En matemáticas

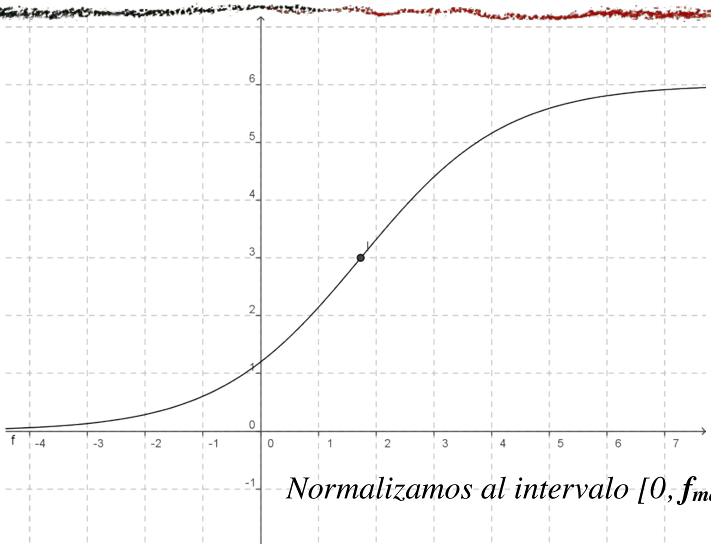
$$y = \begin{cases} 1, & \text{si } a \geq \theta, \\ 0, & \text{si } a < \theta. \end{cases} \quad \text{donde } a = \sum_{i=1}^{n} w_i x_i.$$

Probar: tolerancia a errores/ruido en las entradas

Sequencias de picos

- *Frequencia de generación de picos* $\boldsymbol{f}$
- *Varía entre cero y un máximo (fisiológico)* $\boldsymbol{f_{max}}$
- *Graduación contínua del señal no es posible si usamos simplemente la codificación binaria*

Función sigmoidal



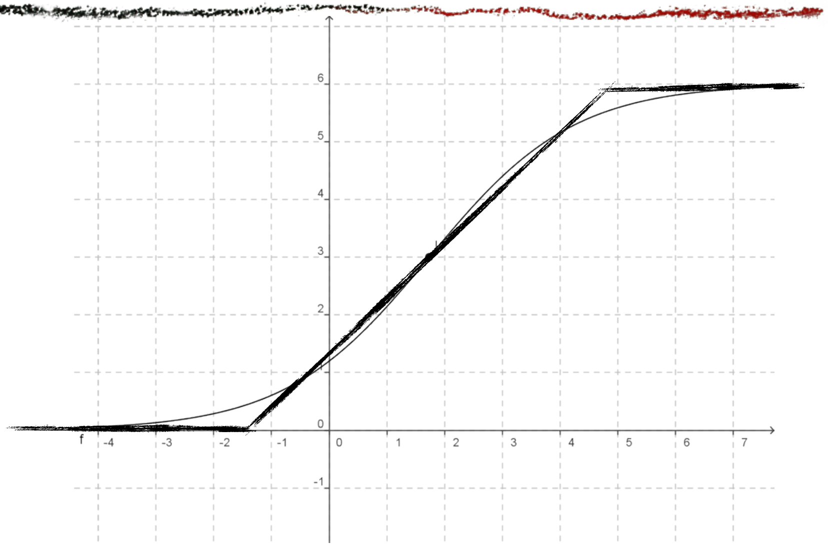
Normalizamos al intervalo [0, $\boldsymbol{f_{max}}$*]*

Formula

$$y = \sigma(a) \equiv \frac{1}{1 + e^{-(a-\theta)/\rho}}$$

Parámetro que determina la forma de la curva (menor valor, mayor pendiente)

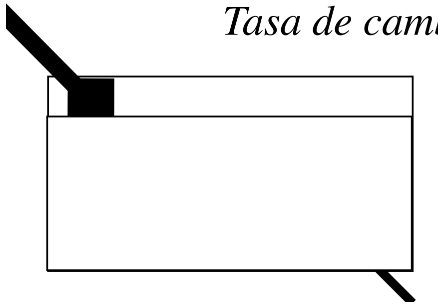
Linearización por partes



Aspecto temporal

$$\frac{da}{dt} = -\alpha a + \beta s, \text{ donde ahora } s = \sum_{i=1}^{n} w_i x_i$$

