 - \frac{1}{2}\left(45\alpha_1^2 - 51\alpha_1\alpha_2 - 51\alpha_2\alpha_1 + 58\alpha_2^2\right)$$

$$s.a. \quad -\alpha_1 + \alpha_2 = 0,\ \alpha_1 \ge 0,\ \alpha_2 \ge 0$$

derivando con respecto a α_1, α_2 y λ, tenemos

$$\begin{cases} 1-45\alpha_1+51\alpha_2+\lambda=0 \\ 1+51\alpha_1-58\alpha_2-\lambda=0 \\ \qquad \alpha_1-\alpha_2=0 \end{cases}$$

de donde $\alpha_1=\alpha_2=2$ y $\lambda=-13$

$$\hat{\mathbf{w}}=\sum_{i\in Sop}\hat{\alpha}_i y_i\Phi(s_i)=(2)(-1)\begin{pmatrix}3\\6\end{pmatrix}+(2)(1)\begin{pmatrix}3\\7\end{pmatrix}=\begin{pmatrix}0\\2\end{pmatrix}$$

$$\hat{\mathrm{w}}_0=-1-(\Phi(s_1)^T\hat{\mathbf{w}})=-1-(3,6)\begin{pmatrix}0\\2\end{pmatrix}=-1-12=-13$$

y la ecuación del hiperplano separador es

$0x_1+2x_2-13=0$, o lo que es igual

$x_2=13/2=6.5$

c) El patrón de coordenadas $(-1,0)^T$ a ser el valor de x_2 menor de 6.5 se clasifica en la clase negativa luego se clasifica bien

Ejercicio 1.- (3 puntos) 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) **(2 puntos)** Calcule los vectores soporte y la ecuación del hiperplano de separación.
b) **(1 punto)** Utilice el algoritmo SVM para clasificar el patrón de coordenadas $(1,1)^T$ ¿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 euclidea entre ellos es mayor 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	$+1$

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

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

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

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

El vector (1,1) 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