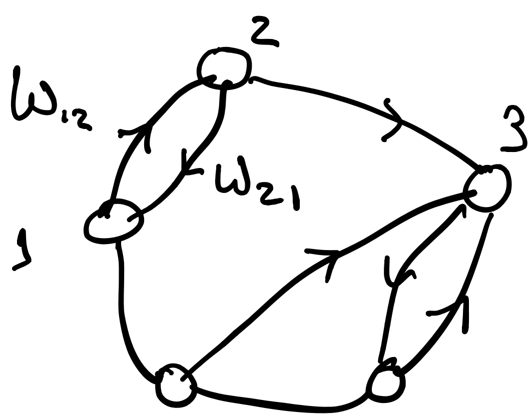
12/10/18

Redes Recurrentes

- No hay una organización jerárquica de las neuronas
- Todas las neuronas pueden procesar información en todas las etapas
- Las neuronas son del tipo Mc Colloch and Pitts
- Pueden loops e incluso puede haber intercambio recíproco entre dos neuronas



- La red tiene N neuronas

y a cada una le asignamos una variable $S_i(t)$. El estado de la red es un vector de N componentes

$$\bar{S}(t) = (S_1(t), S_2(t), \text{———}, S_N(t))$$

- Tenemos una matriz de acoplamiento sináptico

$$W = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & & W_{2N} \\ & & & \\ W_{N1} & W_{N2} & & W_{NN} \end{pmatrix}$$

- La red debe asociar estados iniciales $\bar{S}(t=0)$ con estados finales $\bar{S}(t=\infty)$
- $S(t=\infty)$ puede ser un punto fijo, una oscilación periódica, o un atractor caótico

- $0 \to 1 \to 2 \to 2 \to \cdots \; 2 \to$

  punto fijo

- $$\begin{tikzcd} 0 \arrow[r] & 1 \arrow[r] & 2 \arrow[r] & 3 \arrow[d] \\ & & 5 \arrow[u] & 4 \arrow[l] \end{tikzcd}$$

La red cambia de estado cuando actualizo una vez cada neurona.

Δt : es el tiempo en que actualizo N neuronas

Vamos a suponer que

$$S_i(t) = \begin{cases} +1 & \text{si dispara} \\ -1 & \text{si no dispara} \end{cases}$$

Como son neuronas McCulloch & Pitts

- calcula

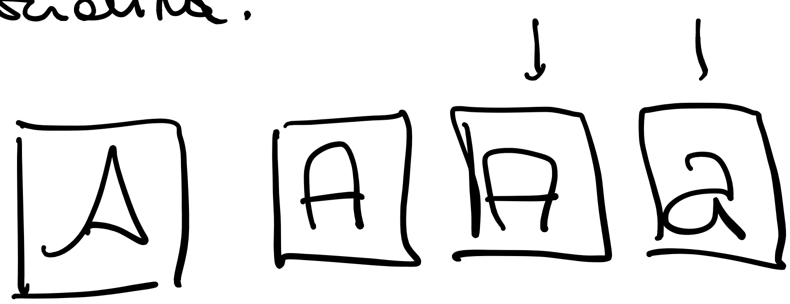
$$h_i(t) = \sum_{j=1}^{N} w_{ij} S_j(t)$$

- Compara $h_i(t)$ con el umbral θ

$$S_i(t+\Delta t) = \text{signo}\left(h_i(t) - \theta\right)$$

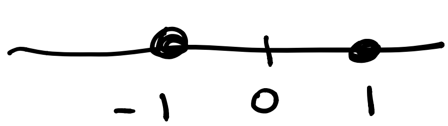
$$\left.\begin{array}{lll} w_{ij} > 0 & & \text{exhitatoria} \\ w_{ij} = 0 & & \text{ausente} \\ w_{ij} < 0 & & \text{inhibitoria} \end{array}\right\} \text{sinapsis}$$

