

Práctica 6: Memorias Asociativas

Evelyn G. Coronel
Redes Neuronales - Instituto Balseiro

(19 de mayo de 2020)

Cuestiones generales del modelo de Hopfield

El modelo de Hopfield contempla la idea de solucionar de la manera más simple el problema de memorizar patrones. Este problema consiste en memorizar un conjunto de patrones ξ_i^μ , de tal manera que si se presenta un patrón ζ_i^μ , la red responde con el patrón almacenado que más se parece a la entrada dada [1].

La red almacena p patrones, que se indexan con $\mu = 1, \dots, p$, mientras que las unidades dentro de la red se indexan mediante $i = 1, \dots, N$. La convención que se toma en este trabajo es que cada patrón ξ_i^μ puede valer $+1$ o -1 .

Sobre la actualización

Se toma una regla de actualización asincrónica, esta implica que:

- En cada época, se selecciona una unidad i a ser actualizada aplicando la regla de la Ec. (3)
- Cada unidad elige independientemente actualizarse o no, con una probabilidad en cada época.

Para el modelo de Hopfield sin ruido, la probabilidad es constante en cada actualización es igual a 0.5, es decir, se actualiza a cada paso con $+1$ o con -1 . En cambio, para el modelo de Hopfield con ruido, la probabilidad de cambio está dada por una función que depende de la salida h_i y de otro parámetro que simula un ruido térmico T .

Salida de la red para varios patrones

Se define la matriz de conexiones w_{ij} , análoga a una matriz de pesos, según la Ec. 1.

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad (1)$$

donde el índice $i, j = 1, \dots, N$ recorre las unidades de la red y el índice $\mu = 1, \dots, p$ recorre el conjunto de patrones.

Dada una entrada ζ_i^μ , definamos el parámetro h_i^ν como $h_i^\nu = \sum_j w_{ij} \zeta_j^\nu$. Para la validación de las redes de esta práctica, se utilizaron los mismos patrones almacenados en la matriz de conexiones como entradas a la red.

EJERCICIO I: EL MODELO DE HOPFIELD SIN RUIDO

Definamos el parámetro de carga $\alpha = p/N$, que es el número de patrones a almacenar como fracción de la cantidad de unidades por patrón. Para el caso del modelo de Hopfield sin ruido, se espera que exista un valor de α crítico, con un valor de $\alpha_c \approx 0.138$, donde la red pierde su capacidad de memorizar patrones.

Para el modelo de Hopfield sin ruido, alimentando la entrada con un patrón ξ_i^ν , la primera iteración se realiza la siguiente operación:

$$S_i := \text{sgn}(h_i^\nu - \theta_i) \quad (2)$$

donde $\text{sgn}(x)$ es la función signo y para este trabajo $\theta_i = 0$. Para el siguiente paso, se toma como entrada los valores de S_i de la iteración anterior y se actualizan según la regla de la Ec.3 hasta que el sistema alcanza la condición de la Ec.4, conocida como condición de estabilidad.

$$\text{Cada iteración } S_i := \text{sgn}(h_i^\nu), \quad (3)$$

$$\text{Converge cuando } S_i := \text{sgn}(h_i^\nu) \quad \forall i \quad (4)$$

Posteriormente, después que el sistema converge, se calcula el overlap m^μ como

$$m^\mu = \frac{1}{N} \sum_i S_i \xi_i^\mu \quad (5)$$

donde ξ_i^μ es el patrón que se usó como condición inicial y S_i es el valor al que convergió el sistema.

A. Resultados

Se utilizaron valores de $N = 500, 1000, 2000, 4000$ con valores de $\alpha = 0.12, 0.14, 0.16, 0.18$, tomando un patrón como condición inicial e iterando el sistema hasta alcanzar la condición de estabilidad de la Ec. 4. Para posteriormente calcular m^μ y obtener su distribución para una cantidad de patrones N_{patrones} , con N y α fijos. Estos histogramas se muestran en las Figs. 1, 2, 3 y 4 para un ancho de bin de 0.05.

Haciendo una analogía con un sistema de espines, consideremos que N se corresponde con el tamaño, donde $N \times N$ es el tamaño de la red. Para distintos tamaños y para $\alpha = 0.12$, se observa que la mayor parte de los overlaps es cercano a 1. Esto indica que el sistema puede reconocer los patrones que se almacenaron en la matriz

de conexiones. En cambio, a partir de $\alpha = 0.14$, para los tamaños estudiados, el histograma tiene una componente alrededor de 0.4. Por que el valor de $\alpha_c \approx 0.138$ es consistente con simulación.

