Almacenamiento de memoria en redes de Hopfield

Julián López Carballal, Jorge García Beni, Rubén de San Juan Morales, Fernando García Sánchez y José Luis Miranda Mora

Red de Hopfield

- Red formada por N neuronas.
- Estados $\sigma_i(t) = \pm 1$.
- Evolución temporal discreta con pasos Δt.
- Interaccionan con pesos J_{ij}.

$$\Phi_i(t) = \sum_{j \neq i} J_{ij} \sigma_j$$

- Su fin: almacenamiento y reproducción de patrones.

$$\mathcal{P}_i^\mu = \pm 1,\ i \in [1, N],\ \mu \in [1, K] \Rightarrow \frac{1}{N} \sum_{i=1}^{N} \mathcal{P}_i^\mu \mathcal{P}_i^\nu = \delta^{\mu\nu}$$

Patrón reproducido si $\sigma_i(t) = \sigma_i(t + \Delta t) = \mathcal{P}_i^\mu$.

Red de Hopfield

Analogía con el modelo de Ising

- Red de Hopfield $\sim$ Modelo de Ising para sistemas magnéticos.
- Hamiltoniano $\equiv$ Ising ($\vec{B} = 0$).

$$H = -\frac{1}{2} \sum_{i} \sum_{j \neq i} J_{ij} \sigma_i \sigma_j$$

imponemos $J_{ij} = J_{ji}$.

Red de Hopfield

- Solapamiento: cuantificación de la coincidencia con el patrón.

$$m^{\mu}(t) = \frac{1}{N}\sum_{i} \mathcal{P}_i^{\mu}\sigma_i(t)$$

valor máximo $m^{\mu} = 1$ cuando el patrón es replicado.

- Elegimos los pesos:

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{K} \mathcal{P}_i^{\mu}\mathcal{P}_j^{\mu}$$

relacionan espines con el patrón.

Generalización del modelo de Hopfield

- Hasta ahora hemos tratado un sistema sin ruido.
- En un sistema con ruido se pueden llegar a estados fuera de la dinámica que tratamos.
- Introducimos el ruido en forma de una temperatura efectiva T.
- Adopta la dinámica de espines de Glauber. La distribución se relaja a la de Gibbs:

$$P(\{\sigma\}) \propto e^{-H(\{\sigma\})/T}$$

este es el **modelo de Hopfield generalizado**.

- Tomamos unidades $k_B = 1$.
- Densidad de energía libre:

$$f(T) = -\frac{T}{N} \langle \log Z \rangle_{\mathcal{P}}$$

- Función de partición:

$$Z = \left(\frac{N}{T}\right)^{\frac{K}{2}} e^{-\frac{K}{2T}} \int \prod_{\mu} \frac{dm^{\mu}}{\sqrt{2\pi}} \exp\left\{ -\frac{N\vec{m}^2}{2T} + \sum_i \log\left[2\cosh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T} \right) \right] \right\}$$

- Si K es finito:

$$-\frac{T \ln Z}{N} = \frac{1}{2}\vec{m}^2 - \frac{T}{N}\sum_i \ln\left[2\cosh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T}\right)\right]$$

- El parámetro de orden viene dado por la ecuación $\partial \ln Z/\partial m^\mu = 0$:

$$\vec{m} = \frac{1}{N}\sum_i \vec{\mathcal{P}}_i \tanh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T}\right)$$

- En LT: $(1/N)\sum_i \to$ promedios sobre $\{\vec{\mathcal{P}}_i\}$.

- Ecuaciones de campo medio:

$$f(T) = \frac{1}{2}\vec{m}^2 - T\left\langle \log\left[2\cosh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T}\right)\right]\right\rangle_{\mathcal{P}}$$

$$\vec{m} = \left\langle \vec{\mathcal{P}}_i \cdot \tanh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T}\right)\right\rangle_{\mathcal{P}}$$

- Promedio térmico de los espines: $\langle\sigma_i\rangle = \tanh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T}\right)$.

$$m^\mu = \left\langle \mathcal{P}_i^\mu \langle\sigma_i\rangle\right\rangle_{\mathcal{P}}$$

es el solapamiento. Similar a la magnetización del modelo de Ising $\tilde{m} = \langle\sigma_i\rangle$.

- Distribución de los patrones:

$$P(\{\mathcal{P}_i^\mu\}) = \prod_{\mu,i} \frac{1}{2}\left[\delta(\mathcal{P}_i^\mu + 1) + \delta(\mathcal{P}_i^\mu - 1)\right]$$

- Desarrollo en potencias de $\vec{m}$ de las ecuaciones de campo medio:

$$\begin{aligned} f =& - T\log 2 + \frac{1}{2}(1-\beta)\vec{m}^2 + o(\vec{m}^4) \\ m^\mu =& \beta m^\mu + \frac{2}{3}\beta^3(m^\mu)^3 - \beta^3 m^\mu \vec{m}^2 + o(\vec{m}^4) \end{aligned}$$

Generalización del modelo de Hopfield

Estados de Mattis

- Distribución de los patrones:

$$P(\{\mathcal{P}_i^\mu\}) = \prod_{\mu,i} \frac{1}{2} \left[\delta(\mathcal{P}_i^\mu + 1) + \delta(\mathcal{P}_i^\mu - 1)\right]$$

- Desarrollo en potencias de $\vec{m}$ de las ecuaciones de campo medio:

$$\begin{aligned} f =& - T \log 2 + \frac{1}{2}(1 - \beta)\vec{m}^2 + o(\vec{m}^4) \\ m^\mu =& \beta m^\mu + \frac{2}{3}\beta^3 (m^\mu)^3 - \beta^3 m^\mu \vec{m}^2 + o(\vec{m}^4) \end{aligned}$$

$\beta \equiv 1/T$.