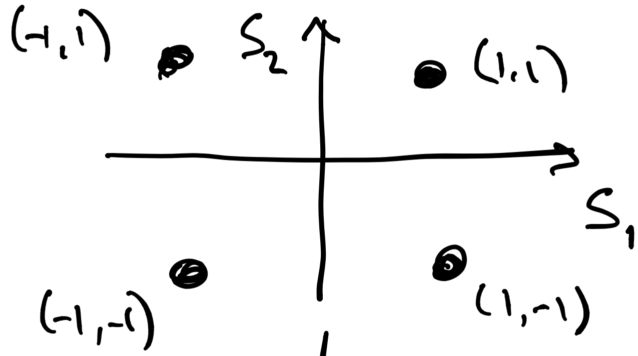
Son especialmente útiles para memoria asociativa.

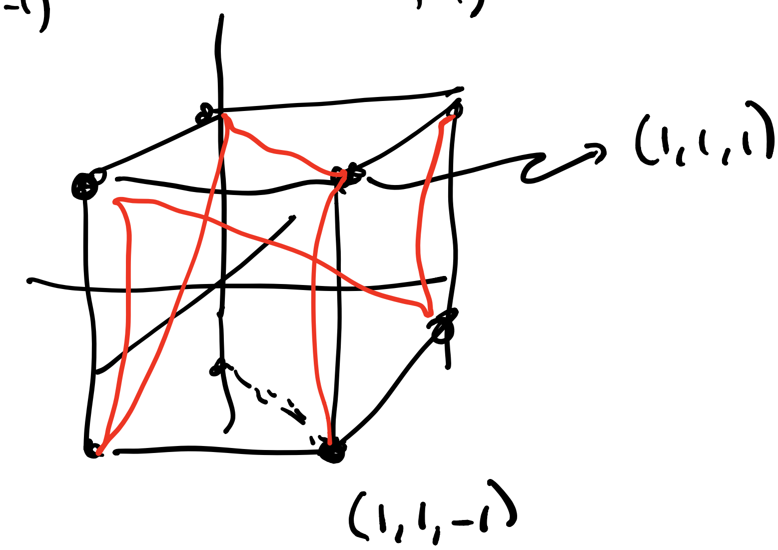


La red tiene 2^N posibles estados.

Los estados se pueden representar como un hipercubo







Memoria asociativa

De entre los 2^N posibles configuraciones $\vec{S}$ de la red, queremos elegir p configuraciones $\left\{\xi_i^{\mu}\right\}^{\mu=1,\dots,p}_{i=1,\dots,N}$ que sean atractores de la dinámica.

En concreto, buscamos la matriz W y los umbrales θ_i que lo logran.

$$\bar{\xi}^1 = (\xi_1^1, \xi_2^1, \ldots, \xi_N^1)$$

Si $S(t) = \bar{\xi}^1$, quiero que se quede allí $\xi_i^1 = \pm 1 \quad \forall i$

lo mismo para μ

$$\bar{\xi}^\mu = (\xi_1^\mu, \xi_2^\mu, \ldots, \xi_N^\mu)$$

y $\xi_i^\mu = \pm 1 \quad \forall i$ y $\forall \mu$.

En 1982 un físico norteamericano propuso el problema y dio una solución. Se llamó John Hopfield.

El modelo de Hopfield

Tenemos N neuronas $S_i = \pm 1$

Tenemos N umbrales nulos $\theta_i = 0 \ \forall i$

Tenemos una matriz W_{NN} simétrica

$$w_{ij} = w_{ji}$$

con diagonal nula

$$w_{ii} = 0 \quad \forall i$$

- La actualización es secuencial, o sea, actualiza una por vez

$$S_i(t+\Delta t) = \text{signo}\left(\sum_{j\neq i} w_{ij} S_j(t)\right)$$

$$\Delta t = \frac{1}{N}$$

<u>Caso 1</u> . Tengo $P=1$ y además

$$\xi_i^1 = +1 \quad \forall i$$

Entonces puedo hacer $w_{ij} = 1 \ \forall i \neq j$

$w_{ii} = 0 \ \forall i$

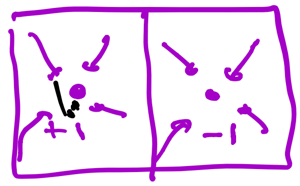
$$S_i(t+\Delta t) = \text{signo}\left(\sum_{j\neq i} 1\, S_j(t)\right)$$

