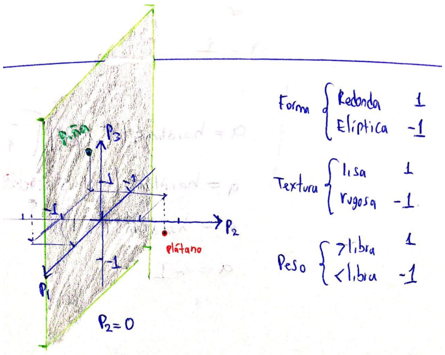
Introducción a las Redes Neurales

Bárcenas Martínez Erick Iván

1. Suponga que queremos distinguir entre plátanos y piñas:

$$P_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \text{ (plátano)}, \qquad P_2 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \text{ (piña)}$$

- **Diseñe un perceptrón para reconocer dichos patrones.**
    1. Graficar los puntos (en 3D) y se propone un límite de decisión (el plano p2=0).



    2. Se ponderan las características y se obtiene la matriz de pesos:
       W = [0 1 0]
    3. El sesgo (bias) es igual a cero, es decir, b=0 debido a que pasa por el origen.
       Se escribe nuestro perceptrón.

$$a = hardlims\left(\begin{bmatrix} w_{1,1} & w_{1,2} & w_{1,3} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} + b \right).$$

Por lo tanto, si es plátano a=1, y si es piña a=-1

$$\boxed{a = hardlims\left(\begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} \right)}$$

- **Diseñe una red de Hamming para reconocer dichos patrones.**

Feedforward Layer

$$W^1 = \begin{bmatrix} p_1^T \\ p_2^T \end{bmatrix} = \begin{bmatrix} -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \qquad b^1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$a^1 = W^1 p + b^1 = \begin{bmatrix} p_1^T \\ p_2^T \end{bmatrix} p + \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} p_1^T p + 3 \\ p_1^T p + 3 \end{bmatrix}$$

Recurrent Layer

$$a^2(0) = a^1 \quad \text{(condición inicial)}$$

$$a^2(t+1) = \text{poslin}(W^2 a^2(t)$$

$$W^2 = \begin{bmatrix} 1 & -\varepsilon \\ -\varepsilon & 1 \end{bmatrix} \qquad \varepsilon = \frac{1}{S-1} \qquad S: \#\text{neuronas}$$

$$a^2(t+1) = \text{poslin}\left(\begin{bmatrix} 1 & -\varepsilon \\ -\varepsilon & 1 \end{bmatrix} a^2(t)\right) = \text{poslin}\left(\begin{bmatrix} a_1^2(t) - \varepsilon a_2^2(t) \\ a_2^2(t) - \varepsilon a_1^2(t) \end{bmatrix}\right)$$

- **Pruebe la operación de las redes con tres distintos vectores de entrada.**

Para el perceptrón

$$a = \text{hardlims}\left(\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}\right)$$

Plátano

$$p_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$a = \text{hardlims}\left(\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}\right)$$

$$a = \text{hardlims}(0+1+0)$$

$$a = \text{hardlims}(1)$$

$a = 1$ (plátano) ✓

Piña

$$p_2 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

$$a = \text{hardlims}\left(\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\right)$$

$$a = \text{hardlims}(0-1+0)$$

$$a = \text{hardlims}(-1)$$

$a = -1$ (piña) ✓

Piña redonda

$$p_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$a = \text{hardlims}\left(\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right)$$

$$a = \text{hardlims}(0-1+0)$$

$$a = \text{hardlims}(-1)$$

$a = -1$ (piña) ✓

Para red de Hamming

plátano

$$p_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$a^1 = \begin{bmatrix} -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$a^1 = \begin{bmatrix} 3+3 \\ -1+3 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

$S = 3$

$$\varepsilon = \frac{1}{S-1} = \frac{1}{2}$$

$$a^2(1) = \text{poslin}\left(W^2 a^2(0)\right) = \begin{cases} \text{poslin}\left(\begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}\begin{bmatrix} 6 \\ 2 \end{bmatrix}\right) \\ \text{poslin}\left(\begin{bmatrix} 5 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 5 \\ 0 \end{bmatrix} \end{cases}$$

para la segunda iteración

$$a^2(2) = \text{poslin}\left(W^2 a^2(1)\right) = \begin{cases} \text{poslin}\left(\begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ 0 \end{bmatrix}\right) \\ \text{poslin}\left(\begin{bmatrix} 5 \\ -2.5 \end{bmatrix}\right) = \begin{bmatrix} 5 \\ 0 \end{bmatrix} \end{cases}$$

Se ve que la red converge al tener como salida $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$

como es mayor el elemento 1,1 este indica que es un plátano

Piña

$$p_2 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

$$a^1 = \begin{bmatrix} -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$a^1 = \begin{bmatrix} -1+3 \\ 3+3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$$

$$S = 3 \quad \varepsilon = \frac{1}{S-1} = \frac{1}{2}$$

$$a^2(1) = \text{poslin}\left(W^2 a^2(0)\right) = \begin{cases} \text{poslin}\left(\begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 6 \end{bmatrix}\right) \\ \text{poslin}\left(\begin{bmatrix} -1 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 5 \end{bmatrix} \end{cases}$$

Para la 2ª iteración

$$a^2(2) = \text{poslin}\left(W^2 a^2(1)\right) = \begin{cases} \text{poslin}\left(\begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 5 \end{bmatrix}\right) \\ \text{poslin}\left(\begin{bmatrix} -2.5 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 5 \end{bmatrix} \end{cases}$$

