Introducción a las Redes Neuronales Artificiales y RNAs Avanzadas

Juan Humberto SOSSA AZUELA

E-mail: hsossa@cic.ipn.mx and humbertosossa@gmail.com

http://sites.google.com/site/cicvision/

Teoría detrás del modelo de M-P:

Ya vimos que la TLU es una neurona de dos estados: 0 y 1:

$$y_i = f\left(\sum_{j=1}^{n} w_{ij}x_j - \theta_j\right)$$

Donde:

x_i : vector de entradas binarias.
w_i: vector de pesos y θ_j es el bias asociado, w_i, θ_j son reales.
La función $f(\cdot)$ es la función escalón.

Para que esta neurona realice algo interesante, w_i, θ_j deben ser seleecionados de forma conveniente.

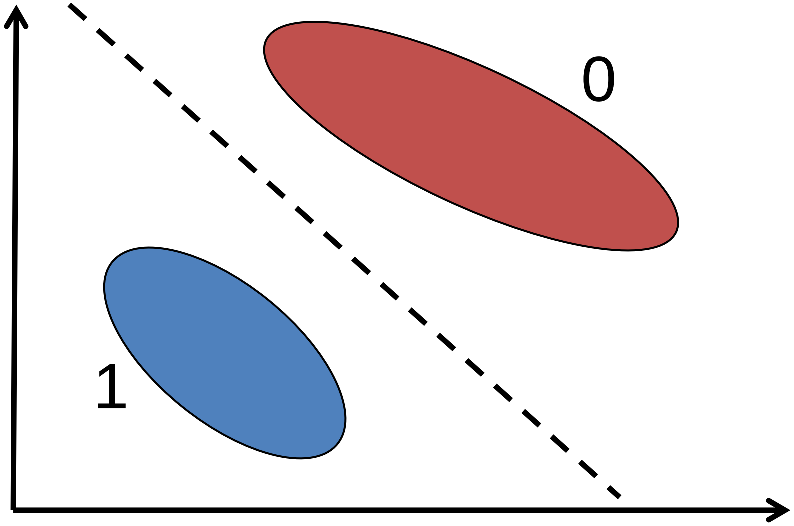
J. McCelland and D. Rumelhart. Parallel Distributed Processes. MIT Cambridge, MA, 1987.

TLUs, separabilidad lineal y vectores:

Interpretación geométrica de la acción de la TLU:

Una TLU separa los patrones de entrada en dos categorías, 0 para un patrón, 1 para el otro patrón.

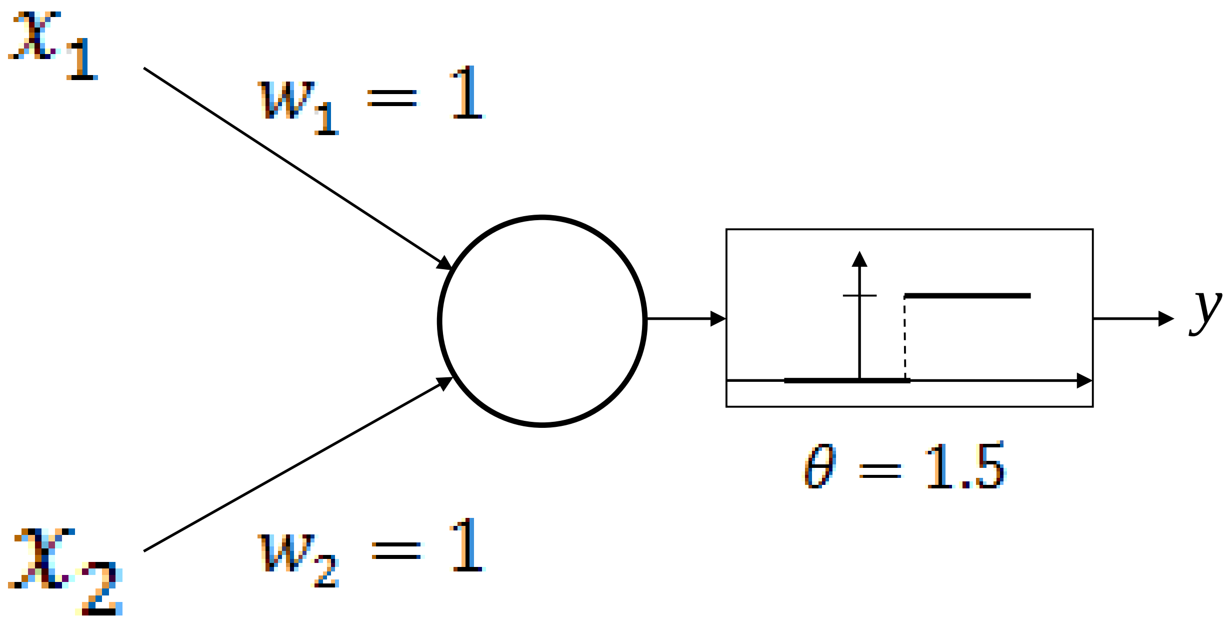
Estas categorías pueden pensarse como dos regiones en un espacio multidimensional separados por un híperplano.



Clasificación de patrones y espacio de entradas.

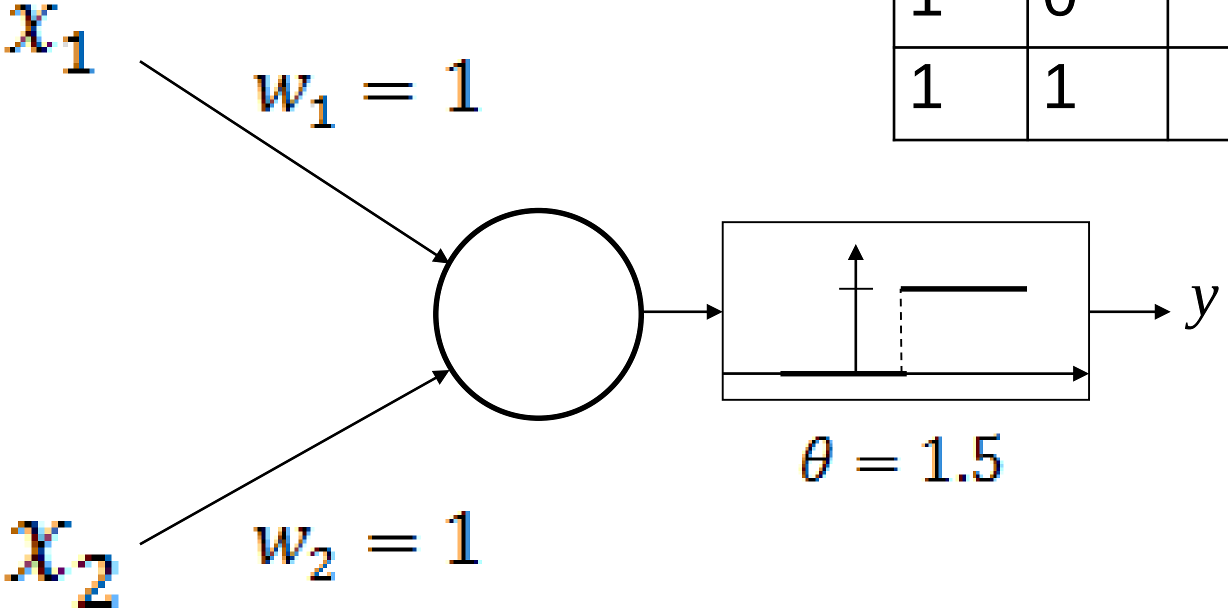
Ejemplo. Sea una TLU de dos entradas, con pesos

$$w_1 = 1 \qquad w_2 = 1 \qquad \theta = 1.5$$

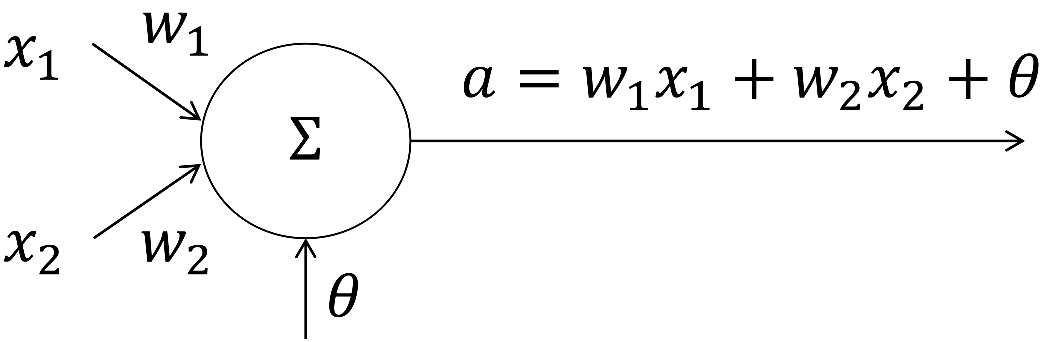


La respuesta a las cuatro posibles entradas se muestra en la tabla:

x_1	x_2	Activación	Salida
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1



Efecto de modificar el bias:



$$s = f(a) = \begin{cases} 1 \; si \; a \geq \theta \\ 0 \; si \; a < \theta \end{cases}$$

Ejemplo 1: Sean $w_1 = 1$ y $w_2 = 1$ y $\boldsymbol{\theta = 0.5}$

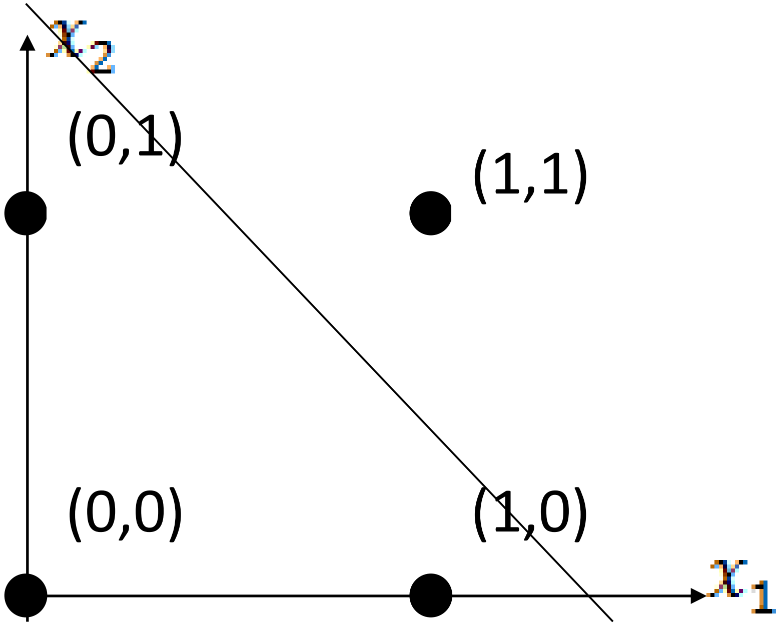
x_1	x_2	a	s
0	0	0.0	0
1	0	1.0	1
0	1	1.0	1
1	1	2.0	1

Ahora la neurona opera de manera diferente…!

Nótese como la TLU permite clasificar las entradas en dos clases, dependiendo del valor de y.

Cada patrón de entrada tiene dos componentes: x_1 y x_2

Estos cuatro patrones se pueden representar en un espacio dos-dimensional.



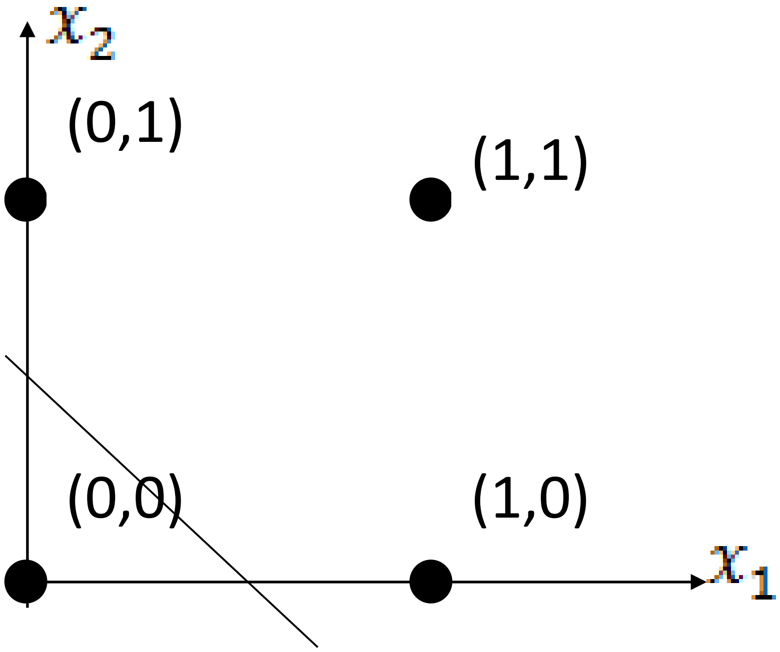
Cada patrón es representado por un punto en este **espacio de patrones** (EP), con dos coordenadas.

