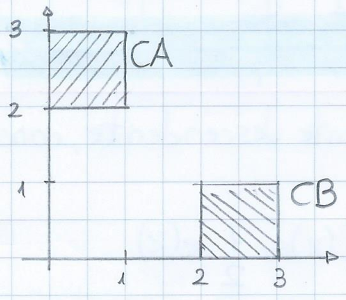
IA Aplicada a la Identificación y Control

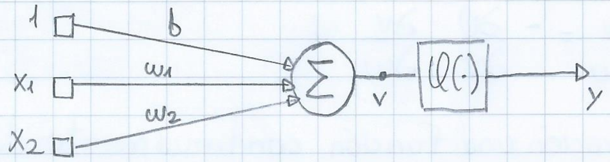
TP N°2 - Diseño de un clasificador

Desarrollar un perceptrón que realice la tarea de clasificación, capaz de clasificar los Clusters A y B



a)_ Realizar un diseño analítico para encontrar la sinápsis del perceptrón y graficar el hiperplano de separación

El perceptrón a desarrollar tendrá la siguiente estructura:



$$v = w_1 . x_1 + w_2 x_2 + b$$

$$y = \varphi(v)$$

Existen muchas funciones de activación posibles, pero para un desarrollo analítico consideraremos la función escalón:

$$\varphi(v) = \begin{cases} 1 & \text{si } v > 0 \\ 0 & \text{si } v \leq 0 \end{cases}$$

En base a esto:

$$y(w,b,x) = w_1 . x_1 + w_2 x_2 + b$$

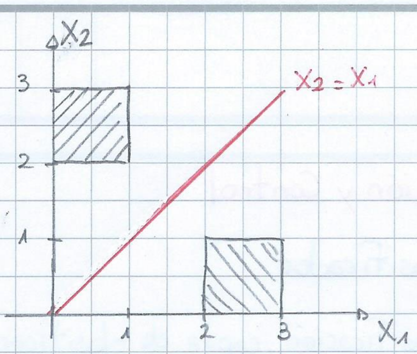
Para despejar x_2 en función del resto de variables y poder graficar el hiperplano de separación, tomaremos $y=0$

$$0 = w_1 . x_1 + w_2 . x_2 + b \Rightarrow x_2 = \underbrace{-\frac{w_1}{w_2}}_{\text{pendiente}} . x_2 - \underbrace{\frac{b}{w_2}}_{\text{ordenada al origen}}$$

Tomando el plano $x_1 = x_2 \Rightarrow -\frac{w_1}{w_2} = 1$ y $-\frac{b}{w_2} = 0 \Rightarrow$ $b = 0$

$w_1 = 1$

$w_2 = -1$



Cabe destacar que no es el único hiperplano que se podría trazar para separar los 2 Clusters.

b)_ La programación para el diseño automático puede verse en el archivo adjunto, desarrollado en lenguaje Python y Jupyter Notebook.

En dicho diseño se empleo el gradiente descendente, considerando el error cuadrótico instantaneo.

$$e(k) = d(k) - y(k) \quad ; \quad C(k) = \frac{1}{2} . e(k)$$

donde $d(k)$ es el valor deseado, $y(k)$ es la salida de la red neuronal y $C(k)$ el funcional de costo.
Para aplicar el gradiente descendente necesito obtener $\frac{\partial C}{\partial w}$, recordando que

$$y(w, b, x) = \varphi(v) = \varphi(w_1 . x_1 + w_2 . x_2 + b) \;, \; y \; \frac{\partial C}{\partial w} = e(k) . \frac{\partial e}{\partial w}$$

$$\frac{\partial e}{\partial w} = - \frac{\partial y}{\partial w} = - \frac{\partial y}{\partial \varphi} . \frac{\partial \varphi}{\partial v} . \frac{\partial v}{\partial w} = - \frac{\partial \varphi}{\partial v} . \frac{\partial v}{\partial w}$$

Empleando como función de activación una función continuamente diferenciable como una sigmoide:

$$\varphi(v) = \frac{1}{1 + e^{-a . v}}$$

$$\frac{\partial e}{\partial w_1} = a . \varphi(v)(1 - \varphi(v)) . x_1 \qquad \frac{\partial e}{\partial b} = - a . \varphi(v)(1 - \varphi(v))$$

$$\frac{\partial e}{\partial w_2} = a . \varphi(v)(1 - \varphi(v)) . x_2$$

Por lo tanto, la actualización de pesos será la siguiente:

$$w_1(k+1) = w_1(k) - \eta \frac{\partial C}{\partial w_1} = w_1(k) - \eta \frac{\partial e}{\partial w_1} . e(k)$$

$$w_2(k+1) = w_2(k) - \eta \frac{\partial C}{w_2} = w_2(k) - \eta \frac{\partial e}{\partial w_2} . e(k)$$

$$b(k+1) = b(k) - \eta \frac{\partial C}{\partial b} = b(k) - \eta \frac{\partial e}{\partial b} . e(k)$$

Donde η es el coeficiente de aprendizaje, relacionado con la velocidad de aprendizaje durante el entrenamiento