Se ve que la red converge al tener como salida $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$

como es mayor el elemento 2,1 este indica que es una piña

Piña redonda

$P_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ $\quad a^1 = \begin{bmatrix} -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

$a^1 = \begin{bmatrix} -3+3 \\ 3+3 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$

$S = 3 \quad \varepsilon = \frac{1}{S-1} = \frac{1}{2}$

$a^2(1) = \mathrm{poslin}(W^2 a^2(0)) = \begin{cases} \mathrm{poslin}\left(\begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 6 \end{bmatrix}\right) \\ \mathrm{poslin}\left(\begin{bmatrix} -3 \\ 6 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 6 \end{bmatrix} \end{cases}$

Para la Segunda Iteración

$a^2(2) = \mathrm{poslin}(W^2 a^2(1)) = \begin{cases} \mathrm{poslin}\left(\begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 6 \end{bmatrix}\right) \\ \mathrm{poslin}\left(\begin{bmatrix} -3 \\ 6 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 6 \end{bmatrix} \end{cases}$

Se ve que la red converge al tener como salida $\begin{bmatrix} 0 \\ 6 \end{bmatrix}$

Como es mayor el elemento 2,1 este indica que es una piña

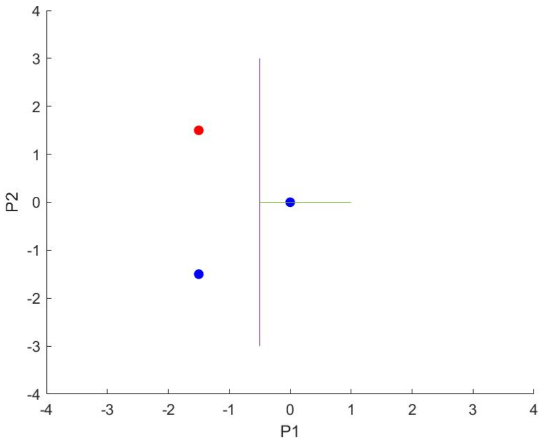
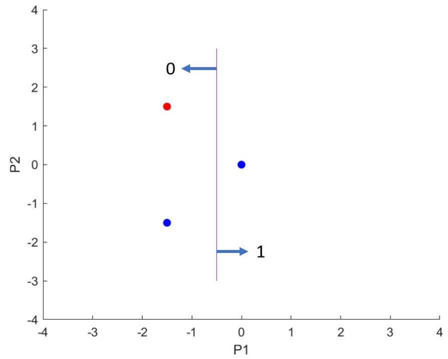
2. Se quiere entrenar un perceptrón con el siguiente conjunto de entrenamiento:

$$\left\{P_1 = \begin{bmatrix} -1.5 \\ -1.5 \end{bmatrix}, t_1 = 0\right\} \left\{P_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, t_2 = 0\right\} \left\{P_3 = \begin{bmatrix} -1.5 \\ 1.5 \end{bmatrix}, t_3 = 1\right\}$$

Las condiciones iniciales para la matriz de pesos y el sesgo son:

$W(0) = [1 \quad 0], \; b(0) = 0.5$

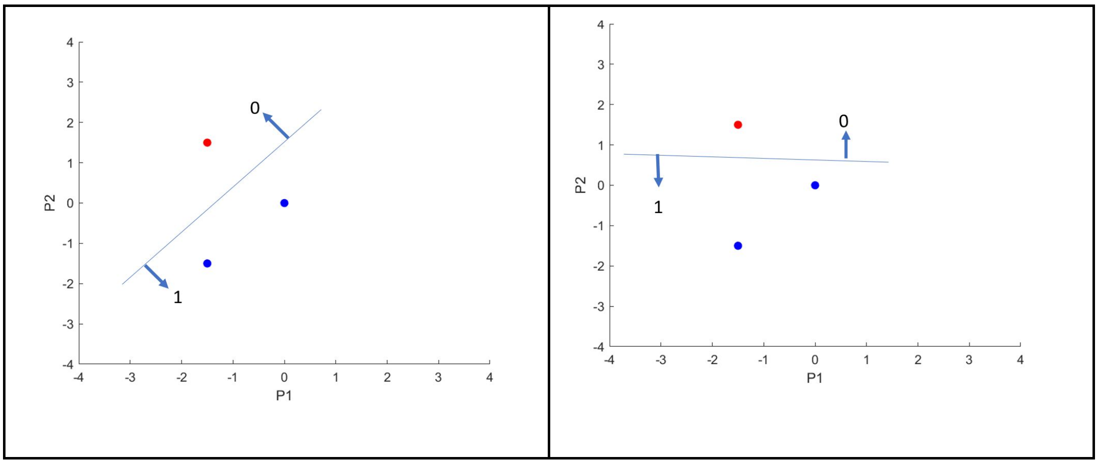
Grafique la frontera de decisión inicial así como el vector de pesos y los vectores de entrada. ¿Cuáles patrones son clasificados correctamente cuando se usan los valores iniciales de pesos y sesgo? Realice su gráfica en Matlab.