En el caso n-dimensional el EP será n-dimensional.

Nótese como la TLU permite clasificar las entradas en dos clases, dependiendo del valor de y.

Cada patrón de entrada tiene dos componentes: x_1 y x_2

Estos cuatro patrones se pueden representar en un espacio dos-dimensional.



Cada patrón es representado por un punto en este **espacio de patrones** (EP), con dos coordenadas.

En el caso n-dimensional el EP será n-dimensional.

Para $n > 3$ el PA no puede ser dibujado.

.

Esto no es un problema ya que las relaciones entre patrones se pueden expresar algebraicamente mediante la noción de **vector**.

Separación lineal entre clases:

Ya que la condición de separación ocurre cuando $a = \theta$ para el caso de dos entradas:

$$w_1 x_1 + w_2 x_2 = \theta$$

Separación lineal entre clases:

Ya que la condición de separación ocurre cuando $a = \theta$ para el caso de dos entradas:

$$w_1 x_1 + w_2 x_2 = \theta$$

Al restar $w_1 x_1$ de ambos lados: $w_2 x_2 = -w_1 x_1 + \theta$

Separación lineal entre clases:

Ya que la condición de separación ocurre cuando $a = \theta$ para el caso de dos entradas:

$$w_1 x_1 + w_2 x_2 = \theta$$

Al restar $w_1 x_1$ de ambos lados: $w_2 x_2 = -w_1 x_1 + \theta$

Al dividir ambos lados por w_2 $x_2 = -\left(\frac{w_1}{w_2}\right) x_1 + \left(\frac{\theta}{w_2}\right)$

Separación lineal entre clases:

Ya que la condición de separación ocurre cuando $a = \theta$ para el caso de dos entradas:

$$w_1 x_1 + w_2 x_2 = \theta$$

Al restar $w_1 x_1$ de ambos lados: $w_2 x_2 = -w_1 x_1 + \theta$

Al dividir ambos lados por w_2 $x_2 = -\left(\frac{w_1}{w_2}\right) x_1 + \left(\frac{\theta}{w_2}\right)$

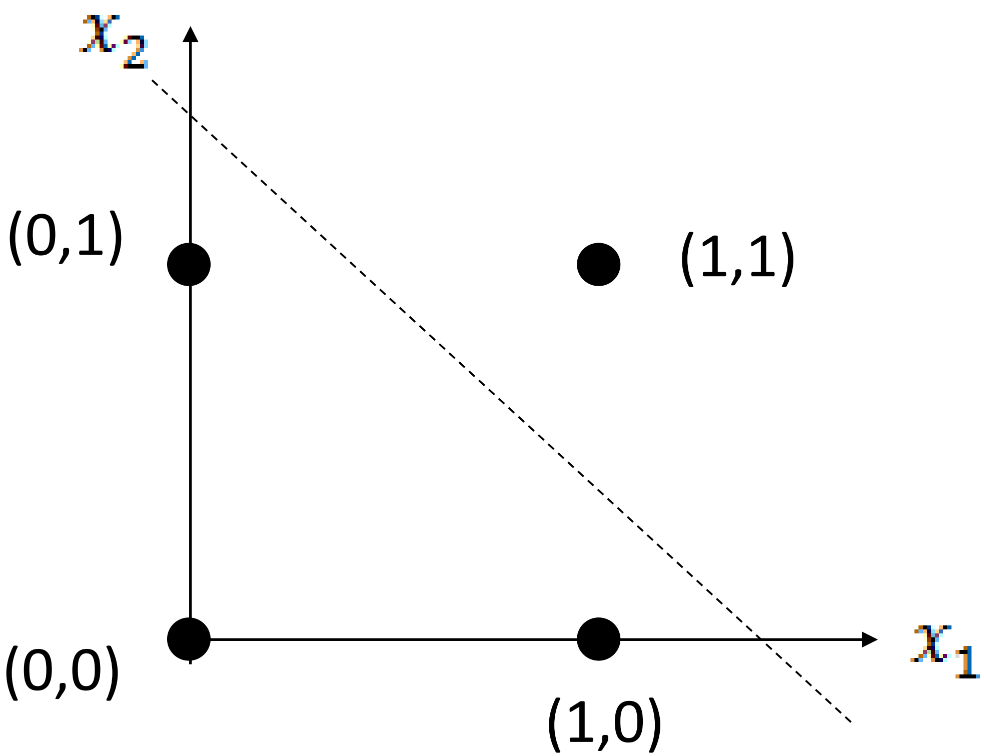
Esta ecuación es de la forma general: $x_2 = a x_1 + b$

Con a y b constantes.

Esta ecuación describe una recta con pendiente a e intercepto b en el eje x_2.

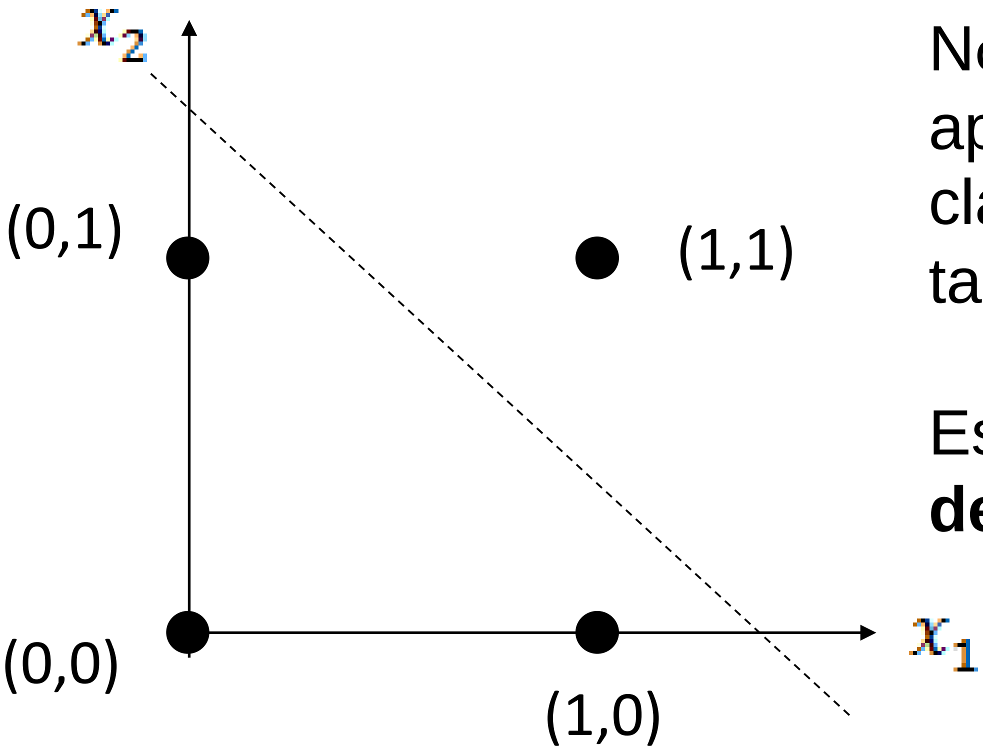
. Para la TLU del ejemplo, al insertar los valores de w_1 , w_2 , θ en la ecuación: $x_2 = -\left(\frac{w_1}{w_2}\right)x_1 + \left(\frac{\theta}{w_2}\right)$.

Se tiene que $a = 1$ y $b = 1.5$, como se muestra:



Nótese como las cuatro clases aparecen separadas en dos clases por esta recta, según la tabla.

Esta recta es llamada la **línea de decisión**.

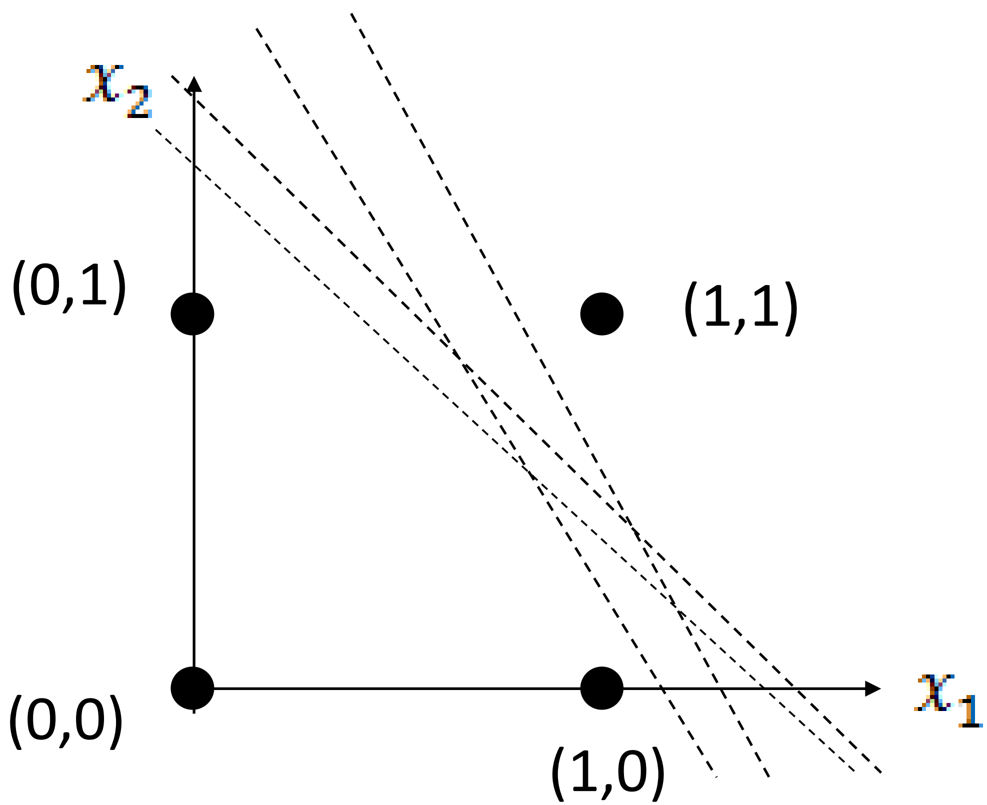


Nótese como las cuatro clases aparecen separadas en dos clases por esta recta, según la tabla.

Esta recta es llamada la **línea de decisión**.

x_1	x_2	Activación	Salida
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

Nótese que un sinnúmero de líneas pudieron haber sido usadas para separar los patrones, sin embargo **la resultante fue del hecho de la selección de los pesos y el umbral**.



En el caso de una TLU de tres entradas la separación se da por un **plano de separación**.

En el caso de una TLU de n entradas la separación se da por un **híperplano de separación**.

Ya que las TLUs están relacionadas con relaciones lineales como la dada por:

$$x_2 = -\left(\frac{w_1}{w_2}\right)x_1 + \left(\frac{\theta}{w_2}\right)$$

se dice que las TLUs son **clasificadores lineales** y su patrones son **linealmente separables**.

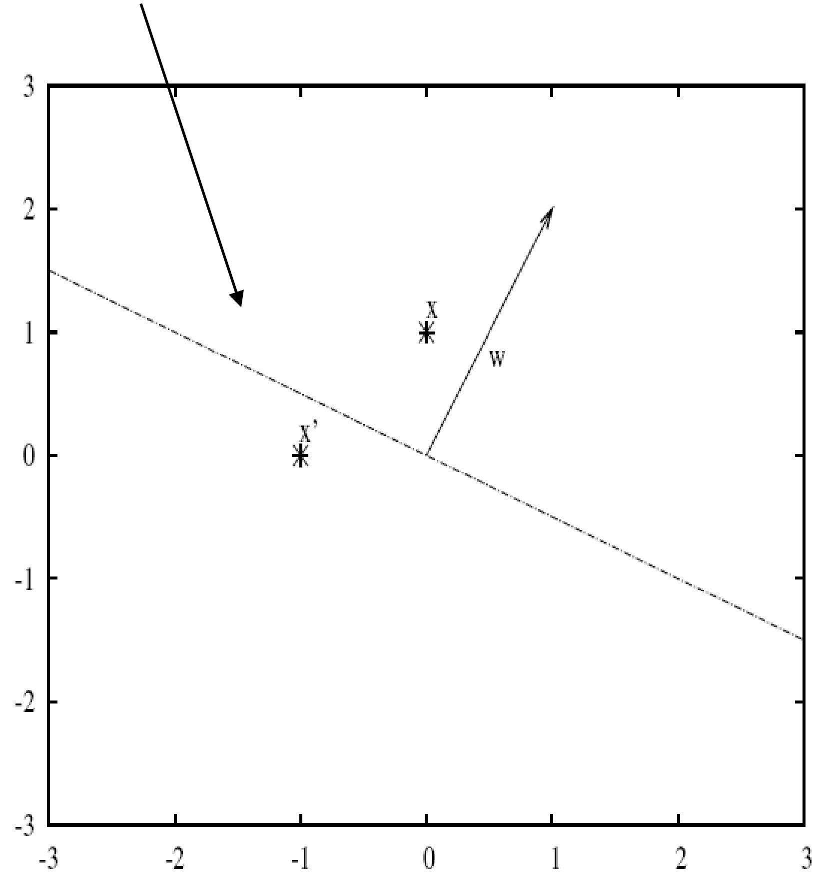
Lo inverso es también cierto: **cualquier clasificación binaria que no puede ser obtenida mediante una superficie de separación lineal puede ser obtenida mediante una TLU.**

.

