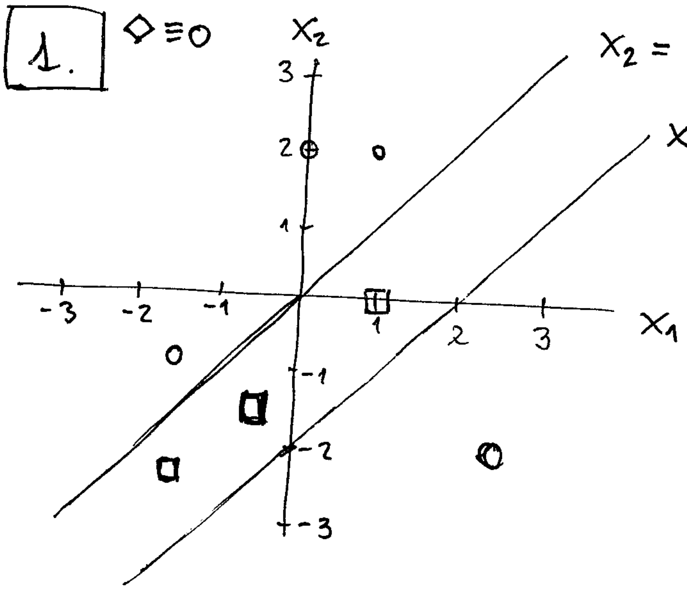
Alejandro Santorum Varela

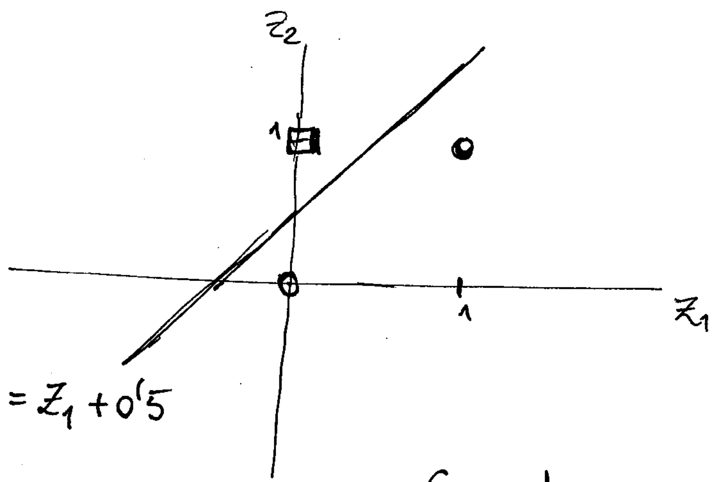
Parcial 4

$\boxed{1.}$ $\diamond \equiv 0$



$x_2 = x_1 \longrightarrow z_1 = \theta(x_2 - x_1)$

$x_2 = x_1 - 2 \longrightarrow z_2 = \theta(x_2 - x_1 + 2)$



$z_2 = z_1 + 0'5$

$\downarrow$

$$\boxed{y = \theta(z_2 - z_1 - 0'5)}$$

Comprobamos que la predicción de clase es la correcta

$z_1 = 1, z_2 = 1 \Rightarrow y = \theta(-0'5) = 0$ ✓

$z_1 = 0, z_2 = 0 \Rightarrow y = \theta(-0'5) = 0$ ✓

$z_1 = 0, z_2 = 1 \Rightarrow y = \theta(0'5) = 1$ ✓

Por tanto:

$V_{01} = 0$

$V_{02} = 2$

$V_{11} = -1$

$V_{12} = -1$

$V_{21} = 1$

$V_{22} = 1$

$W_0 = -0'5$

$W_1 = -1$

$W_2 = 1$

Hemos sacado los coefs. de las fórmulas de z_1, z_2 e y:

$$z_1 = \theta(x_2 - x_1)$$

$$z_2 = \theta(x_2 - x_1 + 2)$$

$$y = \theta(z_2 - z_1 - 0'5)$$

2.1

$$Z_1 = \theta\left(0'5 - 0'5x_1 - \frac{\sqrt{3}}{2}x_2\right)$$

$x_2 = 0 \Rightarrow x_1 = 1$

$x_1 = 0 \Rightarrow x_2 = \dfrac{-0'5}{-\sqrt{3}/2} \approx 0'577$

$\Rightarrow$ pendiente : $m \approx -0'577$

$\Rightarrow \boxed{x_2 = -0'577x_1 + 0'577}$ r_1

$$Z_2 = \theta(1 + x_1)$$

x_2 libre

$x_1 = -1$

$\Rightarrow \boxed{x_1 = -1}$ r_2

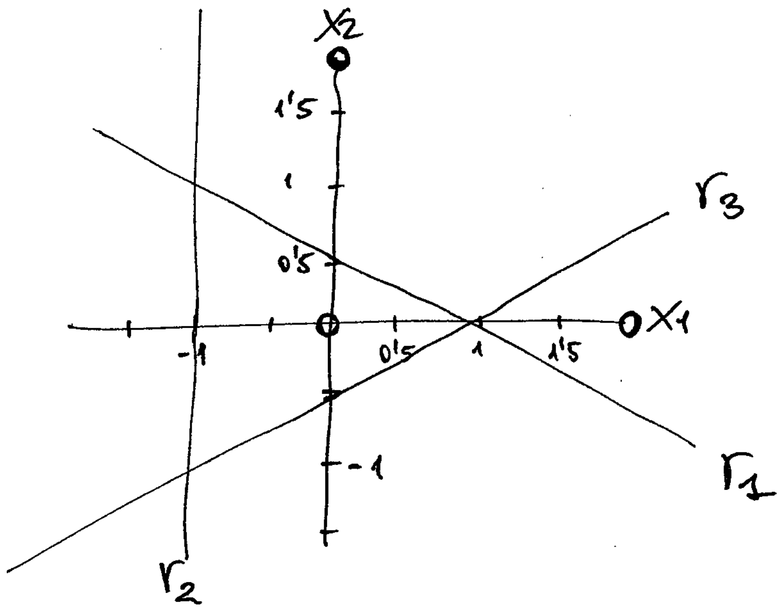
$$Z_3 = \theta\left(0'5 - 0'5x_1 + \frac{\sqrt{3}}{2}x_2\right)$$

$x_2 = 0 \Rightarrow x_1 = 1$

$x_1 = 0 \Rightarrow x_2 = \dfrac{-0'5}{\sqrt{3}/2} \approx -0'577$

$\Rightarrow$ pendiente : $m \approx 0'577$

$\Rightarrow \boxed{x_2 = 0'577x_1 + 0'577}$ r_3



punto $(0,0)$: $\begin{cases} Z_1 = 1 \\ Z_2 = 1 \\ Z_3 = 1 \end{cases} \Rightarrow y = \theta(0'5) = 1$ (clase 1)

pto. $(0,2)$: $\begin{cases} Z_1 = 0 \\ Z_2 = 1 \\ Z_3 = 1 \end{cases} \Rightarrow y = \theta(1'5) = 1$ (clase 1)

pto. $(2,0)$: $\begin{cases} Z_1 = 0 \\ Z_2 = 1 \\ Z_3 = 0 \end{cases} \Rightarrow y = \theta(0'5) = 1$ (clase 1)

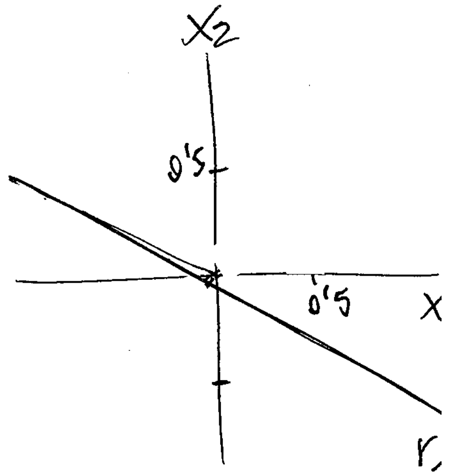
Si la función fuese la sigmoidal para z_1, z_2 cambiaría la frontera de decisión. De forma aproximada, la nueva frontera de decisión sería:

caso $z_1 = \sigma\left(0'5 - 0'5x_1 - \frac{\sqrt{3}}{2}x_2\right)$

buscamos (aprox.): $0'5 - 0'5x_1 - \frac{\sqrt{3}}{2}x_2 = 0'5$

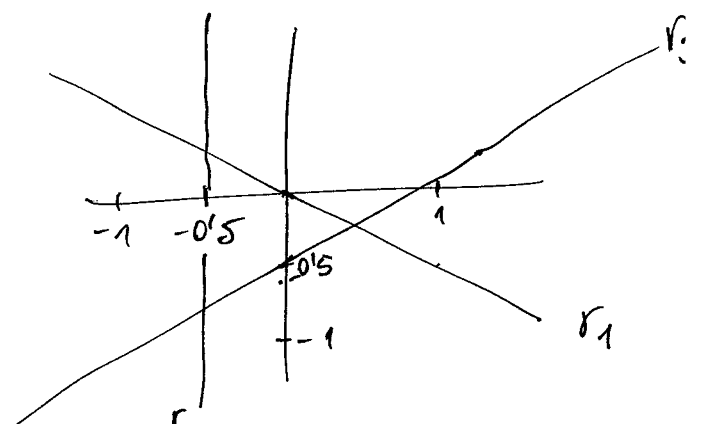
$$\boxed{\frac{\sqrt{3}}{2}x_2 = -0'5x_1}$$