Gráfica con la frontera de decisión inicial y con el vector de pesos	Se proponen las direcciones del vector de pesos
	

Prueba del Perceptrón: evaluación y resultado	T esperada	¿Se acepta el resultado?
Para p1 $a = \text{hardlim}\left([1\ 0]\begin{bmatrix}-1.5\\-1.5\end{bmatrix} + 0.5\right) = 0$ a=harlim(W**p1**)= 0	t1=0	sí (color azul) t1=0,
Para p2 $a = \text{hardlim}\left([1\ 0]\begin{bmatrix}0\\0\end{bmatrix} + 0.5\right) = 1$ a=harlim(W**p2**)=1	t1=0	No (color azul) t2=0
Para p3 $a = \text{hardlim}\left([1\ 0]\begin{bmatrix}-1.5\\1.5\end{bmatrix} + 0.5\right) = 0$ a=harlim(W**p3**)=0	t1=1	No (color rojo) t3=1

El único patrón clasificado correctamente es el **p1**

Como anexo a esta pregunta, yo esperaría que mi límite de decisión clasifique de alguna de la siguiente manera:

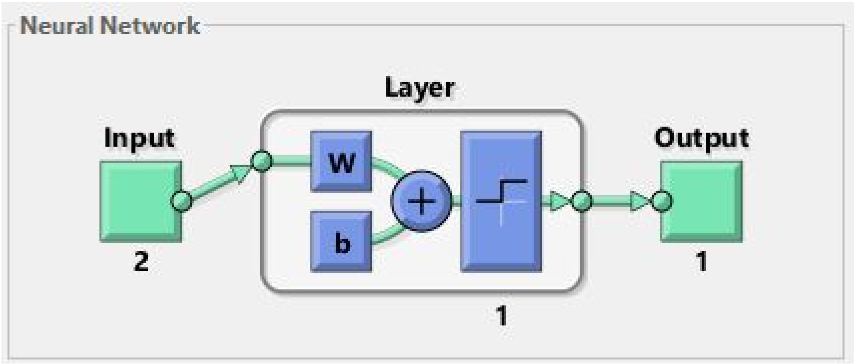


Entrene un perceptrón utilizando la regla de aprendizaje y obtenga el valor de la matriz de pesos y sesgo que hacen que funcione correctamente así como la arquitectura de la red, para ello utilice el toolbox de Matlab.

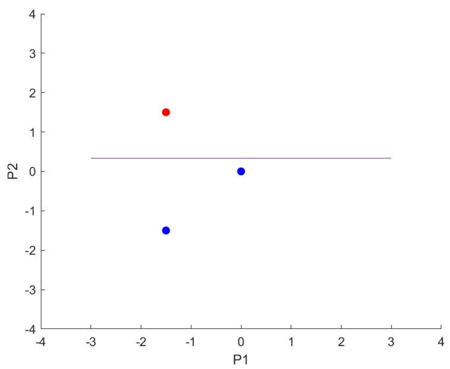
%Inputs
[-1.5 0 -1.5; -1.5 0 1.5]

%Targets
[0 0 1]

%Peso
W= [0 3];
%Bias
b=[-1]



Grafique la frontera de decisión final en Matlab.



SCRIPT

Problema2.m

3. Se tienen seis categorías de vectores de entrada.

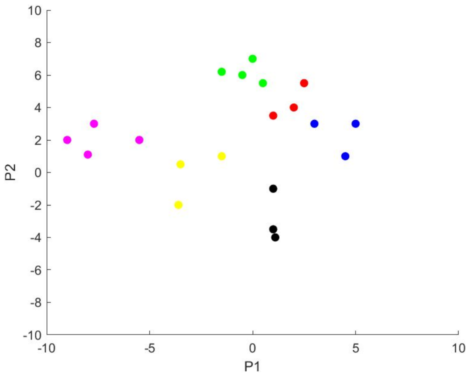
Categoria I: $\left\{\begin{bmatrix}4.5\\1\end{bmatrix}\begin{bmatrix}5\\3\end{bmatrix}\begin{bmatrix}3\\3\end{bmatrix}\right\}$, Categoria II: $\left\{\begin{bmatrix}1\\3.5\end{bmatrix}\begin{bmatrix}2\\4\end{bmatrix}\begin{bmatrix}2.5\\5.5\end{bmatrix}\right\}$,
Categoria III: $\left\{\begin{bmatrix}1\\-1\end{bmatrix}\begin{bmatrix}1\\-3.5\end{bmatrix}\begin{bmatrix}1.1\\-4\end{bmatrix}\right\}$, Categoria IV: $\left\{\begin{bmatrix}0.5\\5.5\end{bmatrix}\begin{bmatrix}0\\7\end{bmatrix}\begin{bmatrix}-0.5\\6\end{bmatrix}\begin{bmatrix}-1.5\\6.2\end{bmatrix}\right\}$
Categoria V: $\left\{\begin{bmatrix}-1.5\\1\end{bmatrix}\begin{bmatrix}-3.5\\0.5\end{bmatrix}\begin{bmatrix}-3.6\\-2\end{bmatrix}\right\}$, Categoria VI: $\left\{\begin{bmatrix}-5.5\\2\end{bmatrix}\begin{bmatrix}-8\\1.1\end{bmatrix}\begin{bmatrix}-7.7\\3\end{bmatrix}\begin{bmatrix}-9\\2\end{bmatrix}\right\}$

- **Diseñe una red perceptrón que reconozca esas cuatro categorías. Grafique las fronteras de decisión y los puntos.**

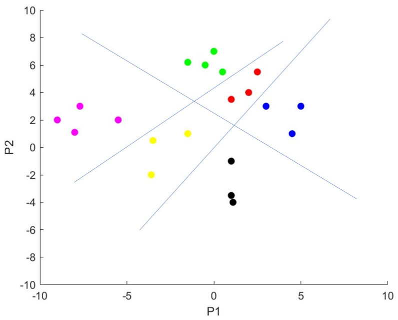
a) Se debe entrenar a un perceptrón con la función hardlim de Matlab para obtener su clasificación respectiva.
b) Entender qué es lo que tenemos y lo que deseamos.