LA TLU divide el espacio de rasgos en dos partes: $f(x) \geq 0$ y $f(x) < 0$.

Estas dos partes se definen por el híper-plano.

Note que **el plano** es siempre **perpendicular** al vector de pesos w.



Por ejemplo si: $w = (1\ 2);\ b = 0$

Para $x\ =\ (0\ 1)$:

$wx + b =$

$$((1\ 2) \cdot (0\ 1)) + 0 = (1 * 0) + (2 * 1) + 0 = 2 > 0$$

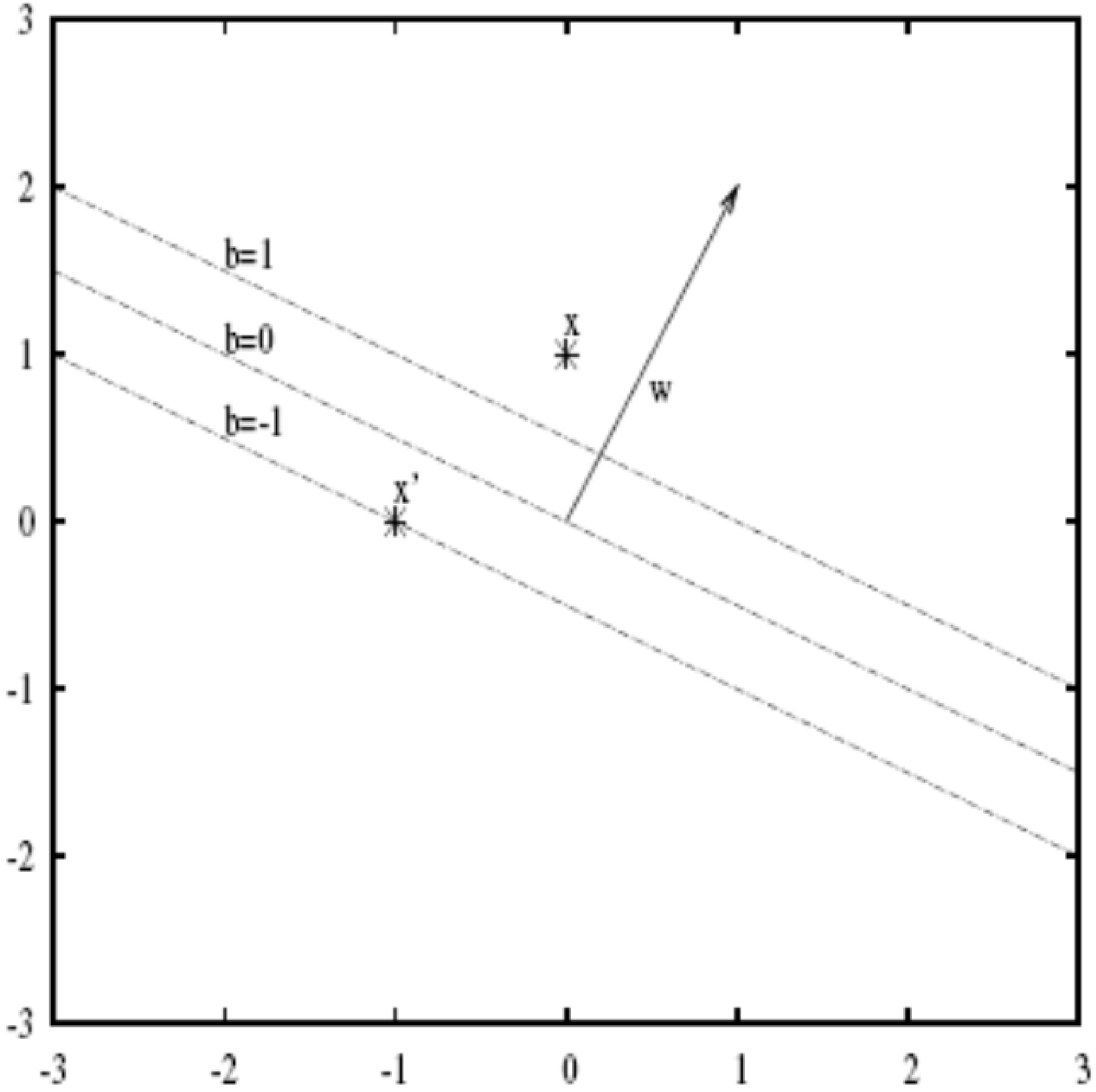
Para $x = (-1\ \ 0)$:

$wx + b =$

$$((1\ 2) \cdot (-1\ 0)) + 0 = (1 * -1) + (2 * 0) + 0 = -1 < 0$$

.

b determina la posición del híper-plano sobre el espacio:



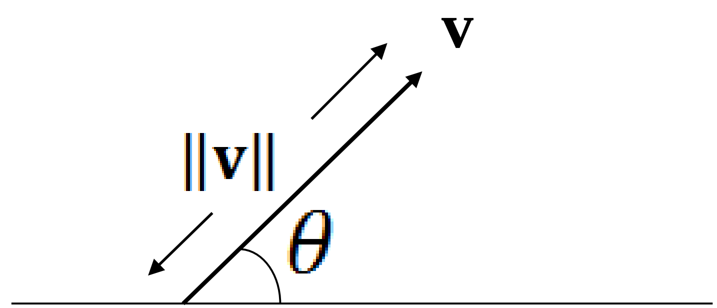
Vectores:

Vectores:

Los vectores son objetos que nos permiten representar cantidades con magnitud y dirección.

Por ejemplo el viento.

Vector simple:



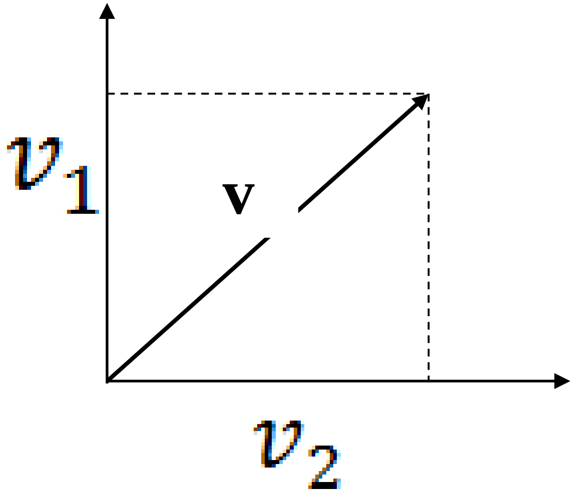
$||\mathbf{v}||$ es la magnitud del vector.

Un vector es definido por una par de números: $(||\mathbf{v}||, \theta)$

θ es el ángulo del vector con respecto a un eje de referencia.

Un vector se puede representar por sus proyecciones con los ejes (componentes):

Un vector se puede representar por sus proyecciones con los ejes (componentes):



El vector es ahora representado por dos números:

$$\mathbf{v} = (v_1, v_2)$$

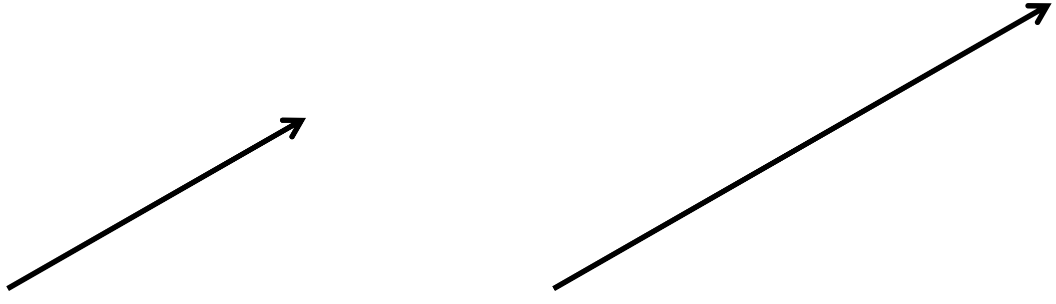
Notar que el orden es importante. (1,3) **no da lo mismo** que (3,1).

En n dimensiones: $\mathbf{v} = (v_1, v_2, \cdots, v_n)$

Suma vectorial y multiplicación escalar.

Multiplicación de un vector por un escalar.

El multiplicar un vector por un escalar modifica su magnitud.

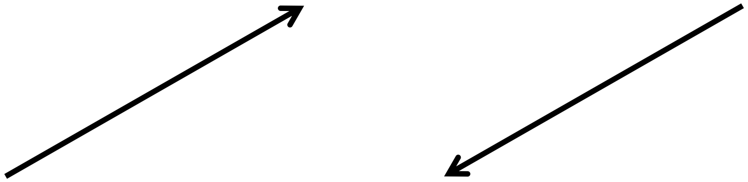


Suma vectorial y multiplicación escalar.

Multiplicación de un vector por un escalar.

El multiplicar un vector por un escalar modifica su magnitud.

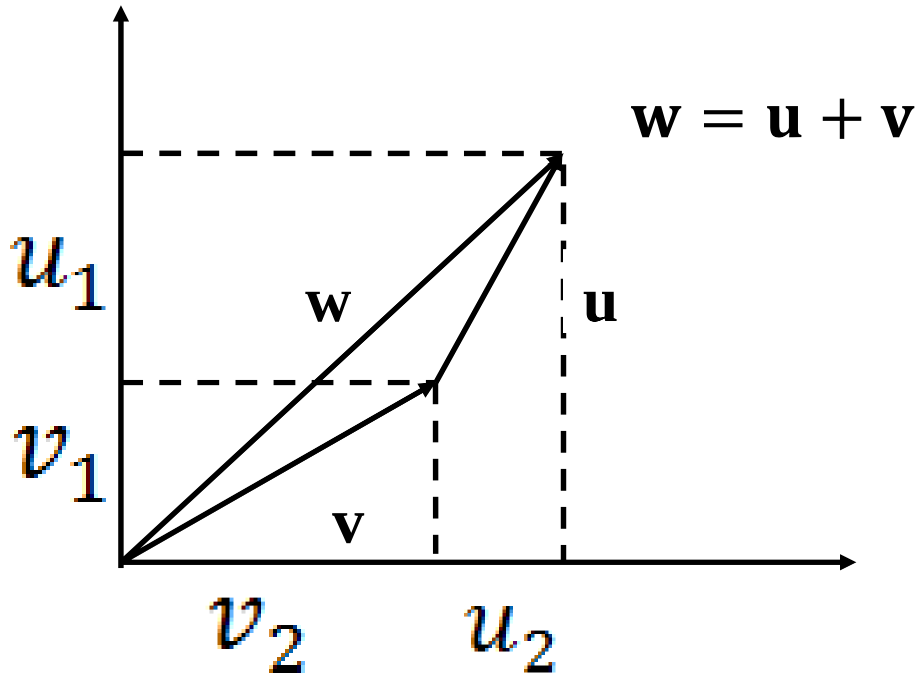
Si el escalar es negativo cambia la dirección del vector.



En términos de componentes: $k\mathbf{v} = (kv_1, kv_2, \cdots, kv_n)$

Suma vectorial.

Geométricamente, dos vectores pueden ser sumados simplemente al conectar uno con el otro:



En términos de componentes:

$$w_1 = u_1 + v_1 \quad w_2 = u_2 + v_2$$

En n dimensiones: $$\mathbf{w} = (u_1 + v_1 \;\; u_2 + v_2 \cdots \; u_n + v_n)$$

Substracción vectorial:

La substracción vectorial es una combinación de adición y multiplicación escalar:

$$u - v = u + (-1)v$$

Longitud de un vector:

En el caso 2-D: $$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2}$$

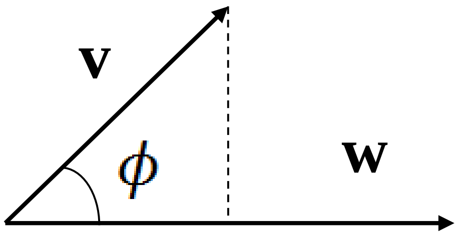
En n dimensiones: $$\|\mathbf{v}\| = \left[\sum_{i=1}^{n} v_i^2\right]^{\frac{1}{2}}$$

Comparación entre vectores.