Si en t la mayoría son $+1$

$$S_i(t+\Delta t) = 1$$

Si en t la mayoría son -1

$$S_i(t+\Delta t) = -1$$



• <u>Caso 2</u> $P=1$ pero ξ_i^1 es una configuración aleatoria

$$\xi_i^1 = \begin{cases} +1 & \text{con prob. } +\frac{1}{2} \\ -1 & \text{'' \quad ''} \quad +\frac{1}{2} \end{cases}$$

$$\boxed{\begin{aligned} w_{ij} &= \frac{\xi_i^1 \xi_j^1}{N} \\ w_{ii} &= 0 \end{aligned}}$$

$$S_i(t+\Delta t) = \text{signo}\left(\sum_{j\neq i} w_{ij}\, S_j(t)\right)$$

$$= \text{signo}\left(\sum_{j\neq i} \frac{1}{N}\, \xi_i^1 \xi_j^1\, S_j(0)\right)$$

Supongamos que en n neuronas no coincide $S(0)$ con ξ^1 y en $N-n$

si coinciden

$$S_i(t+\Delta t) = \text{signo}\left(\xi_i^1 \frac{1}{N}\sum_{\substack{j\neq i}}^{N} \xi_j^1 S_j(0)\right)$$

$$= \text{signo}\left(\xi_i^1 \frac{1}{N}\left[-n + \overset{\sim N\ldots}{\left[N-1-n\right]}\right]\right)$$

Suponemos $N \gg 1$

$$= \text{signo}(\xi_i^1)\, \underbrace{\text{signo}\left(\frac{1}{N}\right)}_{=1} \cdot \text{signo}[N-2n]$$

$$= \text{signo}(\xi_i^1)\, \text{signo}[N-2n]$$

si $n > \frac{1}{2}N$

$$S_i(t+\Delta t) = -\text{signo}(\xi_i^1) = -\xi_i^1$$

$n < \frac{1}{2}N$

$$S_i(t+\Delta t) = \text{signo}(\xi_i^1) = \xi_i^1$$

Moraleja: $\quad w_{ij} = \frac{1}{2}\xi_i^1 \xi_j^1$

hace 2 atractores: ξ^1 y $-\xi^1$

la red ~~lleva~~ a una conf. inicial

al atractor más parecido

<u>Caso 3</u>: se puede ver que si quiero tener dos atractores aleatorios ξ^1 y ξ^2 puedo usar

$$W_{ij} = \frac{\xi_i^1 \xi_j^1 + \xi_i^2 \xi_j^2}{N}$$

Ahora tenemos 4 atractores

$$\xi^1, \ -\xi^1, \ \xi^2, \ -\xi^2$$

- Es una regla local y hebbiana
- Supuso que si tenemos p configuraciones debemos hacer

$$W_{ij} = \frac{1}{N}\left(\xi_i^1 \xi_j^1 + \xi_i^2 \xi_j^2 + \cdots + \xi_i^p \xi_j^p\right)$$

$$\boxed{\begin{aligned} W_{ij} &= \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \\ W_{ii} &= 0 \end{aligned}}$$