Entradas	Salidas
Categoria I: $\left\{\begin{bmatrix}4.5\\1\end{bmatrix}\begin{bmatrix}5\\3\end{bmatrix}\begin{bmatrix}3\\3\end{bmatrix}\right\}$, Categoria II: $\left\{\begin{bmatrix}1\\3.5\end{bmatrix}\begin{bmatrix}2\\4\end{bmatrix}\begin{bmatrix}2.5\\5.5\end{bmatrix}\right\}$, Categoria III: $\left\{\begin{bmatrix}1\\-1\end{bmatrix}\begin{bmatrix}1\\-3.5\end{bmatrix}\begin{bmatrix}1.1\\-4\end{bmatrix}\right\}$, Categoria IV: $\left\{\begin{bmatrix}0.5\\5.5\end{bmatrix}\begin{bmatrix}0\\7\end{bmatrix}\begin{bmatrix}-0.5\\6\end{bmatrix}\right.$ Categoria V: $\left\{\begin{bmatrix}-1.5\\1\end{bmatrix}\begin{bmatrix}-3.5\\0.5\end{bmatrix}\begin{bmatrix}-3.6\\-2\end{bmatrix}\right\}$, Categoria VI: $\left\{\begin{bmatrix}-5.5\\2\end{bmatrix}\begin{bmatrix}-8\\1.1\end{bmatrix}\right.$	Son las que necesito conocer, por lo tanto las debo de proponerlas.

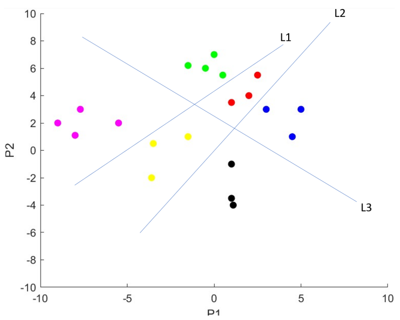
c) Graficar los puntos correspondientes a todas las clases (en Matlab).



d) Trazar líneas a mano que dividan nuestras clases con la intención de clasificar los grupos (no Matlab).



e) Enumerar las líneas con L1, L2 y L3 según nuestro criterio.



f) Establecer salidas correspondientes según la dirección de los pesos **W**$_{SxR}$ que son perpendiculares a las rectas.

g) Donde

h) S: número de neurona.

i) R: número de entrada.

j) Criterio para obtener direcciones:

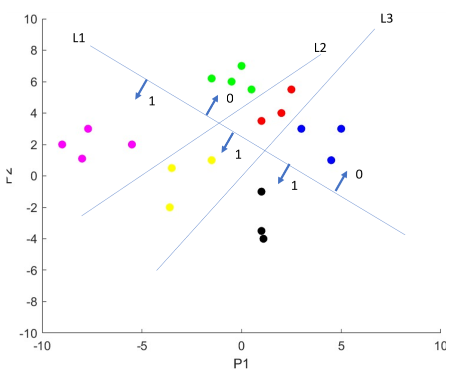
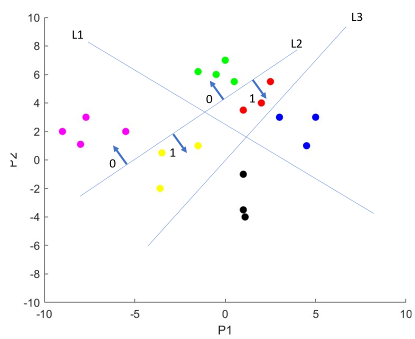
k) a = 0 n< 0

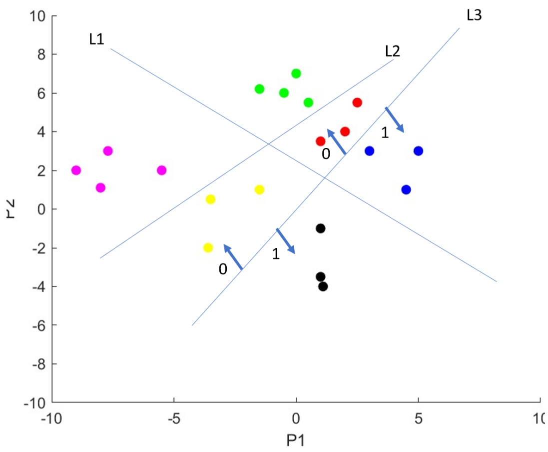
l) a = 1 n>= 1

m) En este caso solo importa saber los valores de a (que son 1 y 0) correspondientes a la función hardlim para conocer qué valor va a tener el vector perpendicular a la recta L1, L2 y L3.

n) Para establecer el valor de 1 se dice que todo vector perpendicular a L1, L2 o L3 que tenga el mismo sentido va a tener el valor 1, y en caso de que el vector apunte en dirección contraria va a tener el valor 0, es decir, si el vector perpendicular apunta en un sentido tiene 1 y si apunta en el otro tiene 0.

Resultados

Recta L1	Recta L2
	

Recta L3


o) Ahora se pueden obtener las salidas correspondientes a cada punto según el criterio anterior. A manera de ejemplo: para los vectores correspondientes a la clase CI se tiene el vector de salida $[0\ 0\ 0]T$ donde el primer 0 corresponde al valor del vector perpendicular a L1, el segundo valor 0 corresponde al valor del vector perpendicular a L2,y el tercer valor 0 corresponde al valor del vector perpendicular a L3.

