Boletín 5. Ejercicio 1

Luis Ardévol Mesa

Observación	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
x_1	2	4	4	4	6	7	8	9	2	6	7	8	9	10	10	12
x_2	6	3	4	6	3	7	4	8	1	2	4	8	1	3	6	4
y	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1
α_i	0	1	0.3333 $\to$ 1/3	0	1	0.1667 $\to$ 1/6	1	1	1	0.5 $\to$ 1/2	1	1	0	0	1	0

- Vectores de soporte $\longrightarrow \alpha_i > 0$
  - En el límite del margen $\longrightarrow \xi_i = 0$, $0 < \alpha_i < C$
  - Resto $\longrightarrow \xi_i > 0$, $\alpha_i = C$

  $\Big\}$ $\underline{C = 1}$

1.) Vectores de soporte $\longrightarrow$ $\boxed{2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 15}$

En el límite del margen $\longrightarrow$ $\boxed{3, 6, 10}$

2.) Para obtener β_0, necesitamos primero $\beta = \sum_{i=1}^{N} \alpha_i y_i x_i$

$$(x_1): \beta_1 = \cancel{4} + \frac{4}{3} + \cancel{6} + \frac{7}{6} + \cancel{8} + \cancel{9} + \cancel{(-2)} + (-3) + \cancel{(-7)} + \cancel{(-8)} + \cancel{(-10)} = \frac{4}{3} + \frac{7}{6} - 3 = -0.5$$

$$(x_2): \beta_2 = \cancel{3} + \frac{4}{3} + \cancel{3} + \frac{7}{6} + \cancel{4} + \cancel{8} + (-1) + (-1) + \cancel{(-4)} + \cancel{(-8)} + \cancel{(-6)} = \frac{4}{3} + \frac{7}{6} - 2 = 0.5$$

$$\boxed{\beta = \begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix}}$$

Obtenemos β_0 usando los vectores en el límite $(\xi_i = 0) \longrightarrow \alpha_i [y_i (x_i^T \beta + \beta_0) - (1 - \cancel{\xi_i})] = 0 \Rightarrow \beta_0 = \frac{1}{y_i} - x_i^T \beta$

$$\text{Observación 3:} \quad \beta_0^{(3)} = 1 - (4 \ \ 4)\begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix} = 1$$

$$\text{Observación 6:} \quad \beta_0^{(6)} = 1 - (7 \ \ 7)\begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix} = 1$$

$$\text{Observación 10:} \quad \beta_0^{(10)} = -1 - (6 \ \ 2)\begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix} = 1$$

$$\boxed{\beta_0 = \frac{1}{3}\left(\beta_0^{(3)} + \beta_0^{(6)} + \beta_0^{(10)}\right) = 1}$$

El valor de M lo obtenemos como $\boxed{M = \|\beta\|^{-1} = \left[\left(1/4 + 1/4\right)\right]^{-1/2} = \sqrt{2} \simeq 1.4142}$

3.) Para las observaciones 1, 4, 13, 14, 16, $\xi_i = 0$ ya que $\alpha_i = 0$; veamos por qué

$$\left.\begin{array}{l} \left.\begin{array}{l} \alpha_i = 0 \\ \alpha_i = C - \mu_i \end{array}\right\} \mu_i = C = 1 \neq 0 \\ \qquad\qquad\quad \uparrow \\ \qquad\qquad\quad \text{dato} \\ \mu_i \xi_i = 0 \end{array}\right\} \Rightarrow \xi_i = 0 \quad \forall i \ / \ \underline{\alpha_i = 0}$$

Por tanto, $\xi_1=\xi_4=\xi_{13}=\xi_{14}=\xi_{16}=0$. Además, también sabemos que para los vectores de soporte en el margen $\xi_i=0$, por tanto $\xi_3=\xi_6=\xi_{10}=0$. Para el resto de observaciones, que son vectores de soporte.

$$\alpha_i\left[y_i(x_i^T\beta+\beta_0)-(1-\xi_i)\right]=0 \overset{\alpha_i\neq 0}{\Rightarrow} \xi_i = 1-y_i(x_i^T\beta+\beta_0)$$

$$\xi_2 = 1-\left[\begin{pmatrix}4 & 3\end{pmatrix}\begin{pmatrix}-0.5\\0.5\end{pmatrix}+1\right] = 1-\left[-2+\frac{3}{2}+1\right]=\frac{1}{2} < 1$$

$$\xi_5 = 1-\left[\begin{pmatrix}6 & 3\end{pmatrix}\begin{pmatrix}-0.5\\0.5\end{pmatrix}+1\right] = 1-\left[-3+3/2+1\right]=\frac{3}{2} > 1$$

$$\xi_7 = 1-\left[\begin{pmatrix}8 & 4\end{pmatrix}\begin{pmatrix}-0.5\\0.5\end{pmatrix}+1\right] = 1-\left[-4+2+1\right]=2 > 1$$

$$\xi_8 = 1-\left[\begin{pmatrix}9 & 8\end{pmatrix}\begin{pmatrix}-0.5\\0.5\end{pmatrix}+1\right] = 1-\left[-\frac{9}{2}+4+1\right]=\frac{1}{2} < 1$$

$$\xi_9 = 1+\left[\begin{pmatrix}2 & 1\end{pmatrix}\begin{pmatrix}-0.5\\0.5\end{pmatrix}+1\right] = 1+\left[-1+\frac{1}{2}+1\right]=\frac{3}{2} > 1$$

$$\xi_{11} = 1+\left[\begin{pmatrix}7 & 4\end{pmatrix}\begin{pmatrix}-0.5\\0.5\end{pmatrix}+1\right] = 1+\left[-\frac{7}{2}+2+1\right]=\frac{1}{2} < 1$$

$$\xi_{12} = 1+\left[\begin{pmatrix}8 & 8\end{pmatrix}\begin{pmatrix}-0.5\\0.5\end{pmatrix}+1\right] = 1+1=2 > 1$$

$$\xi_{15} = 1+\left[\begin{pmatrix}10 & 6\end{pmatrix}\begin{pmatrix}-0.5\\0.5\end{pmatrix}+1\right] = 1+\left[-5+3+1\right]=0 < 1$$

Hagamos una tabla con las observaciones y sus respectivas ξ_i. Indicaremos las observaciones mal clasificadas ($\xi_i>1$) en rojo.

Observación	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
ξ_i	0	0.5	0	0	1.5	0	2	0.5	1.5	0	0.5	2	0	0	0	0