Regla de Hopfield

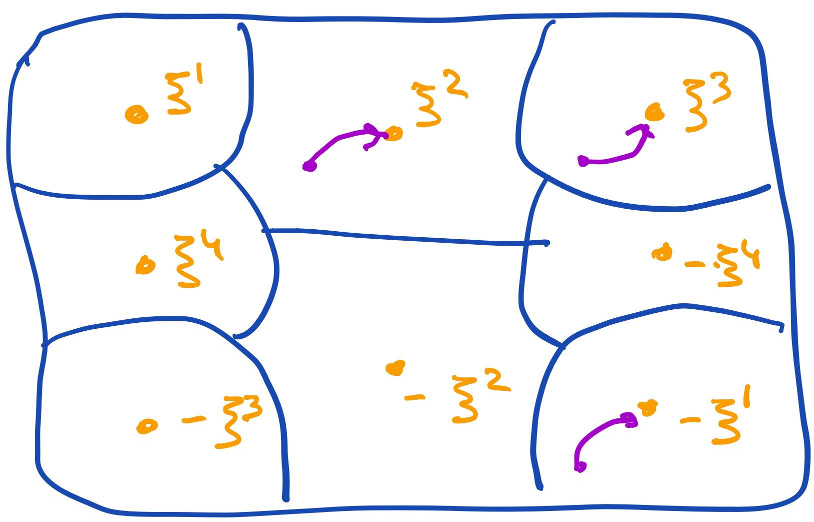
Una forma canónica de resolver esto es la siguiente.

$$d = \frac{1}{N} \sum_i^N \frac{(S_i - \xi_i)^2}{4}$$

Para cada $\mu = 1, \dots p$ calculamos las distancias

$$d^\mu = \frac{1}{N} \sum_i \frac{(S_i(0) - \xi_i^\mu)^2}{4}$$

y asociamos $\bar{S}(0)$ con el μ tal que d^μ sea mínimo.



hiper cubo de los 2^N posibles estados

Análisis de estabilidad

Supongamos que $\vec{S}(0) = \vec{\xi}^{\nu}$

el estado inicial es la memoria $\vec{\xi}^{\nu}$

$\vec{\xi}^{\nu}$: memoria ν

$$h_i^{(\nu)} = \sum_j w_{ij} S(0) = \sum_j w_{ij} \xi_j^{\nu}$$

$$= \sum_j \left[\frac{1}{N} \sum_{\mu=1}^{p} \xi_i^{\mu} \xi_j^{\mu} \right] \xi_j^{\nu}$$

$$= \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^{\mu} \sum_{j=1} \xi_j^{\mu} \xi_j^{\nu}$$

$$= \frac{1}{N} \left[\xi_i^{\nu} \underbrace{\sum_{j=1}^{N} \overbrace{\xi_j^{\nu} \xi_j^{\nu}}^{=1}}_{N} + \sum_{\mu \neq \nu} \xi_i^{\mu} \sum_{j=1}^{N} \xi_j^{\mu} \xi_j^{\nu} \right]$$

$$h_i^{(\nu)} = \frac{1}{N} \left[N \xi_i^{\nu} + \sum_{\mu \neq \nu} \xi_i^{\mu} \sum_{j=1}^{N} \xi_j^{\mu} \xi_j^{\nu} \right]$$

$$h_i^{(\nu)} = \xi_i^\nu + \frac{1}{N} \underbrace{\sum_{\mu \neq \nu}}_{\substack{P-1 \\ \text{términos}}} \underbrace{\sum_{\substack{j=1 \\ j \neq i}}}_{\substack{N-1 \\ \text{términos}}} \underbrace{\xi_i^\mu \xi_j^\mu \xi_j^\nu}_{\pm 1}$$

$$S_i = \text{signo}\left(h_i^{(\nu)}\right) = \text{signo}\left(\xi_i^\nu + \frac{1}{N} \sum_{\mu \neq \nu} \sum_{i \neq j} \xi_i^\mu \xi_j^\mu \xi_j^\nu\right)$$

$$(\xi_i^\nu)^2 = 1$$

$$= \text{signo}\left(\xi_i^\nu \left[1 + \frac{\xi_i^\nu}{N} \sum_{\mu \neq \nu} \sum_{j \neq i} \xi_i^\mu \xi_j^\mu \xi_j^\nu\right]\right)$$