Vectores de entrada	Salidas	
Categoria I: $\left\{\begin{bmatrix}4.5\\1\end{bmatrix}\begin{bmatrix}5\\3\end{bmatrix}\begin{bmatrix}3\\3\end{bmatrix}\right\}$, Categoria II: $\left\{\begin{bmatrix}1\\3.5\end{bmatrix}\begin{bmatrix}2\\4\end{bmatrix}\begin{bmatrix}2.5\\5.5\end{bmatrix}\right\}$, Categoria III: $\left\{\begin{bmatrix}1\\-1\end{bmatrix}\begin{bmatrix}1\\-3.5\end{bmatrix}\begin{bmatrix}1.1\\-4\end{bmatrix}\right\}$, Categoria IV: $\left\{\begin{bmatrix}0.5\\5.5\end{bmatrix}\begin{bmatrix}0\\7\end{bmatrix}\begin{bmatrix}-0.5\\6\end{bmatrix}\right.$ Categoria V: $\left\{\begin{bmatrix}-1.5\\1\end{bmatrix}\begin{bmatrix}-3.5\\0.5\end{bmatrix}\begin{bmatrix}-3.6\\-2\end{bmatrix}\right\}$, Categoria VI: $\left\{\begin{bmatrix}-5.5\\2\end{bmatrix}\begin{bmatrix}-8\\1.1\end{bmatrix}\right.$	$T_1=\left\{\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix}\right\}$ $T_2=\left\{\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ $T_3=\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix}\right\}$	$T_4=\left\{\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$ $T_5=\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix}\right\}$ $T_6=\left\{\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$

Use el toolbox de Matlab.¿Cuáles son los valores del peso y sesgo obtenidos?

- Entrenar a un perceptrón con la ayuda de nntool (de Matlab) para obtener los el peso (W) y el sesgo (b).
  - Construir el vector de entradas

    [4.5 5 3 1 2 2.5 1 1 1.1 0.5 0 -0.5 -1.5 -1.5 -3.5 -3.6 -5.5 -8 -7.7 -9;
    1 3 3 3.5 4 5.5 -1 -3.5 -4 5.5 7 6 6.2 1 0.5 -2 2 1.1 3 2]
  - Construir el vector de targets (propuesto)

    [0 0 0 0 0 0 1 1 1 0 0 0 0 1 1 1 1 1 1 1;
    1 1 1 1 1 1 1 1 1 0 0 0 0 1 1 1 0 0 0 0;
    1 1 1 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0];

  Obtener W y b con nntool

W = [-2.5 -3; 1.4 -2.5; 6.5 -4.5]; b = [1; 11; -2];

Graficar las líneas en Matlab.

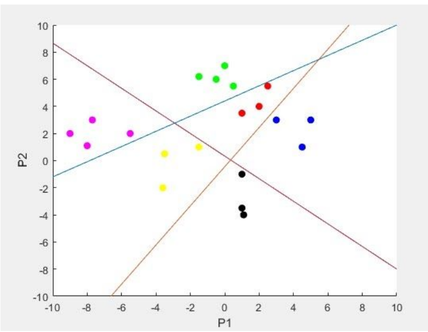
a. Crear la rectas para graficar

L1	L2
-2.5p1-3p2+1=0 Para el punto 1 si P1 = -10 P2 = 8.6666 Para el punto 2 si P1 = 10 P2 = -8	1.4p1-2.5p2+1=0 Para el punto 1 si P1 = -10 P2 = -1.2 Para el punto 2 si P1 = 10 P2 = 10

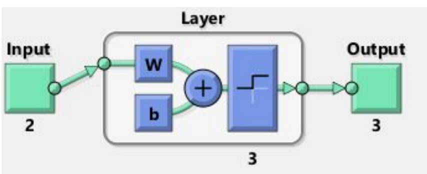
L3	
6.5p1-4.5p2-2=0 Para el punto 1 si P1 = -10 P2 = -14.8888 Para el punto 2 si P1 = 10 P2 = 14	%recta 1 plot([-10 10],[8.6666 -8]) %recta 2 plot([-10 10],[-1.2 10]) %recta 3 plot([-10 10],[-14.8888 14])

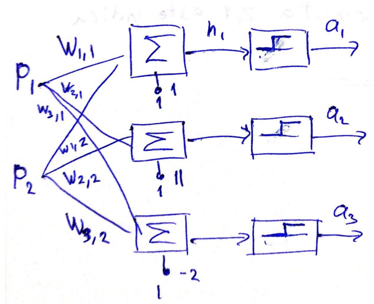
b. Correr el programa.

c. Ver el resultado (la hipótesis se cumple).



- **Anexar el diagrama de la arquitectura de la red.**

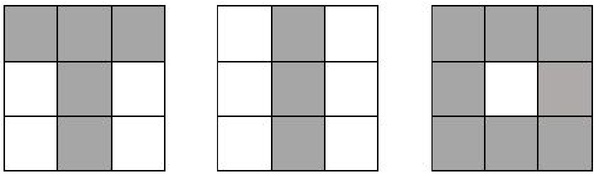




SCRIPT

Problema3.m

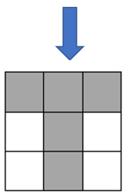
4. **Diseñe una red autoasociativa utilizando regla de Hebb y regla de la pseudoinversa que aprenda las tres letras mostradas. Pruebe la salida de la red diseñada con regla de Hebb y con regla de la pseudoinversa con cada una de las entradas y concluya sobre los resultados obtenidos. Elabore un archivo en MATLAB e imprima los procedimientos y resultados.**



P1 P2 P3

Los patrones que queremos almacenar se muestran desplegados en una cuadrícula de 3x3.
Para escribir el vector **p1** se toma la siguiente consideración:

Comenzando de arriba hacia abajo los cuadros en blanco= -1, los cuadros en gris =1, posteriormente comenzaremos con la columna de la derecha y así sucesivamente.