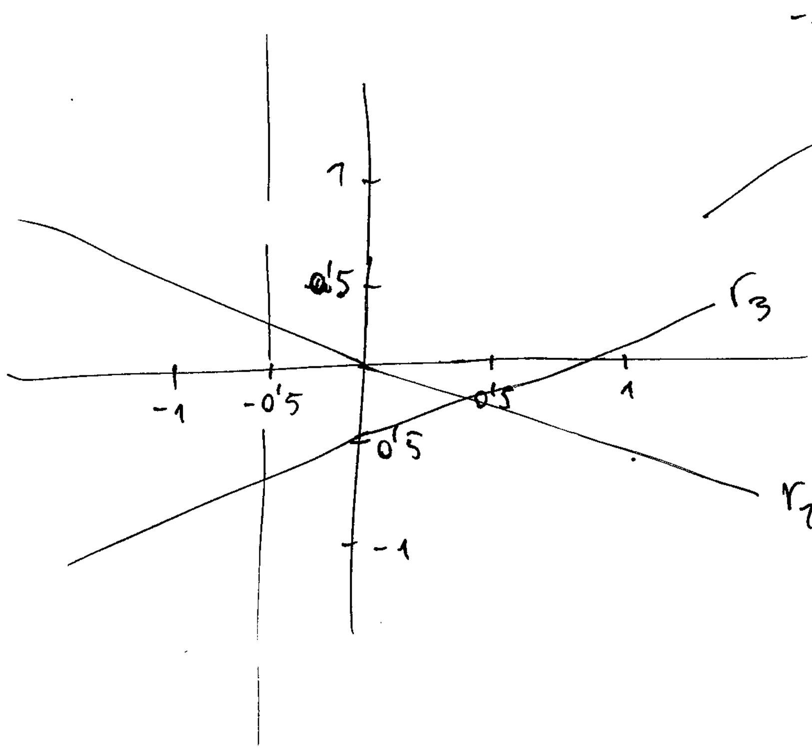
$$\Rightarrow \boxed{x_2 \approx -0'577\, x_1} \quad r_1$$



caso $z_2 = \sigma(1+x_1)$

buscamos (aprox.): $1 + x_1 = 0'5 \Rightarrow \boxed{x_1 = -0'5}$





$\boxed{3.}$ Y entrada
X oculta
1 salida

$\left.\begin{array}{l} Y \text{ entrada} \\ X \text{ oculta} \\ 1 \text{ salida} \end{array}\right\}$ tarda t

Coste clasificación NN: (teoría) : $O(XY) \Rightarrow$

$\Rightarrow O(XY) = t \Rightarrow O(\frac{X}{2}Y) = \frac{1}{2}O(XY) = \frac{1}{2}t$

$\Rightarrow$ tardaría aproximadamente la mitad del tiem

ALEJANDRO SANTORUM VARELA PARCIAL 4

4. Lista $[1, -2, 8, -3, 10, -5]$

Ind.	Números	Fitness	Esquema
100110	$1, -3, 10$	$1/8 = 0'125$	H_1
110100	$1, -2, -3$	$1/4 = 0'25$	H_1, H_2
010110	$-2, -3, 10$	$1/5 = 0'2$	
011001	$-2, 8, -5$	$1 = 1$	H_2

Fitness medio población: $\frac{0'125 + 0'25 + 0'2 + 1}{4} = 0'394 := \bar{f}(t)$

Fitness medio H_1 : $\bar{f}_{H_1}(t) = \frac{0'125 + 0'25}{2} = \boxed{0'188}$

Fitness medio H_2 : $\bar{f}_{H_2}(t) = \frac{0'25 + 1}{2} = \boxed{0'625}$

$$\mathbb{E}_s[n_{H_1}(t)] = n_{H_1}(t) \cdot \frac{\bar{f}_{H_1}(t)}{\bar{f}(t)} = 2 \cdot \frac{0'188}{0'394} = \boxed{0'954}$$

$$\mathbb{E}_s[n_{H_2}(t)] = n_{H_2}(t) \cdot \frac{\bar{f}_{H_2}(t)}{\bar{f}(t)} = 2 \cdot \frac{0'625}{0'394} = \boxed{3'173}$$

4.2 Los pesos de una red neuronal son V y W

V es una matriz $(D+1) \times J$, donde D es el número de atributos y J el número de neuronas de la capa oculta

W es un vector $(J+1) \times 1$.