- Para $T = 1$, $f = -T \log 2$ y $m^\mu \approx (\beta - \beta^3 \vec{m}^2) m^\mu \Rightarrow \vec{m} = 0$.
- Se mantiene para $T > 1$.
- Por debajo de $T_c = 1$: aparecen soluciones $m^\mu \neq 0$. Definimos la dimensionalidad de $\vec{m}$, n.
- Alterar el signo o el orden de las n componentes no nulas no altera las soluciones.

- Caso $n = 1$.
- Suponemos $m^1 = m \neq 0$ y $m^\mu = 0 \ \forall \mu > 1$.

$$f = \frac{1}{2} m^2 - T \log \left[2 \cosh(\beta m) \right]$$

$$m = \tanh(\beta m)$$

- Se corresponden a un estado dónde:

$$\langle \sigma_i \rangle = \mathcal{P}_i^1 \tanh(\beta m)$$

Termodinámicamente equivalente al estado ferromagnético del modelo de Ising. Existen $2K$ de estos estados, los estados de Mattis.

Generalización del modelo de Hopfield

Estados de Mattis

- Los estados de Mattis son mínimos globales de la energía libre $\Rightarrow$ estados estables.
- Calculamos en $T = 0$ para $n = 1$:

$$\lim_{x\to\infty} \tanh mx = 1$$
$$\lim_{x\to\infty} \frac{1}{x} \log\left(2\cosh mx\right) = 1$$

$$m_{(n=1)}(T = 0) = 1$$

$$f_{(n=1)}(T = 0) = \frac{1}{2}$$

- En campo medio:

$$\lim_{T\to 0} T \log\left[2\cosh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T}\right)\right] = \vec{m}\cdot\vec{\mathcal{P}}_i$$

$$\lim_{T\to 0} \tanh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T}\right) = \mathrm{sgn}\left(\vec{m}\cdot\vec{\mathcal{P}}_i\right)$$

$$\vec{m}_{\mathrm{CM}}(T=0) = \left\langle \vec{\mathcal{P}}_i \tanh\left(\frac{\vec{m}\cdot\vec{\mathcal{P}}_i}{T}\right)\right\rangle_{\mathcal{P}}$$

$$f_{\mathrm{CM}}(T=0) = \frac{1}{2}\vec{m}^2$$

- $m^\mu \leq 1$

$$\vec{m}^2 = \sum_{\mu=1}^{K} (m^\mu)^2$$

Implica $\vec{m}^2 \leq 1$. La igualdad sólo se satisface si $n = 1$ (caso estados de Mattis) $\rightarrow$ $f_{\text{CM}}(T = 0) = f_{(n=1)}(T = 0) = 1/2$

Generalización del modelo de Hopfield

Estados de Mattis

- Los estados de Mattis son mínimos globales de f (son estables) para $0 \leq T \leq 1$.
- Se corresponden a la reproducción de un patrón.
- Existen estados metaestables, soluciones con $n > 1$, que serán mínimos locales de la energía libre.

Ecuación de Langevin

Formulación continua

- Necesitamos que $\sigma_i \to \pm 1$ conforme $t \to \infty$
- Incluimos fluctuaciones térmicas con $\xi_i(t)$
- Fluctuaciones independientes $\langle \xi_i(t) \xi_j(t') \rangle = 2T\delta(t - t')\delta_{ij}$

$$\frac{\partial \sigma_i}{\partial t} = -\lambda(\sigma_i^2 - 1)\sigma_i + \sum_j J_{ij}\sigma_j + h_i\sigma_i + \xi_i(t)$$

Ecuación de Langevin

Formulación discreta

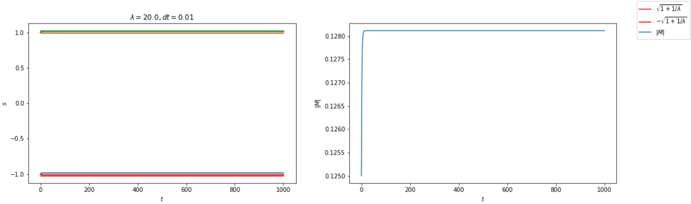
- Diferencial estocástico $\mu(dW_i) = 0, \sigma^2(dW_i) = dt$

$$d\sigma = \left(-\lambda(\sigma_i^2 - 1)\sigma_i + \sum_j J_{ij}\sigma_j + h_i\sigma_i \right) dt + \sqrt{2T}dW_i$$

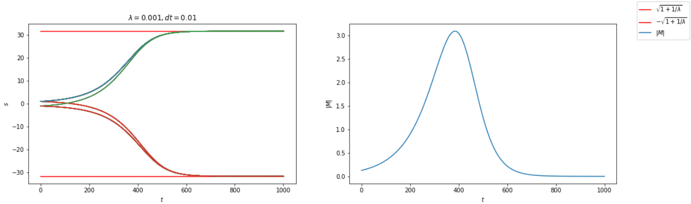
Ecuación de Langevin

Valor de λ

- Valores demasiado grandes de λ hacen que no haya evolución



- Valores demasiado grandes de λ hacen no se acerce a ± 1



Ecuación de Langevin
Valor de λ

- Valores cercanos a 1 dan una evolución razonable