$$= \text{signo}\left(\xi_i^\nu \left(1 + C_i^\nu\right)\right)$$

$$C_i^\nu = \frac{1}{N} \sum_{\mu \neq \nu} \sum_{j \neq i} \xi_i^\nu \xi_i^\mu \xi_j^\mu \xi_j^\nu$$

la distrib. de prob. tiende a una normal de media 0 y de varianza $\sigma = \frac{1}{N}\sqrt{(N-1)(P-1)}$

si $N \gg 1$ y $P \gg 1$

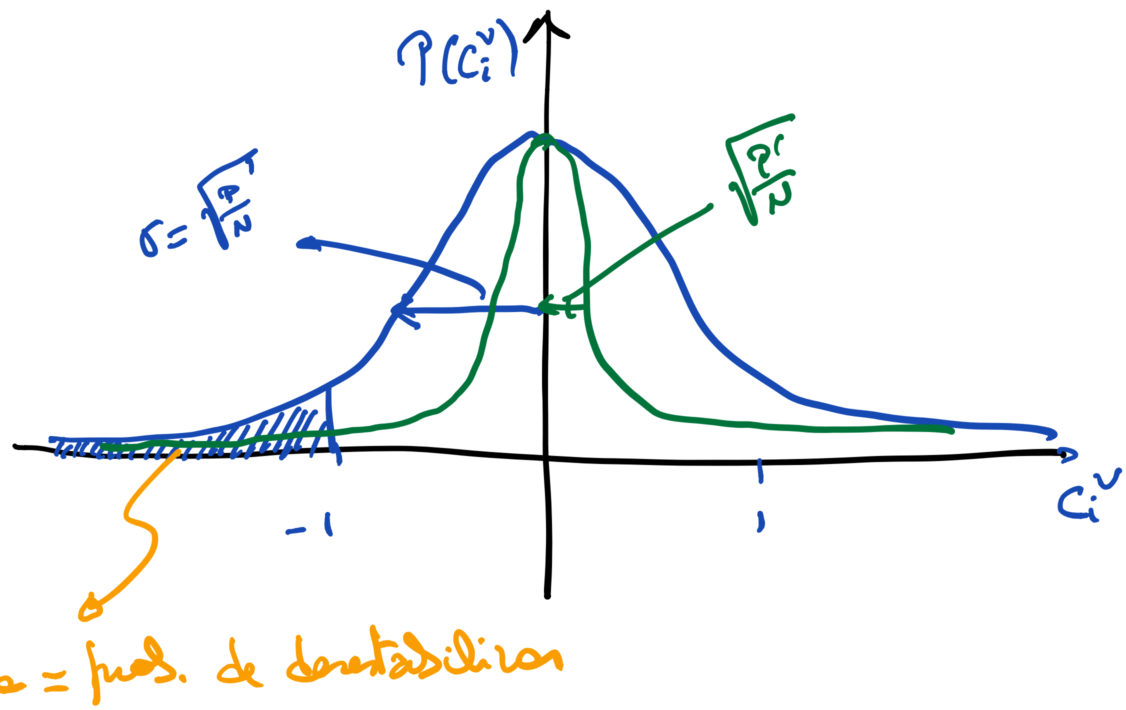
$$\sigma \simeq \frac{1}{N}\sqrt{NP} = \sqrt{\frac{P}{N}}$$

$$P(C_i^\nu) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(C_i^\nu - \langle C_i^\nu \rangle)^2}{2\sigma^2}}, \quad \langle C_i^\nu \rangle = 0$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(C_i^\nu)^2}{2\sigma^2}}$$