En total tenemos $\boxed{(D+1)J + (J+1) \text{ pesos}}$ (lo denotaremos por n-pesos)

Los pesos de una red neuronal pueden ser cualquier número real, por lo tanto podemos representar un individuo (cadena de pesos) como un $\boxed{\text{vector real n\_pesos-dimens}}$ es decir, una lista de números reales de tamaño n_pesos

La inicialización puede ser una muestra de tamaño n_pesos de una distribución uniforme [-0'5, 0'5], así sería idéntica a la inicialización del algoritmo vis en teoría.

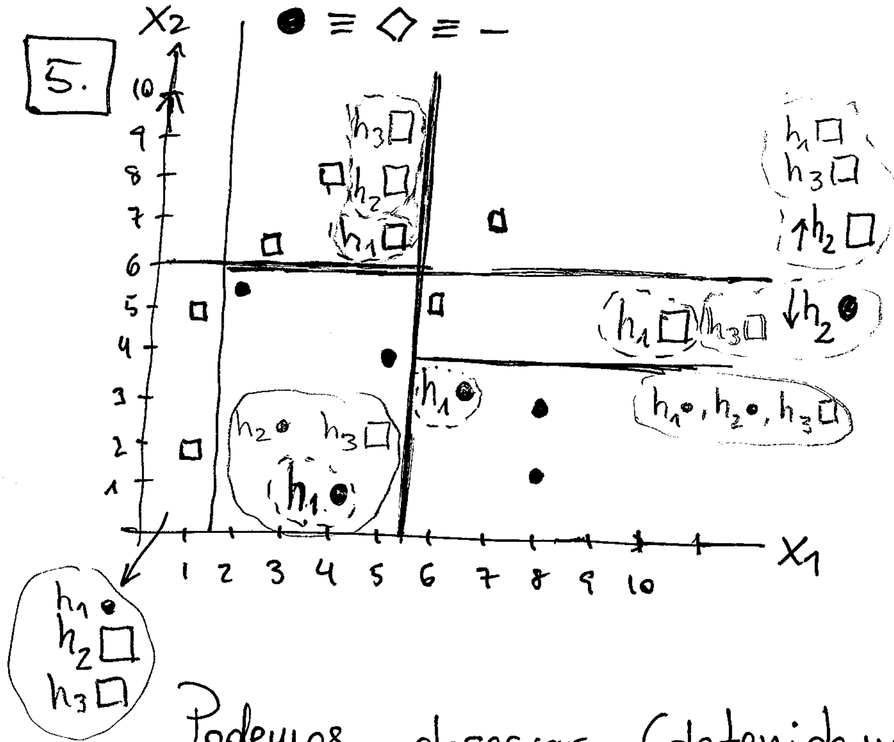
El fitness sería en función del error final en la neurona de salida, es decir: $\hat{y}$:= valor predicho red neuronal
y := valor real

Entonces $$f(v, w \mid \vec{x}) = \frac{1}{|\hat{y} - y| + 1}$$

Hemos escogido esta fórmula para maximizar el fitness a la vez que minimizamos el error. Añadimos un '+1' en el denominador para evitar división por cero.

Alejandro Santorum Varela Parcial 4

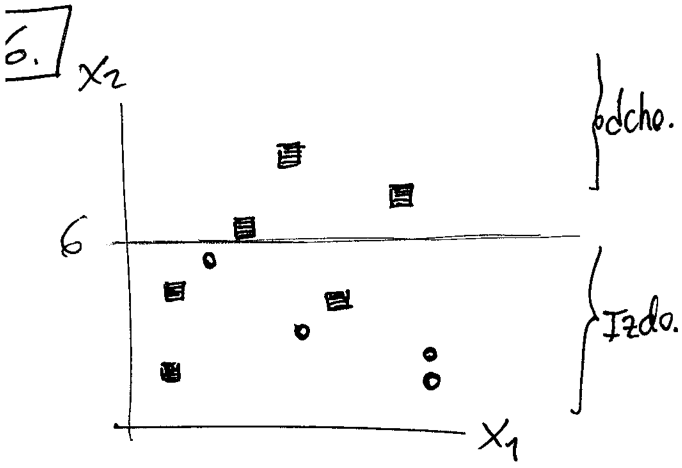
5.



No dibujamos las fronteras de h_3 ya que siempre predice positivo.

Podemos observar (detenidamente) que todos los ejemplos están bien clasificados gracias al voto por mayoría. $\Longrightarrow$ Número de errores = cero.

6.



$P_L = \frac{7}{10} = 0'7$

$i(t_L) = 2 \cdot \frac{4}{7} \cdot \frac{3}{7} = 0'49$

$P_R = \frac{3}{10} = 0'3$

$i(t_R) = 2 \cdot 0 \cdot 1 = 0$

$i(t) = 2 \cdot \frac{4}{10} \cdot \frac{6}{10} = 0'48$

$\Delta i(s,t) = 0'48 - 0'7 \cdot 0'49 - 0'3 \cdot 0 = \boxed{0'137}$