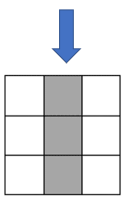


p1

$$\mathbf{p1} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \qquad \mathbf{p1 = pt}$$

De la misma forma se obtiene el vector **p2:**

Comenzando de arriba hacia abajo los cuadros en blanco= -1, los cuadros en gris =1, posteriormente comenzaremos con la columna de la derecha y así sucesivamente.

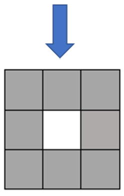


p2

$$\mathbf{p2} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \qquad \mathbf{p2 = pt}$$

De la misma forma se obtiene el vector **p3:**

Comenzando de arriba hacia abajo los cuadros en blanco= -1, los cuadros en gris =1, posteriormente comenzaremos con la columna de la derecha y así sucesivamente.



p3

$$p3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad p3 = pt$$

Considerando los patrones anteriores

Solución en Matlab, se declaran los tres patrones de entrada.

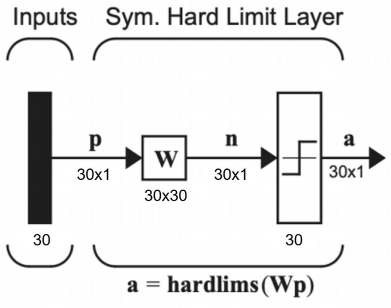
```
% Problema #4.
clear all, clc
p1=[1; -1; -1; 1; 1; 1; 1; -1; -1];
p2=[-1; -1; -1; 1; 1; 1; -1; -1; -1];
p3=[1; 1; 1; 1; -1; 1; 1; 1; 1];
```

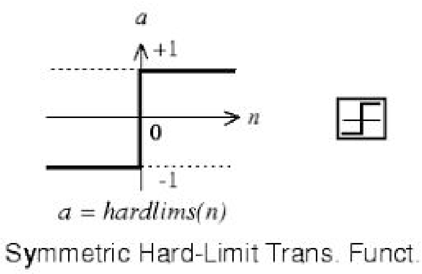
Como no se comprueba si son vectores ortonormales entre sí, se utiliza una función de activación hardlims para forzar a la red a tomar valores de 1 y -1.

Usando regla de Hebb para diseñar un autoasociador para los patrones (p1,p3 y p3).

Regla de Hebb → $\mathbf{W} = \mathbf{TP}^T$

Como los vectores de entrada están formados por dos valores únicamente, se modificó el asociador lineal de tal forma que a la salida sólo se puedan obtener únicamente valores de 1 y -1. De manera que los vectores de entrada sean iguales a las salidas objetivo en el conjunto de entrenamiento.El resultado se muestra en la siguiente imagen:





Entonces: $\mathbf{P} = \mathbf{T}$, $\quad \mathbf{W} = \mathbf{TP}^T$, $\mathbf{a} = \mathbf{hardlims}(\mathbf{W} \cdot \mathbf{p_t})$

```
%calculando el peso
```

```
P=[p1,p2,p3];
T=P;
Wh=P*transpose(T)

Wh =

     3     1     1     1    -1     1     3     1     1
     1     3     3    -1    -3    -1     1     3     3
     1     3     3    -1    -3    -1     1     3     3
     1    -1    -1     3     1     3     1    -1    -1
    -1    -3    -3     1     3     1    -1    -3    -3
     1    -1    -1     3     1     3     1    -1    -1
     3     1     1     1    -1     1     3     1     1
     1     3     3    -1    -3    -1     1     3     3
     1     3     3    -1    -3    -1     1     3     3


% calculando la salida de la red usando la regla de Hebb, cuando se presenta el vector p1
a1=hardlims(Wh*p1);
a1 =

     1
    -1
    -1
     1
     1
     1
     1
    -1
    -1


% calculando la salida de la red usando la regla de Hebb, cuando se presenta el vector p2
a2=hardlims(Wh*p2);
a2 =

    -1
    -1
    -1
     1
     1
     1
    -1
    -1
    -1


% calculando la salida de la red usando la regla de Hebb, cuando se presenta el vector p3
a3=hardlims(Wh*p3);
a3 =

     1
     1
     1
     1
    -1
     1
```

1
1
1

Usando regla de la pseudoinversa para diseñar un autoasociador para para los patrones (p1,p3 y p3).