Para que se desestabilice la neurona i

$$C_i^\nu < -1$$

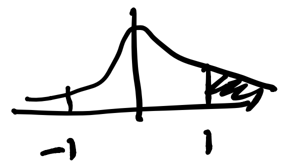


Queremos calcular la probabilidad de que una neurona se desestabilice

$$P_{error} = P(C_i^{\nu} < -1) = P(C_i^{\nu} > +1)$$ por simetría

$$P_{error} = \int_1^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx$$

$$= 1 - \int_{-\infty}^{1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx$$



$$= 1 - \frac{1}{2} - \int_0^{1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx$$

$$u^2 = \frac{x^2}{2\sigma^2} \qquad du = \frac{dx}{\sqrt{2}\sigma}$$

$$P_{error} = \frac{1}{2} - \int_0^{\frac{1}{\sqrt{2\sigma^2}}} e^{-u^2} \frac{du}{\sqrt{\pi}}$$

$$= \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{1}{\sqrt{2\sigma^2}}\right)$$

$$= \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{1}{\sqrt{2\sigma^2}}\right)\right]$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^{x} e^{-x^2} dx$$

Dado un error fijo, o una tolerancia P_{error} puedo indagar cual debería ser el valor maximo de $\frac{P}{N}$.

Si fijo N, estoy indagando P_{max}

P_{error}	P_{max}
0.001	$0.105 * N$
0.01	$0.185 * N$
0.05	$0.37 * N$
0.1	$0.61 * N$

Una forma más rigurosa es preguntarse por la estabilidad de las N

Pido

$$(1 - P_{error})^N > 0.99$$

$$\downarrow$$

$$(1 - P_{error})^N = \sum_{n=0}^{N} \binom{N}{n} 1^{N-n} (-P_{error})^n$$

Si $P_{error} \ll 1$

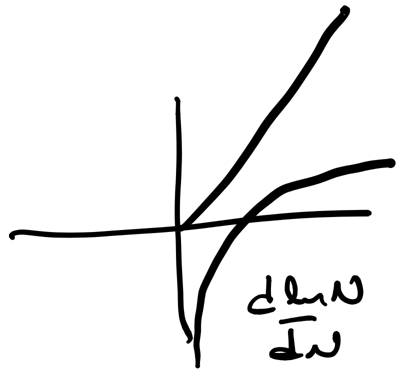
$$(1 - P_{error})^N \approx 1 - N P_{error}$$

|

hacer la cuenta

$\downarrow$

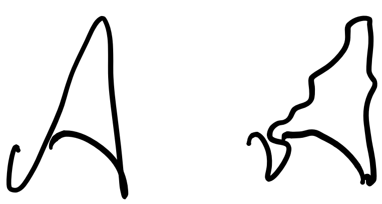
$$P_{max} = \frac{N}{2 \ln N}$$

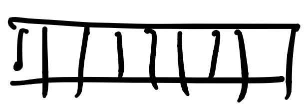


$$\frac{d \ln N}{dN} = \frac{1}{N}$$

Otra cuenta posible es pedir que las p sean estables

$$P_{max} = \frac{N}{4 \ln N}$$





```
for t:=1 to tmax
    for i:=1 to N
        h:=0
        for j:=1 to N
            h:= h + W[i,j]*S[j]
        end
        if h>0 then
            S[i]:= +1
        else
            S[i]:= -1
        end
    end
        Puedo chequear si llegó a punto fijo
end
```

$$m^{\mu} = \frac{1}{N} \sum_{i=1}^{N} \xi_i^{\mu} S_i(t=\infty)$$

$$-1 \leq m^{\mu} \leq 1$$

$$d^{\mu} = \frac{1}{4N} \sum \left(\xi_i^{\mu} - S_i(t=\infty)\right)^2$$

$$= \frac{1}{4N} \sum \Big(\underbrace{\xi_i^2}_{=1} - 2\, \xi_i^\mu S(t=\infty) + \underbrace{S(t=\infty)^2}_{=1} \Big)$$

$$= \frac{2}{4N} \sum_i \Big(1 - \xi_i^\mu S(t=\infty) \Big)$$

$$= \frac{1}{2} \left(1 - m^\mu \right)$$