Producto interno (forma geométrica).

Producto interno (forma algebraica).

Comparación entre vectores.



Producto interno (forma geométrica).

Supongamos dos vectores **v** y **w** separados por un ángulo ϕ, el **producto interno** (producto escalar) entre **v** y **w**, se define como:

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos\phi$$

Notar que: $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$

Nos dice que tan alineados están dos vectores.

Si las longitudes de los vectores son fijas entonces $\mathbf{v} \cdot \mathbf{w}$ depende sólo de $\cos\phi$

Casos:

1 Si los dos vectores están más o menos alineados el **producto interno es positivo**.

Casos:

1 Si los dos vectores están más o menos alineados el **producto interno es positivo**.

2 Cuando **v** y **w** forman un ángulo de 90 grados: $\mathbf{v} \cdot \mathbf{w} = \mathbf{0}$

Casos:

1 Si los dos vectores están más o menos alineados el **producto interno es positivo**.

2 Cuando **v** y **w** forman un ángulo de 90 grados: $\mathbf{v} \cdot \mathbf{w} = \mathbf{0}$

3 Cuando los dos vectores apuntan más o menos en direcciones opuestas, el **producto interno es negativo**.

Producto interno (forma algebraica).

Se pude demostrar que $\mathbf{v} \cdot \mathbf{w} = u_1 v_1 + u_2 v_2$

Para cualesquiera dos vectores v, w en 2-D.

En n dimensiones es claro que: $\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^{n} w_i v_i$

Este valor se debe interpretar igual que en el caso 2-D:

1 SI $\mathbf{v} \cdot \mathbf{w}$ es positivo los dos vectores apuntan más o menos en la misma dirección.

2 SI $\mathbf{v} \cdot \mathbf{w}$ es negativo los dos vectores apuntan en direcciones opuestas, y

3 SI $\mathbf{v} \cdot \mathbf{w} = \mathbf{0}$ forman un ángulo recto.

No hay que tratar de visualizar esto, pero hacer la analogía con el cado 2-D.

Es como usar imágenes 2-D para representar escenas 3-D.

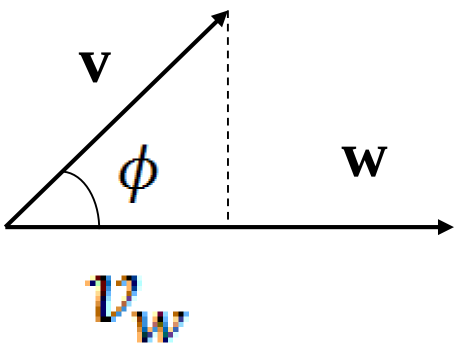
Esto ayuda a pensar en las propiedades geométricas de los objetos.

Si **v=w**:

$$\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^{n} w_i v_i \longrightarrow \mathbf{v} \cdot \mathbf{v} = \sum_{i=1}^{n} v_i v_i = \|\mathbf{v}\|^2$$

NOTA: El cuadrado de la longitud de un vector es igual al producto interno del vector consigo mismo.

Proyección de un vector.

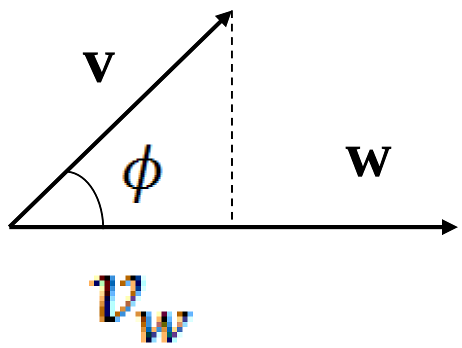


Sean dos vectores **v** y **w**:

¿Cuánto de **v** yace en la dirección de **w**?

Este segmento es llamado la proyección v_w de **v** en **w**.

Proyección de un vector.



Sean dos vectores **v** y **w**:

¿Cuánto de v yace en la dirección de **w**?

Este segmento es llamado la proyección v_w de **v** en w.

Usando la definición del coseno:

$$v_w = \|\mathbf{v}\| \cos \phi$$

Se puede reformular esto en términos del producto interno al multiplicar y dividir por la magnitud de **w**:

$$v_w = \frac{\|\mathbf{v}\| \|\mathbf{w}\| \cos \phi}{\|\mathbf{w}\|} = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|}$$

TLUs y separabilidad lineal:

TLUs y separabilidad lineal.

La discusión anterior está motivada por el hecho de probar que TLU es separador lineal universal independiente de la dimensionalidad del espacio de los patrones.

Al usar el concepto de producto interno, la activación de un TLU de n entradas se puede expresar como:

$$a = \mathbf{w} \cdot \mathbf{x}$$

Desviaciones.

Al momento de los cálculos puede ocurrir que a la salida de la TLU haya un resultado diferente al esperado.

En general para un x arbitrario, su proyección sobre **w** viene dada por

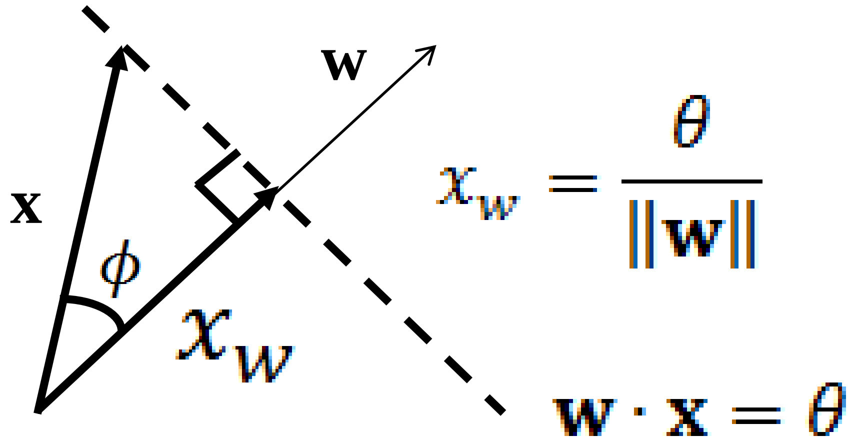
$$x_w = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|}$$

De acuerdo a lo visto, la condición de disparo ocurre cuando:

$$\mathbf{w} \cdot \mathbf{x} = \theta$$

Luego, entonces: $$x_w = \frac{\theta}{\|\mathbf{w}\|}$$

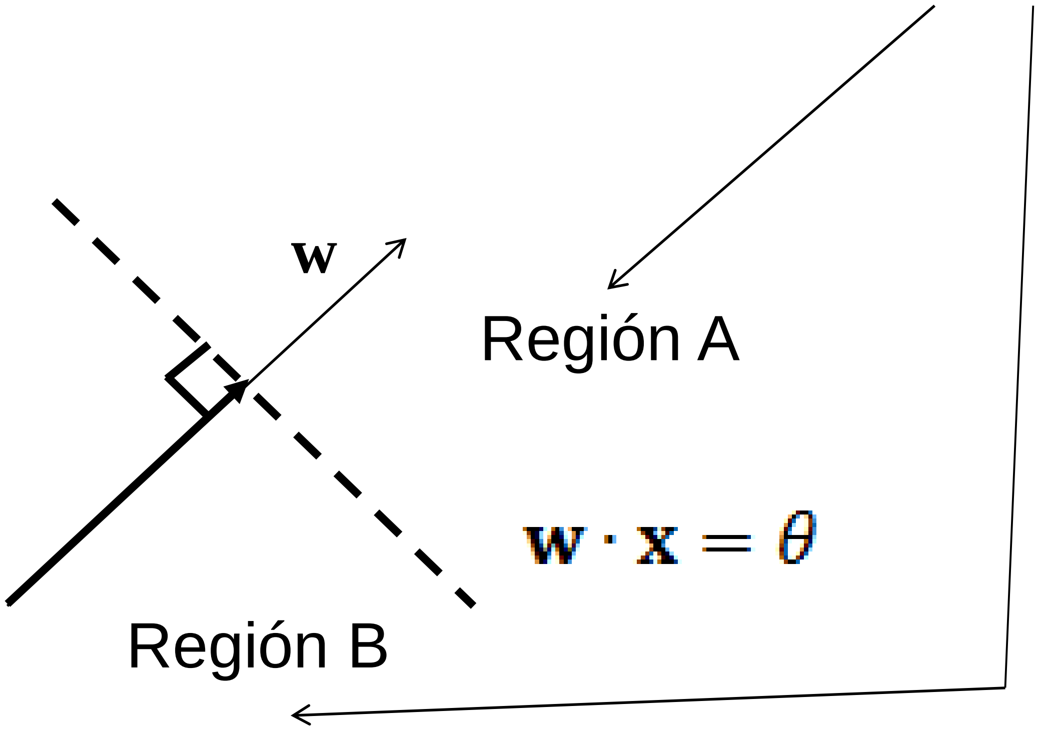
Si se asume que **w** y θ son constantes, la proyección v_w es constante y en 2D, **x** DEBE llegar exactamente a la perpendicular al vector de pesos, como se muestra:



La relación $\mathbf{v} \cdot \mathbf{w} = \theta$ define una línea recta perpendincular al vector **w**.

Todo lo dicho es válido en n dimensiones.

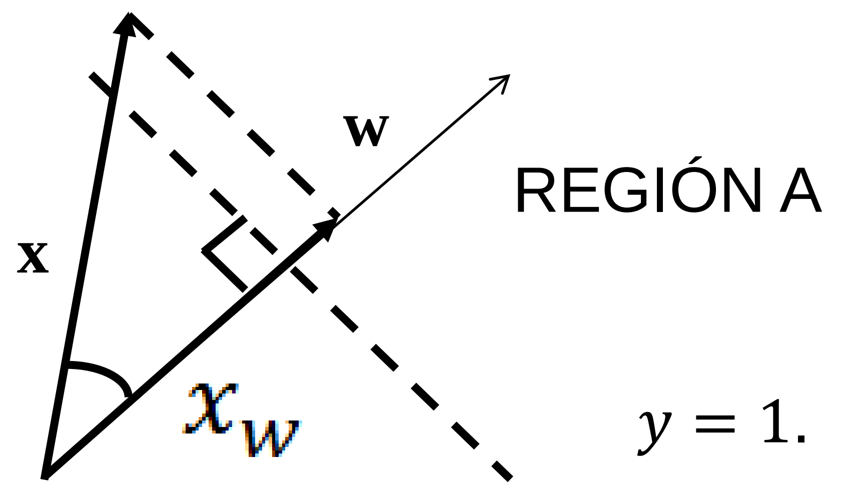
El plano $\mathbf{v} \cdot \mathbf{w} = \theta$ **divide** al espacio en dos regiones: A y B:



Cuando $\mathbf{w} \cdot \mathbf{x} = \theta$, $y = 1$.

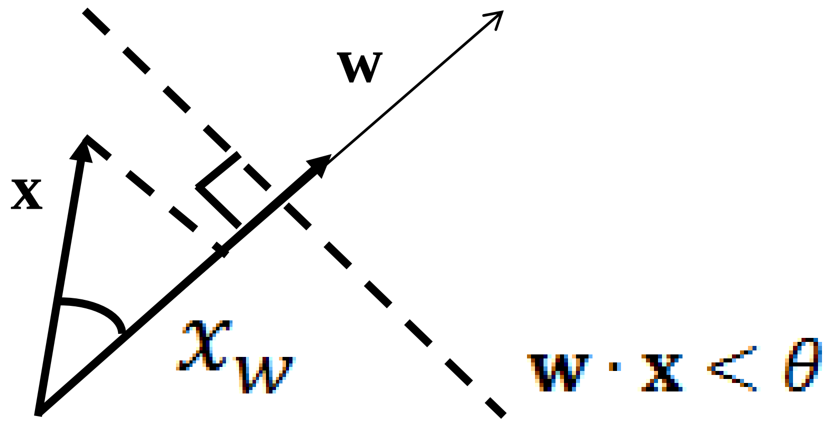
Veamos que pasa a cada lado de la línea (hiper-plano):

Supongamos que $x_w > \dfrac{\theta}{\|\mathbf{w}\|}$ entonces la proyección de **x** es más larga y, por tanto, **x** debe yacer en la región A.



Supongamos que $x_w < \frac{\theta}{\|\mathbf{w}\|}$ entonces la proyección de **x**

Es más corta y, por tanto, **x** debe yacer en la región B.



REGIÓN B, $y = 0$.

NOTA:
Los resultados son generales y son independientes de entradas a la TLU.

Resumen:

La activación de una TLU viene dada como el producto interno del vector de pesos y el vector de entrada: $a = \mathbf{w} \cdot \mathbf{x}$

Resumen:

La activación de una TLU viene dada como el producto interno del vector de pesos y el vector de entrada: $a = \mathbf{w} \cdot \mathbf{x}$

La relación $\mathbf{w} \cdot \mathbf{x} = \theta$ define un híper-plano en el espacio de patrones que es perpendicular al vector de pesos.