Regla de la Pseudoinversa → $\mathbf{W} = \mathbf{T}\mathbf{P}^{+}$, $\mathbf{P}^{+} = (\mathbf{P}^{T}\mathbf{P})^{-1}\mathbf{P}^{T}$

```
%Peso con regla de la pseudoinversa
%P=[p1,p2,p3];
%T=P;
W=T*pinv(P)
W =
```

0.5000	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.5000	-0.0000	-0.0000
-0.0000	0.2000	0.2000	0.0000	-0.2000	0.0000	-0.0000	0.2000	0.2000
-0.0000	0.2000	0.2000	0.0000	-0.2000	0.0000	-0.0000	0.2000	0.2000
0.0000	-0.0000	0.0000	0.5000	-0.0000	0.5000	0.0000	0.0000	0.0000
0.0000	-0.2000	-0.2000	-0.0000	0.2000	-0.0000	0.0000	-0.2000	-0.2000
0.0000	-0.0000	0.0000	0.5000	-0.0000	0.5000	0.0000	0.0000	0.0000
0.5000	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.5000	-0.0000	-0.0000
-0.0000	0.2000	0.2000	0.0000	-0.2000	0.0000	-0.0000	0.2000	0.2000
-0.0000	0.2000	0.2000	0.0000	-0.2000	0.0000	-0.0000	0.2000	0.2000

```
%calculando la salida de la red usando la regla de la pseudoinversa, cuando se presenta el vector p1
a1_pseudo=hardlims(W*p1)
a1_pseudo =

     1
    -1
    -1
     1
     1
     1
     1
    -1
    -1
```

```
%calculando la salida de la red usando la regla de la pseudoinversa, cuando se presenta el vector p2
a2_pseudo=hardlims(W*p2)
a2_pseudo =

    -1
    -1
    -1
     1
     1
     1
    -1
```

-1
-1

%calculando la salida de la red usando la regla de la pseudoinversa, cuando se presenta el vector p3
a3_pseudo=hardlims(W*p3)
a3_pseudo =

1
1
1
1
-1
1
1
1
1

En conclusión, las redes diseñadas funcionan como se esperaban, tanto para la regla de Hebb como para regla de la pseudoinversa el error no existe pues
a1 = a1_pseudo = p1
a2 = a2_pseudo = p2
a3 = a3_pseudo = p3
esto se debe a que se utilizan patrones sencillos.

SCRIPT

Problema4.m

5. Se tiene el siguiente conjunto de entrenamiento:

$$\left\{P_1 = \begin{bmatrix}1\\0\end{bmatrix}, t_1 = 1\right\}\left\{P_2 = \begin{bmatrix}1\\1\end{bmatrix}, t_2 = -1\right\}\left\{P_3 = \begin{bmatrix}0\\1\end{bmatrix}, t_3 = 1\right\}$$

- **Muestre por qué este problema no puede ser resuelto si la red no tiene sesgo.**

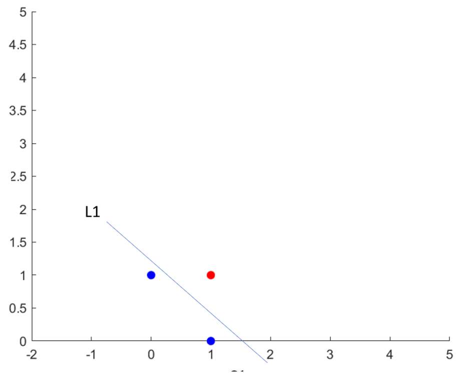
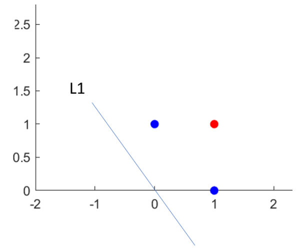
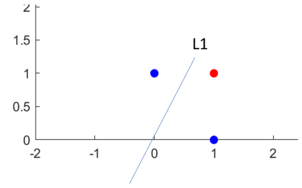
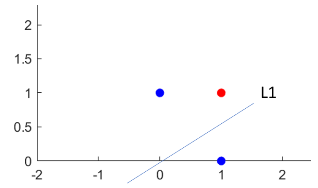
Si recordamos, la frontera de decisión de un perceptrón está dada por la recta definida por:

$$\mathbf{Wp} + b = 0$$

Si no hay sesgo, entonces b=0 y la frontera es definida por:

$$\mathbf{Wp} = 0$$

Lo que significa que la recta pasa por el origen, si consideramos los tres vectores **p1, p2 y p3**, dados en el problema y los graficamos en el plano, podemos ver que no existe una recta sin sesgo que pueda separar a los vectores, por lo tanto es necesario usar un sesgo para resolver el problema.

<table>
<tr><th>DESEADO
(SE NECESITA EL SESGO)</th><th>REALIDAD
(SIN SESGO)</th></tr>
<tr><td></td><td>No existe ninguna frontera de la decisión correcta.</td></tr>
</table>

- **Use la regla de la pseudoinversa para diseñar un perceptrón para ese conjunto de entrenamiento. Verifique que la red funcione correctamente. Realice sus procedimientos en una hoja blanca.**

  Trataremos al sesgo como otro peso, con entrada igual a 1, como se muestra en la arquitectura de la red. Entonces tendremos que aumentar a los vectores de entrada un elemento final con valor de 1.

  Por lo que ahora tendremos:

$$p_1' = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad p_2' = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad p_3' = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

  Recordando las salidas deseadas:

$$t_1 = 1 \quad t_2 = -1 \quad t_3 = 1$$

  De manera que:

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad T = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$$

No será necesario normalizar los vectores de entrada para usar la regla de la pseudoinversa.

$$\mathbf{W} = \mathbf{T}\mathbf{P}^{+},$$

$$\mathbf{P}^{+} = (\mathbf{P}^{T}\mathbf{P})^{-1}\mathbf{P}^{T}$$

$$W = TP^{+} = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}^{+}$$

$$P^{+} = (P^{T}P)^{-1} P^{T}$$

$$P^{T}.P = \left(\underset{3\times3}{\begin{bmatrix} \boxed{1 \quad 0 \quad 1} \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}} \underset{3\times3}{\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}} \right) = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \underset{3\times3}{\begin{bmatrix} 2 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}}$$