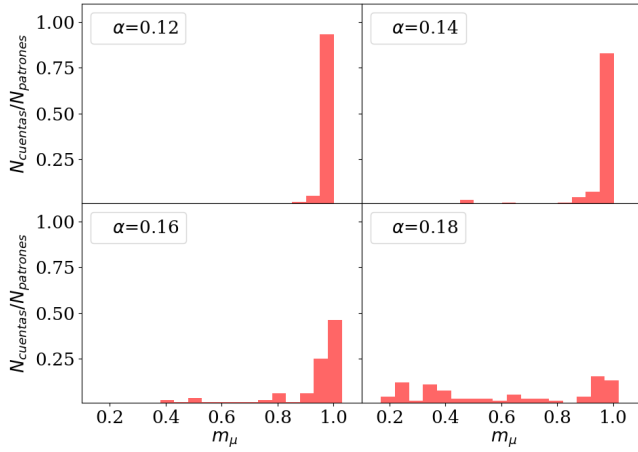


Fig. 1: Para $N = 500$

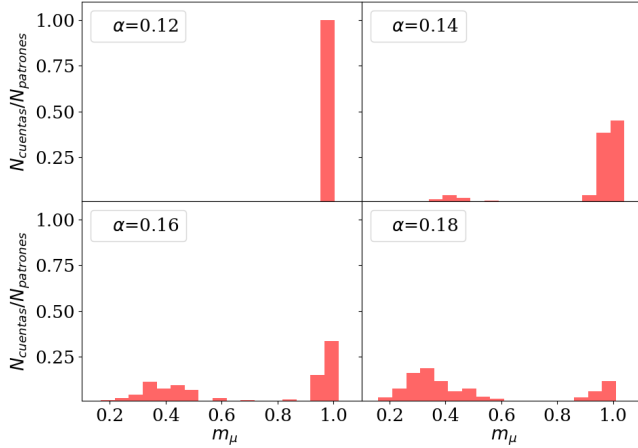


Fig. 2: Para $N = 1000$

A medida que aumenta el valor de α , la componente del histograma en la región de $m^\mu \approx 0.3$ también aumenta. Esto es consecuencia de que pasamos el valor de α_c , donde la red ya no puede reconocer los patrones con los que estamos alimentando la entrada, es decir que el sistema ya no puede almacenar patrones para está cantidad de patrones.

Para el valor más alto de α estudiado, se observan distintos comportamientos del histograma para distintos tamaños de red. Comparando las Figs. 1 y 4, donde para $N = 500$ se observa que el sistema reconoce aproximadamente un 20 % de los patrones presentados, en cambio para $N = 4000$ el sistema no es capaz de almacenar ninguno de patrones utilizados para la verificación del sistema.

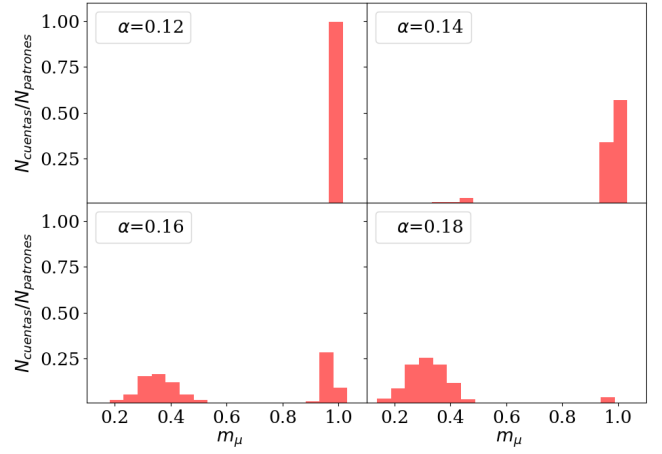


Fig. 3: Para $N = 2000$

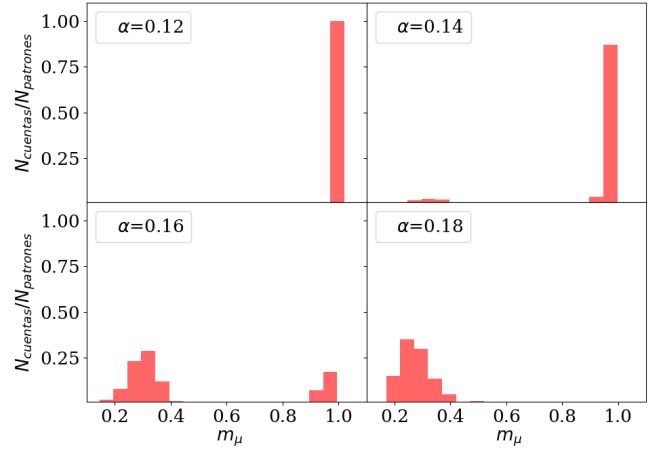


Fig. 4: Para $N = 4000$

Otro aspecto en particular de los histogramas por encima de α_c , es que se espera que el overlap sea nulo, en cambio se observa que sistema converge a un valor de aproximadamente 0.3.