Resumen:

La activación de una TLU viene dada como el producto interno del vector de pesos y el vector de entrada: $a = \mathbf{w} \cdot \mathbf{x}$

La relación $\mathbf{w} \cdot \mathbf{x} = \theta$ define un híper-plano en el espacio de patrones que es perpendicular al vector de pesos.

A un lado de este hiper-plano yacen todos los patrones que son clasificados por la TLU como “1”, mientras que todos los patrones que yacen al otro lado del hiper-plano son clasificados como “0”.

Resumen:

La activación de una TLU viene dada como el producto interno del vector de pesos y el vector de entrada: $a = \mathbf{w} \cdot \mathbf{x}$

La relación $\mathbf{w} \cdot \mathbf{x} = \theta$ define un híper-plano en el espacio de patrones que es perpendicular al vector de pesos.

A un lado de este hiper-plano yacen todos los patrones que son clasificados por la TLU como “1”, mientras que todos los patrones que yacen al otro lado del hiper-plano son clasificados como “0”.

El hiper-plano es la superficie de decisión de la TLU. Ya que esta superficie es una versión n-dimensional de una línea recta, **la TLU es un clasificador lineal**.

Resumen:

La activación de una TLU viene dada como el producto interno del vector de pesos y el vector de entrada: $a = \mathbf{w} \cdot \mathbf{x}$

La relación $\mathbf{w} \cdot \mathbf{x} = \theta$ define un híper-plano en el espacio de patrones que es perpendicular al vector de pesos.

A un lado de este hiper-plano yacen todos los patrones que son clasificados por la TLU como “1”, mientras que todos los patrones que yacen al otro lado del hiper-plano son clasificados como “0”.

El hiper-plano es la superficie de decisión de la TLU. Ya que esta superficie es una versión n-dimensional de una línea recta, **la TLU es un clasificador lineal**.

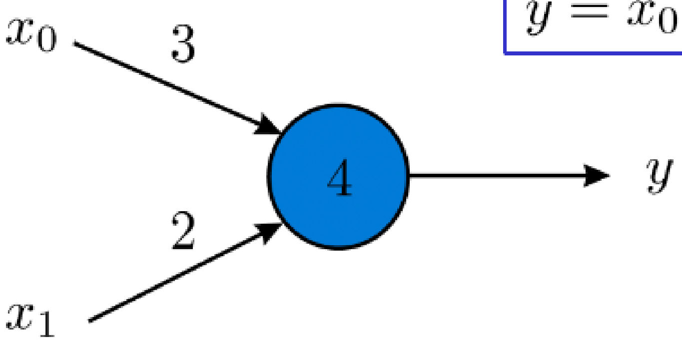
Si los patrones no pueden ser separados por un hiper-plano, entonces la TLU no podrá clasificarlos.

Representación de compuertas lógicas mediante TLUs:

Otros ejemplos:

Threshold Logic Unit for the Conjunction:

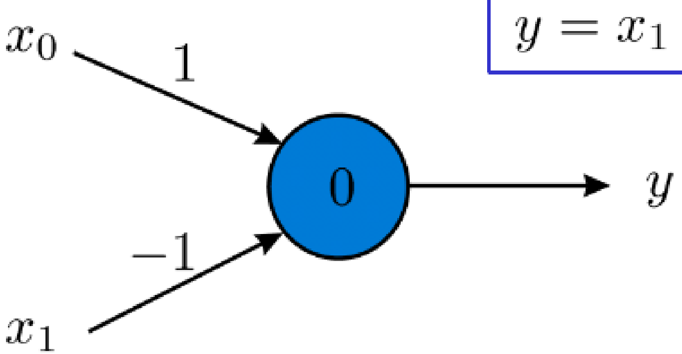
$$\boxed{y = x_0 \wedge x_1}$$



x_0	x_1	$3\,x_0 + 2\,x_1$	y
0	0	0	0
0	1	2	0
1	0	3	0
1	1	5	1

Threshold Logic Unit for the Implication:

$$\boxed{y = x_1 \longrightarrow x_0}$$

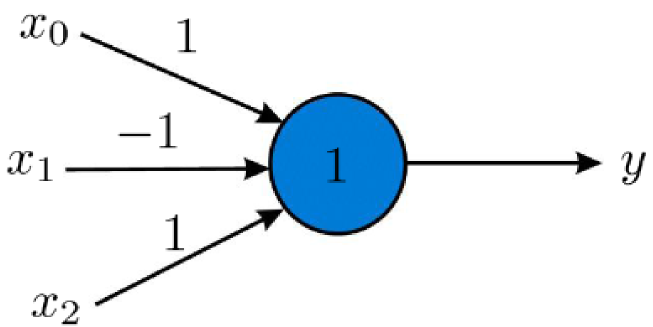


x_0	x_1	$x_0 - x_1$	y
0	0	0	1
0	1	-1	0
1	0	1	1
1	1	0	1

Otros ejemplos:

Further Sophisticated Example:

$$\boxed{y = (x_0 \wedge \overline{x}_1) \vee (\overline{x}_1 \wedge x_2) \vee (x_0 \wedge x_2)}$$

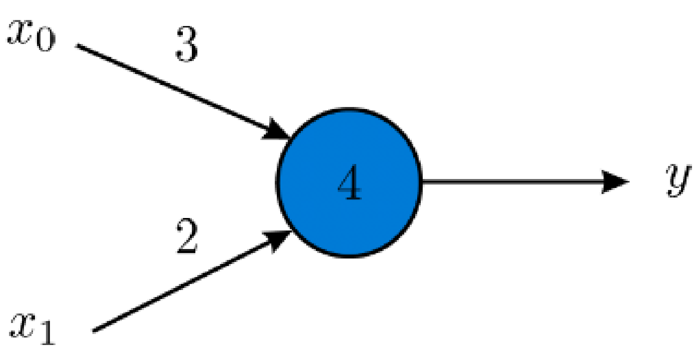


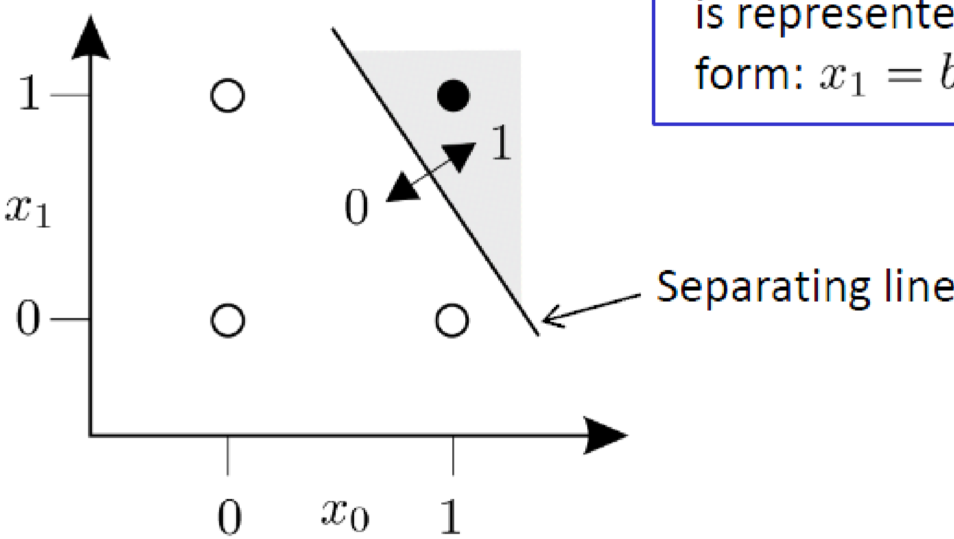
- Conventional implementation requires at least ***two OR logic gates*** each with two inputs.

x_0	x_1	x_2	$x_0 - x_1 + x_2$	y
0	0	0	0	0
0	0	1	1	1
0	1	0	-1	0
0	1	1	0	0
1	0	0	1	1
1	0	1	2	1
1	1	0	0	0
1	1	1	1	1

Interpretación geométrica (1/2):

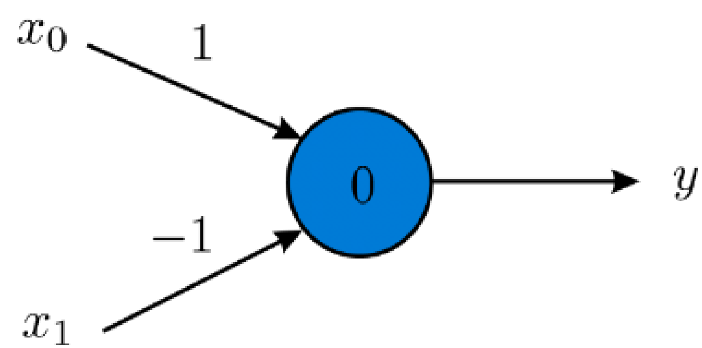
Conjunction:

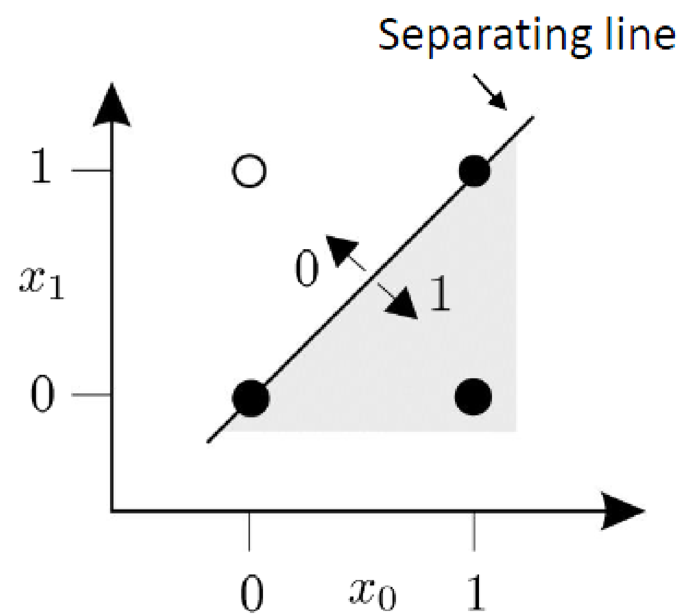




The ***separating line*** is represented in the form: $x_1 = b\,x_0 + c\,.$

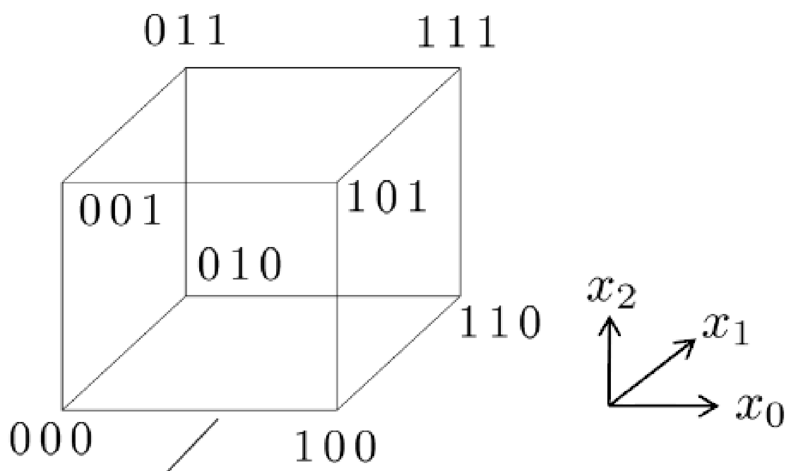
Implication:





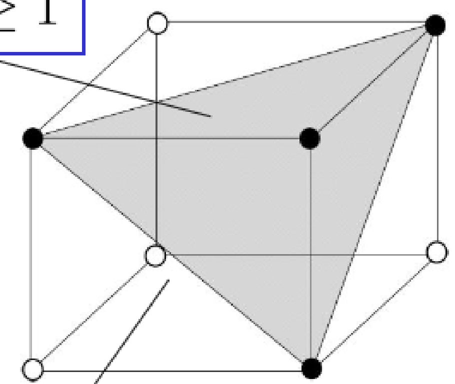
Interpretación geométrica (2/2):

Three Dimensional Unit Cube:

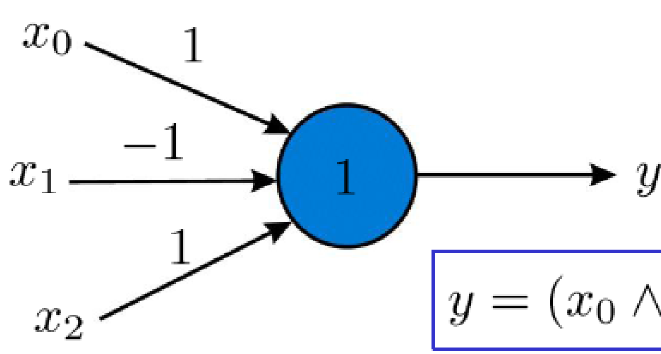


Logical values of three variables

$$\boxed{x_0 - x_1 + x_2 \geq 1}$$



Separation of the input space (separating plane)



$$\boxed{y = (x_0 \wedge \overline{x}_1) \vee (\overline{x}_1 \wedge x_2) \vee (x_0 \wedge x_2)}$$

Limitaciones de las TLUS:

Limitaciones:

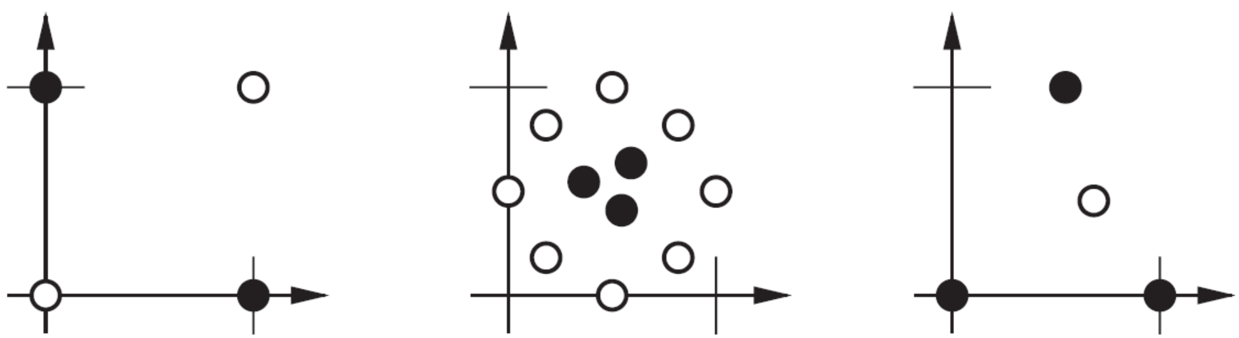
Un perceptrón no puede resolver problemas no linealmente separables.

Su frontera de decisión viene dada como: ${}_i\mathbf{w}^T\mathbf{p} + b_i = 0$

Ejemplo: Problema XOR:

$$\left\{\mathbf{p}_1 = \begin{bmatrix}0\\0\end{bmatrix}, t_1 = 0\right\} \left\{\mathbf{p}_2 = \begin{bmatrix}0\\1\end{bmatrix}, t_2 = 1\right\} \left\{\mathbf{p}_3 = \begin{bmatrix}1\\0\end{bmatrix}, t_3 = 1\right\} \left\{\mathbf{p}_4 = \begin{bmatrix}1\\1\end{bmatrix}, t_4 = 0\right\}.$$

Otros:



Se requiere de perceptrones combinados.

Problema XOR: En este caso la TLU debe satisfacer las siguientes 4 desigualdades:

$$
\begin{array}{lclcl}
x_0 = 0,\, x_1 = 0 & \longmapsto & w_0\, x_0 + w_1\, x_1 \;=\; 0 & \longrightarrow & 0 < \Phi\,, \\
x_0 = 1,\, x_1 = 0 & \longmapsto & w_0\, x_0 + w_1\, x_1 \;=\; w_0 & \longrightarrow & w_0 > \Phi\,, \\
x_0 = 0,\, x_1 = 1 & \longmapsto & w_0\, x_0 + w_1\, x_1 \;=\; w_1 & \longrightarrow & w_1 > \Phi\,, \\
x_0 = 1,\, x_1 = 1 & \longmapsto & w_0\, x_0 + w_1\, x_1 \;=\; w_0 + w_1 & \longrightarrow & w_0 + w_1 < \Phi\,.
\end{array}
$$

Problema XOR: En este caso la TLU debe satisfacer las siguientes 4 desigualdades:

$$
\begin{array}{llllll}
x_0 = 0,\, x_1 = 0 & \longmapsto & w_0\, x_0 + w_1\, x_1 & = & 0 & \longrightarrow \quad 0 < \Phi\,, \\
x_0 = 1,\, x_1 = 0 & \longmapsto & w_0\, x_0 + w_1\, x_1 & = & w_0 & \longrightarrow \quad w_0 > \Phi\,, \\
x_0 = 0,\, x_1 = 1 & \longmapsto & w_0\, x_0 + w_1\, x_1 & = & w_1 & \longrightarrow \quad w_1 > \Phi\,, \\
x_0 = 1,\, x_1 = 1 & \longmapsto & w_0\, x_0 + w_1\, x_1 & = & w_0 + w_1 & \longrightarrow \quad w_0 + w_1 < \Phi\,.
\end{array}
$$

Se puede ver que las primeras tres desigualdades son satisfechas para valores positivos de w_0 y w_1.

Sin embargo, la cuarta desigualdad ($w_0 + w_1 < \Phi$) NO PUEDE ser satisfecha para valores positivos de w_0 y w_1.

Problema XOR: En este caso la TLU debe satisfacer las siguientes 4 desigualdades:

$$
\begin{array}{llllll}
x_0 = 0,\ x_1 = 0 & \longmapsto & w_0\,x_0 + w_1\,x_1 & = & 0 & \longrightarrow & 0 < \Phi\,, \\
x_0 = 1,\ x_1 = 0 & \longmapsto & w_0\,x_0 + w_1\,x_1 & = & w_0 & \longrightarrow & w_0 > \Phi\,, \\
x_0 = 0,\ x_1 = 1 & \longmapsto & w_0\,x_0 + w_1\,x_1 & = & w_1 & \longrightarrow & w_1 > \Phi\,, \\
x_0 = 1,\ x_1 = 1 & \longmapsto & w_0\,x_0 + w_1\,x_1 & = & w_0 + w_1 & \longrightarrow & w_0 + w_1 < \Phi\,.
\end{array}
$$

Se puede ver que las primeras tres desigualdades son satisfechas para valores positivos de w_0 y w_1.

Sin embargo, la cuarta desigualdad ($w_0 + w_1 < \Phi$) NO PUEDE ser satisfecha para valores positivos de w_0 y w_1.

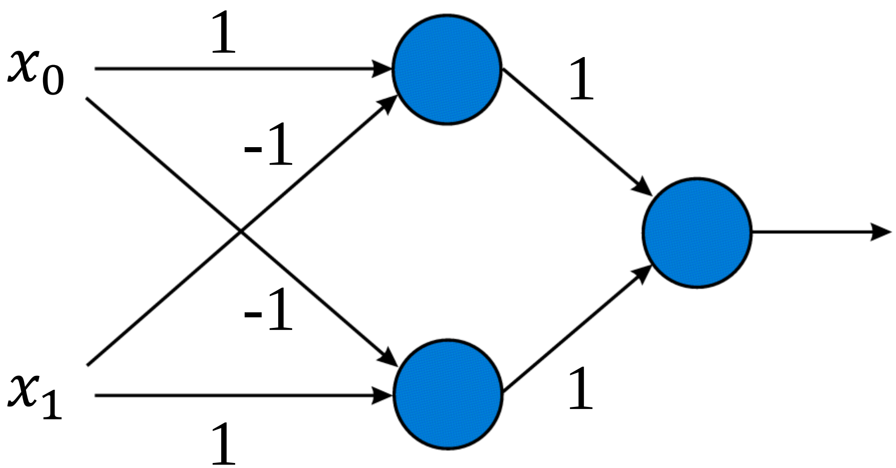
Luego, entonces, el problema XOR no puede ser resuelto con una sola TLU.

Solución al problema del XOR mediante TLUS:

Solución ideal mediante tres neuronas TLU:

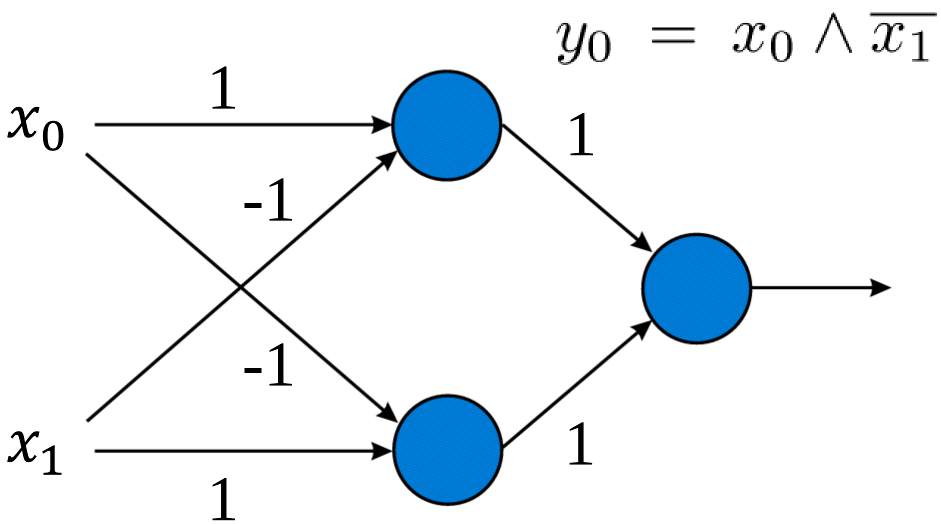
Solución ideal mediante tres neuronas TLU:

Por descomposición lógica:



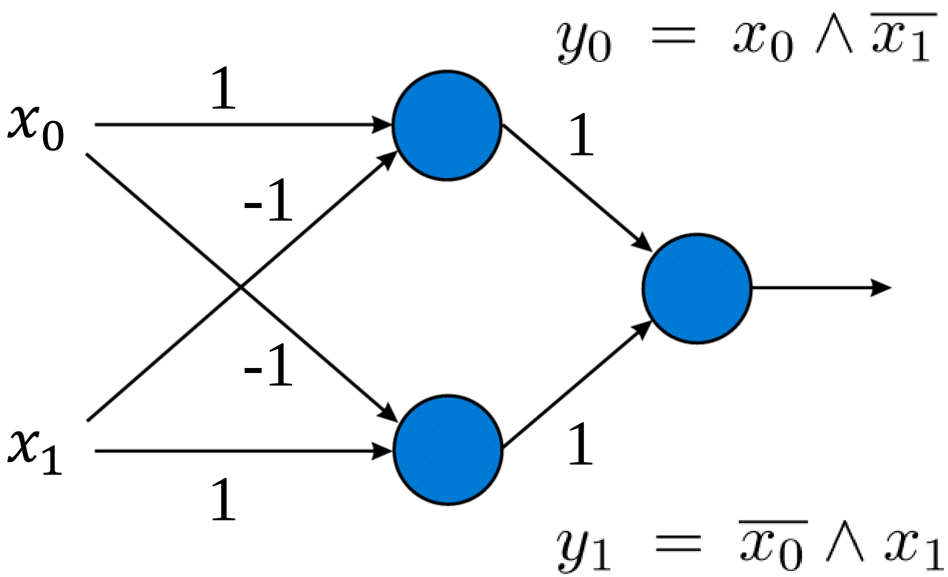
Solución ideal mediante tres neuronas TLU:

Por descomposición lógica:



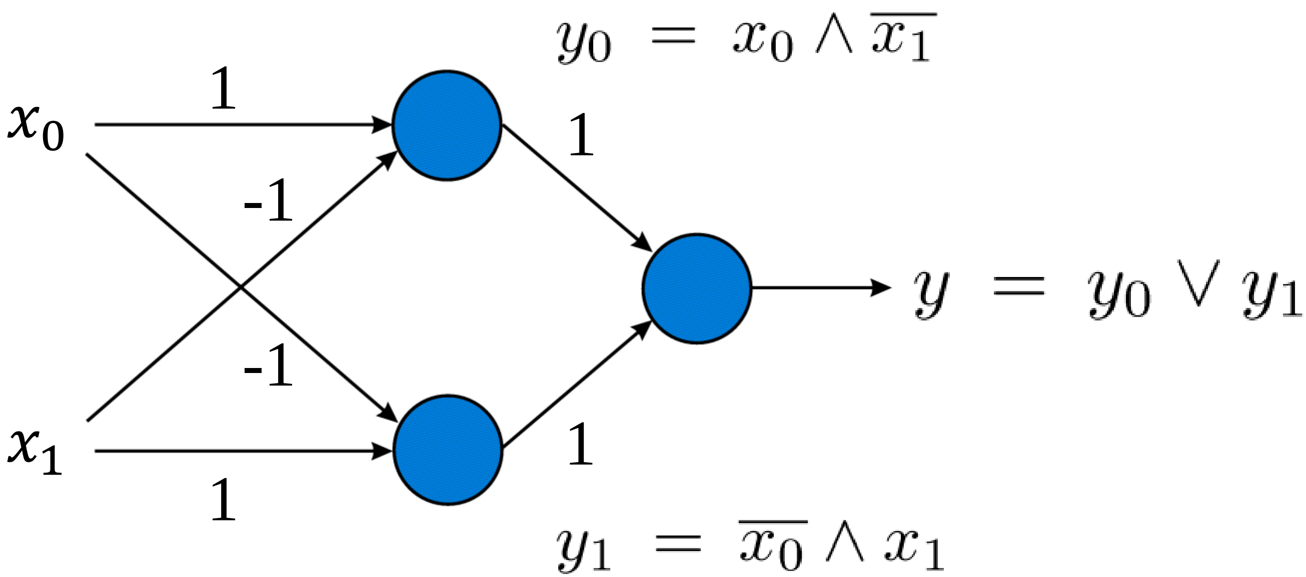
Solución ideal mediante tres neuronas TLU:

Por descomposición lógica:



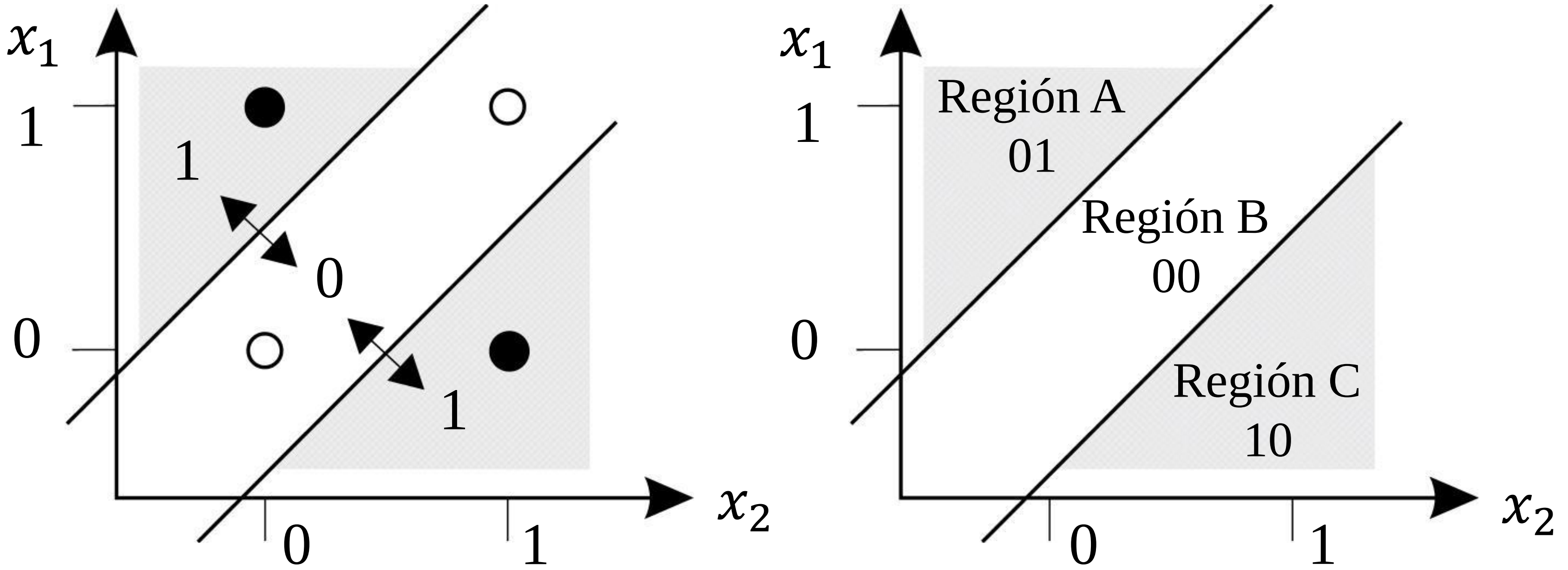
Solución ideal mediante tres neuronas TLU:

Por descomposición lógica:



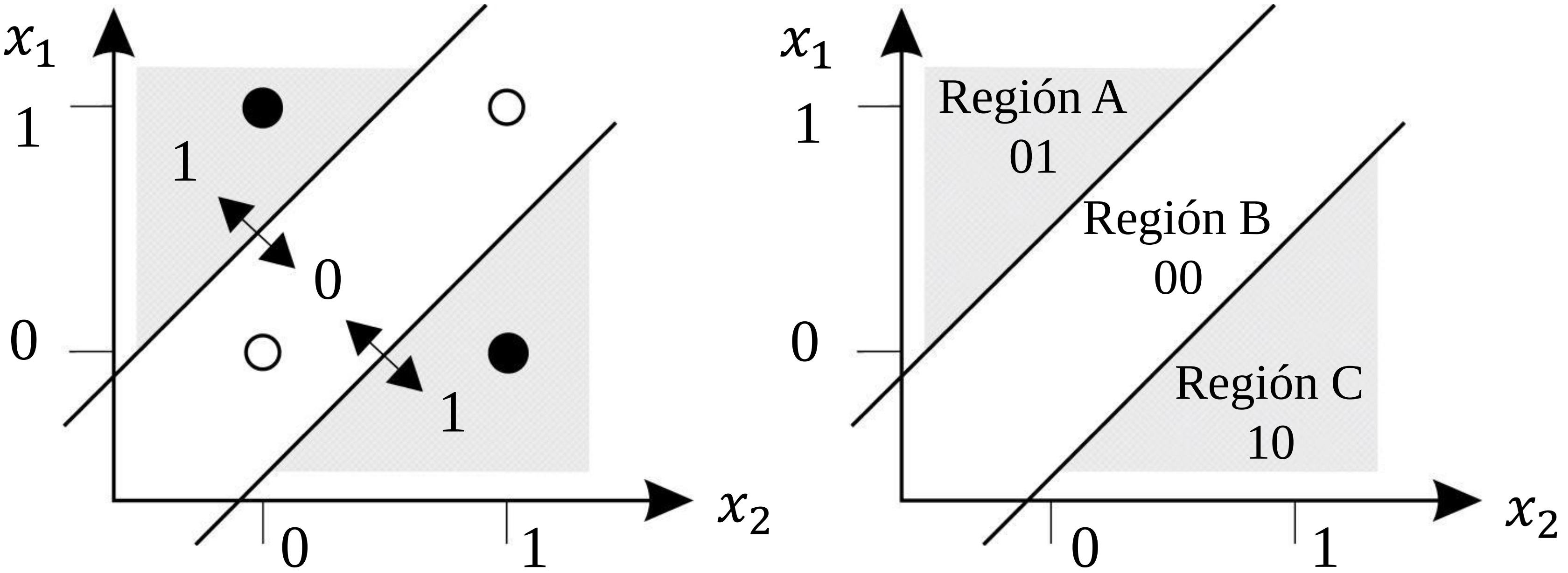
En resumen, la red consta de tres neuronas que en conjunto llevan a cabo la función lógica: $(x_0 \wedge \overline{x_1}) \vee (\overline{x_0} \wedge x_1)$.

Interpretación geométrica:



Las dos neuronas de la primera capa producen el etiquetado de la región en la cual la entrada se localiza.

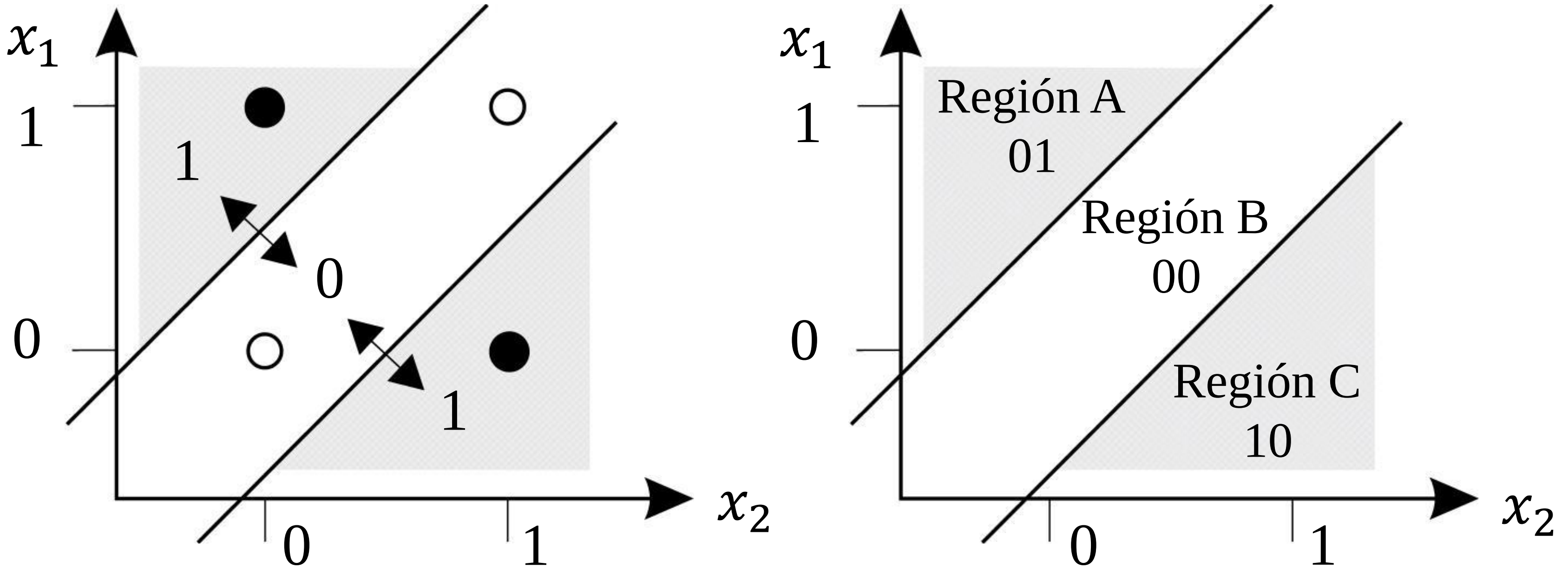
Interpretación geométrica:



Las dos neuronas de la primera capa producen el etiquetado de la región en la cual la entrada se localiza.

La unidad de salida decodifica las etiquetas de las tres regiones (en este caso se aplica una función OR).

Interpretación geométrica:

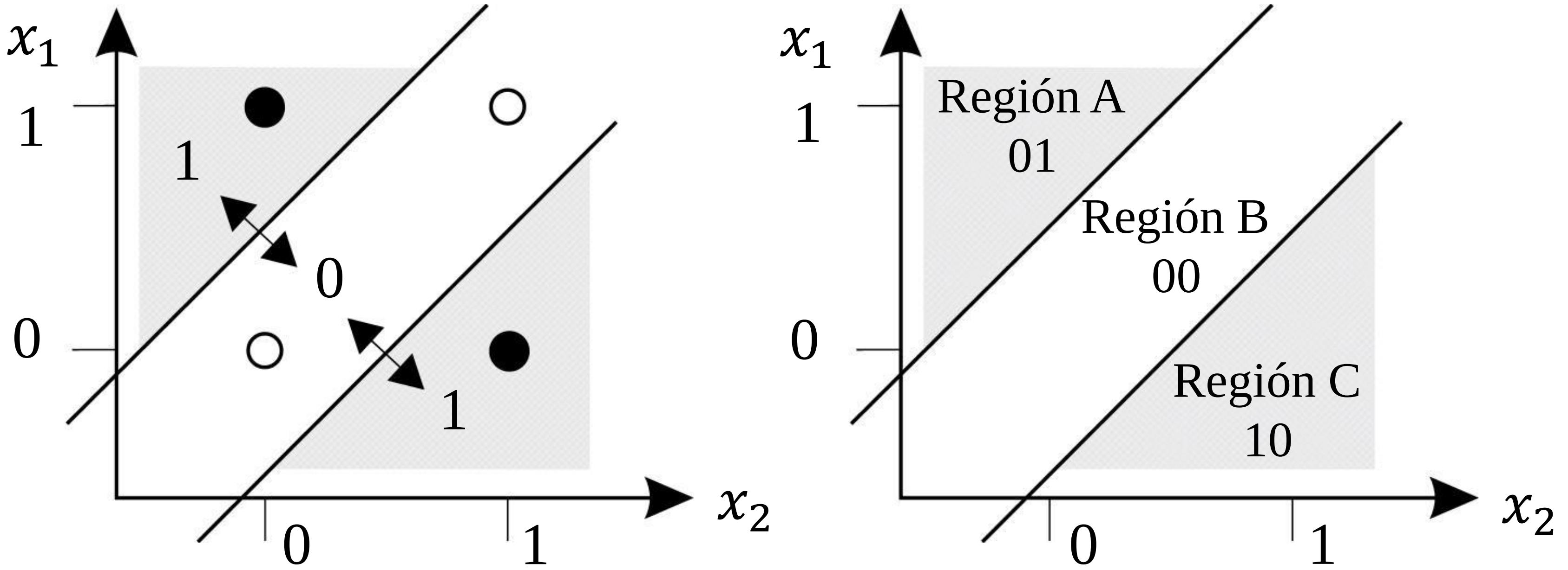


Las dos neuronas de la primera capa producen el etiquetado de la región en la cual la entrada se localiza.