Tasa de cambio instantáneo



Integrador con fuga

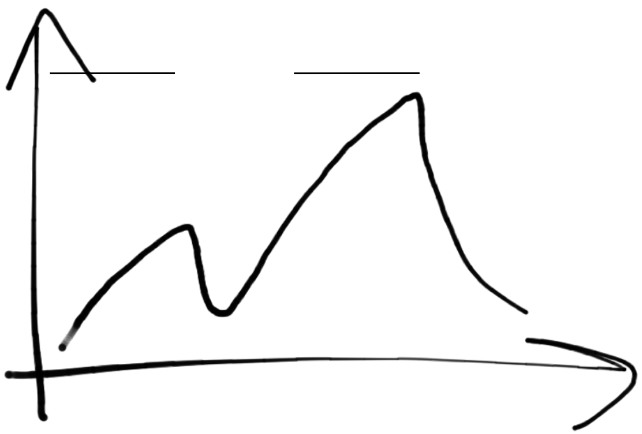
Efecto de "memoria"

- *Ahora cuando se apaga la entrada, la salida sigue por un tiempo antes de tocar cero*
- *De la misma forma, al prender la entrada, la salida tarda en llegar a su nivel máximo*

Equilibrio

$$a_{\mathrm{eqm}} = \left(\frac{\beta}{\alpha}\right) s$$

Respuesta a dos impulsos



Espacio de patrones de entrada

- ***n*** *entradas,* ***n*** *dimensiones*
- *representamos las entradas como componentes de un vector*

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

Separabilidad lineal

- *En dos dimensiones, tenemos una línea* $w_1 x_1 + w_2 x_2 = \theta$
  - *Esta separa las activadas de las ignoradas*
  - *La distancia entre un punto y la línea se mide por* Δx_i
- *En* ***n*** *dimensiones, será un híperplano*
  - *… que hará exactamente lo mismo*
  - *Se llama* ***plano de decisión***

Aritmética vectorial

- *Sumar, restar y multiplicar por componentes*
- *Largo/norma:* $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}$
- *Producto interno (escalar):*

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

Proyección vectorial

- *Cuánto de un véctor* ***v*** *corresponde a la dirección de otro véctor* ***w***

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

Activación de un TLU: híperplano

$$\begin{aligned} a &= \mathbf{w} \cdot \mathbf{x} \\ \theta &= \mathbf{w} \cdot \mathbf{x} \\ x_w &= \frac{\theta}{\|\mathbf{w}\|} \end{aligned}$$

Ahora válido tal cual en n dimensiones.

Clasificación con el superficie de decisión

$$y = \begin{cases} 1, & \text{si } x_w \geq \dfrac{\theta}{\|\mathbf{w}\|}, \\ 0, & \text{si } x_w < \dfrac{\theta}{\|\mathbf{w}\|}. \end{cases}$$

Entrenamiento supervisado

- *Propósito: determinar el superficie de decisión adecuado.*
- *Típicamente un proceso iterativo:*
  - *Se presentan ejemplos preclasificados* $\{\mathbf{x}_i\}$.
  - *Se realizan ajustes pequeños para acercar las respuestas a las deseadas* $\{\mathbf{t}_i\}$.

Cambio notacional

- *Observamos que $\mathbf{w} \cdot \mathbf{x} - \theta = 0$.*
- *Reescribimos **w** con n+1 elementos, agregando θ al final.*
- *Al final de **x** agregamos un -1 constante.*
- *Ahora con estas modificaciones tenemos $\mathbf{w} \cdot \mathbf{x} \geq 0 \Rightarrow y = 1$, y en otro caso $y = 0$.*

Regla de entrenamiento de perceptrón

- *Supongamos que hemos presentado una muestra **v** al TLU y nos ha producido $y = 0$ aunque debería haber dado $y = 1 = t$.*
  - *Habrá que rotar **w** para acercar el TLU a la clasificación deseada:*
    - $\mathbf{w'} = \mathbf{w} + \alpha \mathbf{v}$;
    - *$\alpha \in (0, 1)$ es una constante que se llama la tasa de aprendizaje.*
- *Si ocurre lo contrario (el TLU produce $y = 1$ aunque se esperaba que dara $y = 0 = t$), rotamos en la dirección contraria:*
  - $\mathbf{w'} = \mathbf{w} - \alpha \mathbf{v}$.
- *Combinando esto en términos de t, obtenemos una sola regla:*
  - $\mathbf{w'} = \mathbf{w} + \alpha(t - y)\mathbf{v}$.

Teorema de convergencia de perceptrón

*Si dos clases de véctores son linealmente separables, la aplicación iterativa de la regla de entrenamiento de perceptrón a estas clases eventualmente producirá un véctor de ponderación **w** cuyo híperplano de decisión separa las clases.*

Rosenblatt, 1962.