$c_{11} = 1 + 0 + 1 = 2$

$c_{12} = 1 + 0 + 1 = 2$

$c_{13} = 0 + 0 + 1 = 1$

$c_{21} = 1 + 0 + 1 = 2$

$c_{22} = 1 + 1 + 1 = 3$

$c_{23} = 0 + 1 + 1 = 2$

$c_{31} = 0 + 0 + 1 = 1$

$c_{32} = 0 + 1 + 1 = 2$

$c_{33} = 0 + 1 + 1 = 2$

$$(P^{T}P)^{-1} = \left(\begin{bmatrix} 2 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 2 \end{bmatrix}$$

$$(P^+P)^{-1} = \frac{1}{\det(P^+P)}\operatorname{adj}(P^+P) \qquad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \rightarrow C^T$$

$$\det(P^+P) = |P^+P| = \begin{vmatrix} 2 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{vmatrix} = 2\times3\times2 + 2\times2\times1 + 2\times2\times1 - 1\times3\times1 - 2\times2\times2 - 2\times2\times2 = 1$$

$C_{11} = 3\times2 - 2\times2 = 6-4 = 2$

$C_{12} = -[2\times2 - 1\times2] = -[4-2] = -2$

$C_{13} = 2\times2 - 1\times3 = 4-3 = 1$

$C_{21} = -[2\times2 - 2\times1] = -[4-2] = -2$

$C_{22} = 2\times2 - 1\times1 = 4-1 = 3$

$C_{23} = -[2\times2 - 1\times2] = -[4-2] = -2$

$C_{31} = 2\times2 - 3\times1 = 4-3 = 1$

$C_{32} = -[2\times2 - 2\times1] = -[4-2] = -2$

$C_{33} = 2\times3 - 2\times2 = 6-4 = 2$

$$\operatorname{adj}(P^+P) = \begin{bmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 2 \end{bmatrix}$$

$$(P^+P)^{-1} = \frac{\begin{bmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & 2 & 2 \end{bmatrix}}{1} = \begin{bmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & 2 & 2 \end{bmatrix}$$

$$P^+ = (P^TP)^{-1}P^T = \begin{bmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$a_{11} = 2 - 2 + 0 = 0$

$a_{12} = 0 - 2 + 1 = -1$

$a_{13} = 2 - 2 + 1 = 1$

$a_{21} = -2 + 3 + 0 = 1$

$a_{22} = 0 + 3 - 2 = 1$

$a_{23} = -2 + 3 - 2 = -1$

$a_{31} = 1 - 2 + 0 = -1$

$a_{32} = 0 - 2 + 2 = 0$

$a_{33} = 1 - 2 + 2 = 1$

Para la pseudoinversa de P

$$P^+ = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$$

Ps = 0. -1 1
1 1 -1
-1 0 1

Calculando la matriz W usando la regla de la pseudoinversa, se tiene:

$$W = TP^+ = \underbrace{\begin{bmatrix} 1 & -1 & 1 \end{bmatrix}}_{1\times3} \underbrace{\begin{bmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}}_{3\times3} = \underbrace{\begin{bmatrix} j_1 & j_2 & j_3 \end{bmatrix}}_{1\times3}$$

$j_1 = 0 - 1 - 1 = -2$

$j_2 = -1 - 1 + 0 = -2$

$j_3 = 1 + 1 + 1 = 3$

$W = [-2 \;\; -2 \;\; 3]$ Sesgo (b) $b = 3$ $W = [-2 \;\; -2]$, $b = 3$	**W=T*Ps** W = [-2 -2 3]
	Donde el último elemento de la matriz de pesos corresponde la valor del sesgo. Por lo tanto, la solución queda como: W = [-2 -2], b=3

$$W = [\,-2 \quad -2\,],\ b=3$$

Verificando que la red funciona correctamente se puede ver que la solución encontrada reconoce a los vectores dados.

$$-2P_1 - 2P_2 + 3 = 0$$

$$P_2 = \frac{-2P_1+3}{2}$$

Para el primer punto

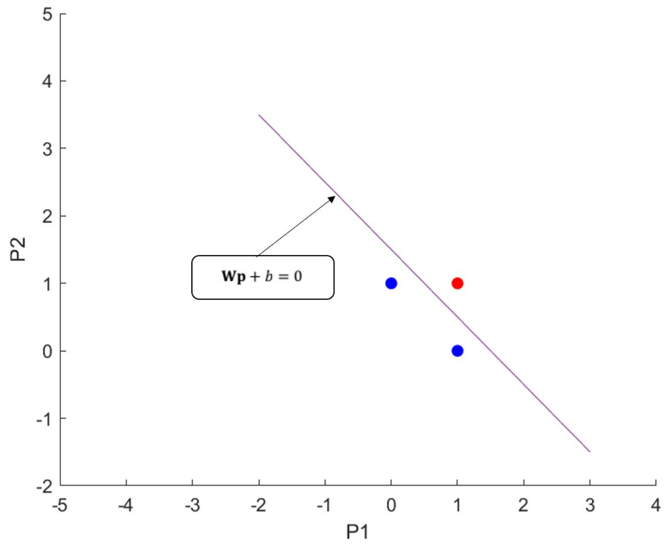
Si $P_1 = -2$

$P_2 = \frac{7}{2}$

Para el segundo punto

Si $P_1 = 3$

$P_2 = -\frac{3}{2}$



SCRIPT

Problema5.m