La unidad de salida decodifica las etiquetas de las tres regiones (en este caso se aplica una función OR).

La primera capa mapea el vector de entrada a la segunda etapa, llamada de clasificación.

Interpretación geométrica:



Las dos neuronas de la primera capa producen el etiquetado de la región en la cual la entrada se localiza.

La unidad de salida decodifica las etiquetas de las tres regiones (en este caso se aplica una función OR).

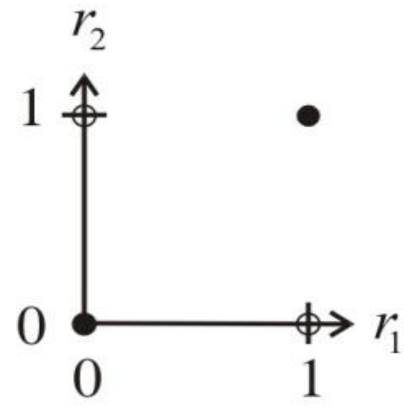
La primera capa mapea el vector de entrada a la segunda etapa, llamada de clasificación.

La última capa decodifica esta clasificación y entrega la salida.

Función de la capa interna de la RNA de TLUs para el problema del XOR:

El OR exclusivo:

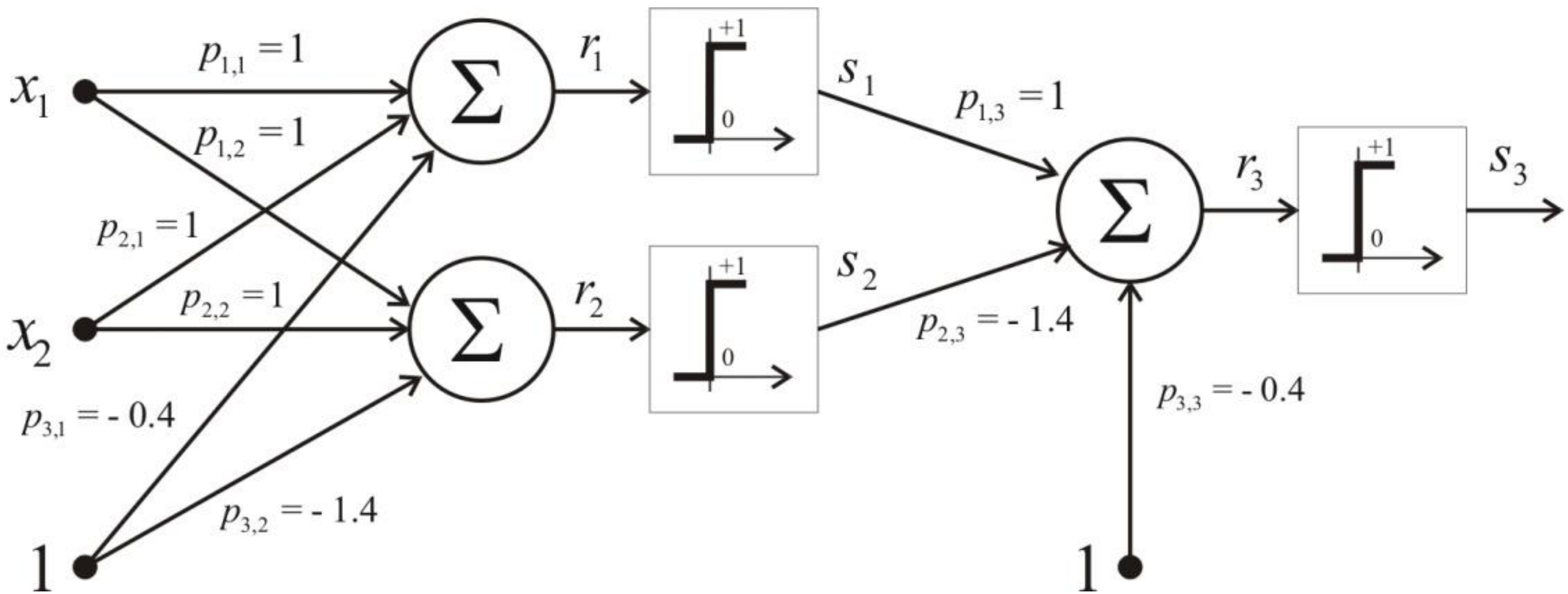
x_1	x_2	r_1	r_2	s_1	s_2	r_3	s_3
0	0	-0.4	-1.4	0	0	-0.4	0
0	1	0.6	-0.4	1	0	0.6	1
1	0	0.6	-0.4	1	0	0.6	1
1	1	1.6	0.6	1	1	-0.8	0

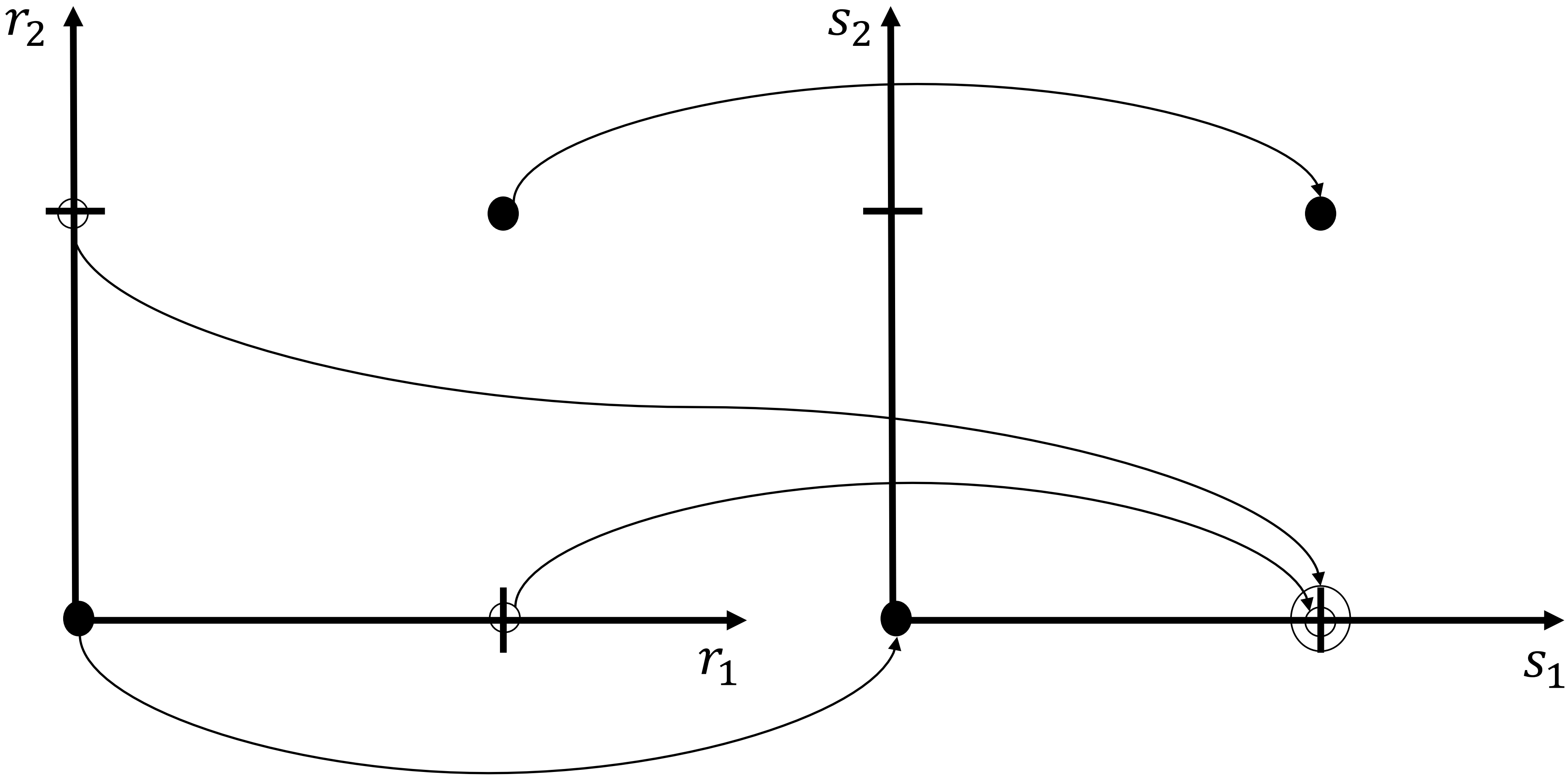


(a)

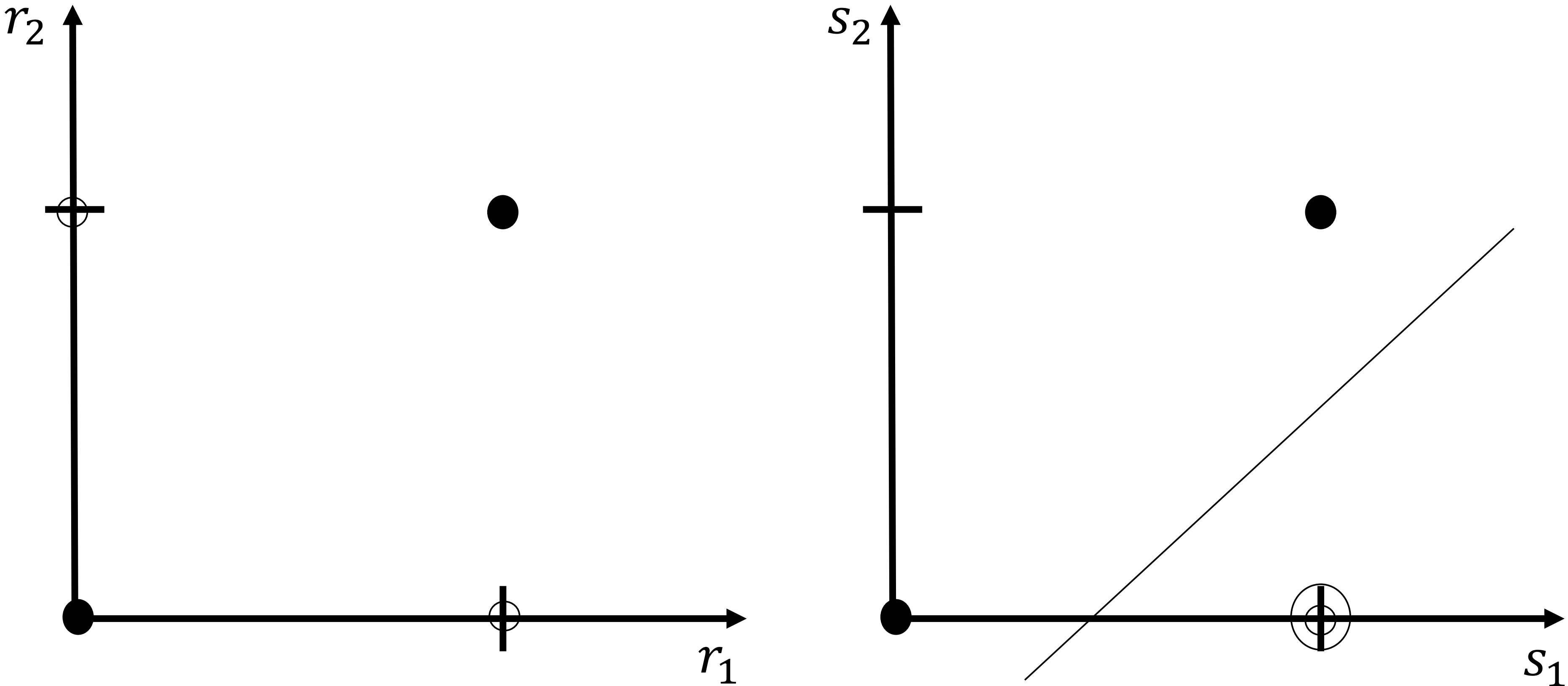
r_1	r_2	Clase de pertenencia
0	0	C^2
0	1	C^1
1	0	C^1
1	1	C^2

(b)





x_1	x_2	r_1	r_2	r_2	s_1	s_2	r_1
0	0	-0.4	-1.4	0	0	-0.4	0
0	1	0.6	-0.4	1	0	0.6	1
1	0	0.6	-0.4	1	0	0.6	1
1	1	1.6	0.6	1	1	-0.8	0



r_2
r_1
s_2
s_1