B. Modificaciones a los pesos

En este caso, las conexiones entre unidades son ± 1 , manteniendo el signo del valor de la definición 1. Para $N = 1000$ y para $\alpha = 0.05, 0.09, 0.1, 0.12$, como se muestra en la Fig.5. Se observa que al cambiar los pesos a variables enteras ± 1 , el α_c está entre 0.1 y 0.12, por lo que disminuyó ligeramente con respecto al caso anterior.

La ventaja de este algoritmo sobre el caso anterior es que se necesita menos memoria para almacenar la matriz de conexiones, ya que solo son dos valores enteros. Entonces la memoria de la máquina que realiza la simulación puede asociarse a un solo bit a cada entrada del matriz $w_{ij}[1]$.

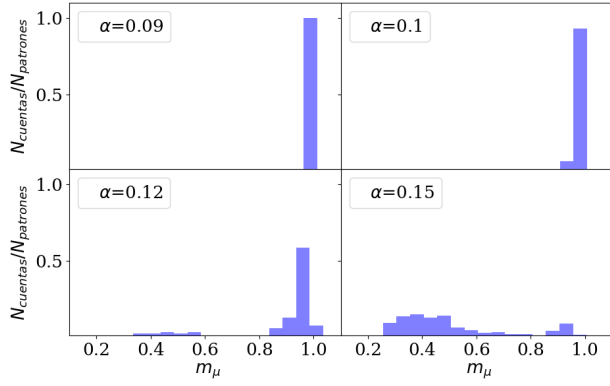


Fig. 5: Para $N = 1000$ y los pesos modificados

EJERCICIO II: EL MODELO DE HOPFIELD CON RUIDO

Para el análisis de esta sección, utilizamos la analogía a un sistema de espines. En el modelo de Hopfield con ruido tiene una diferencia a la hora de la actualización de los S_i . En este caso, tenemos la regla de actualización dada por la Ec. 6.

$$S(t+1)_i = \begin{cases} +1 & \text{Probabilidad } P(h_i, t, +1) \\ -1 & \text{Probabilidad } P(h_i, t, -1) \end{cases}, \quad (6)$$

$$P(h_i, t, \pm 1) = \frac{\exp(\pm \beta h_i(t))}{\exp(-\beta h_i(t)) + \exp(\beta h_i(t))} \quad (7)$$

donde la ecuación de $P(h_i, t, \pm 1)$ viene dada por la Ec. 7. Nótese que $P(h_i, t, +1) + P(h_i, t, -1) = 1$. El parámetro β es inversamente proporcional al nivel de ruido del sistema. Para la implementación del algoritmo, se toma un número P_{ref} con distribución uniforme entre 0 y 1, y se compara

con $P(h_i, t, +1)$, si $P(h_i, t, +1) \geq P_{ref}$ se hace el cambio a +1, caso contrario a -1.

1. Resultados

Tomando los parámetros iguales a $N = 4000$, $p = 40$ y $N = 1000$ con $p = 100$, para temperaturas T en el rango $[0.1, 2.0]$, tomando como $k_B = 1$ por lo que $\beta = 1/T$. El resultado de las simulaciones en función de la temperatura se muestran en la Fig. 6.

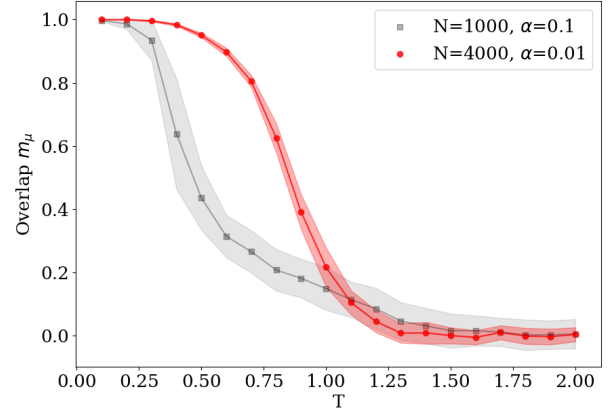


Fig. 6: El overlap en función de T

En la Fig. 6 se ve que para temperaturas bajas, el overlap de la red es cercana a 1, pero a medida que aumenta el ruido térmico la red empieza a perder la capacidad de almacenar patrones. Dependiendo del valor de α del sistema, la temperatura a la que pierde esta capacidad es distinta. Mediante estas dos simulaciones se ve que a menor α , mayor es la temperatura donde ocurre el quiebre en m_μ en función de T .