

REDES NEURONALES

Practico 1

Redes de Hopfield

Prof: F. A. Tamarit
Autor: Eric N. Jurio

October 10, 2010

Introducción

Características del modelo de Hopfield

Aquí presentamos un resumen de las características del modelo de Hopfield según lo visto en el teórico.

- En este modelo en particular todas las neuronas están conectadas a todas.
- La matriz de pesos en el conexionado de las neuronas (W) es simétrica, lo cual es irreal desde el punto de vista biológico, pero facilita mucho las cosas a la hora de hacer cuentas sobre el modelo.
- Este tipo de redes es mayormente usado como memoria asociativa. Donde para agregar un concepto e^u (representado normalmente por un arreglo de bits) a la matriz de pesos W se realiza la siguiente operación

$$W_{i,j} = e_i^u \cdot e_j^u$$

- Para que la red funcione bien, los conceptos tienen que ser notablemente distintos. Tienen que estar distribuidos aleatoriamente en el espacio de configuraciones.
- Los atractores son puntos fijos, que en el caso determinista, la red se estabiliza y en caso de neuronas estocásticas la red queda oscilando alrededor del atractor.
- Para el modelo tradicional de red neuronal de Hopfield, al ser inicializada cerca de un concepto, la red reconoce bien si se cumple

$$\frac{N^{\circ}conceptos}{N^{\circ}neuronas} \leq 0.138$$

- En el caso de neuronas estocásticas, dependiendo del nivel de ruido, pueden saltar barreras o rugosidades en el espacio de configuraciones llegando a mínimos mas profundos, pero a medida que la relación $\frac{N^{\circ}conceptos}{N^{\circ}neuronas}$ decrece, el nivel ruido al cual la red deja de reconocer también decrece.

Estos últimos puntos serán observados en los gráficos resultados de los experimentos.

Objetivos

En este informe analizaremos el grado de reconocimiento de las redes neuronales de Hopfield usando tanto neuronas deterministas como estocásticas, para distintas cantidades de neuronas y conceptos (variando el nivel de ruido en el caso estocástico).

Llamaremos grado de reconocimiento de la red ($\langle m \rangle$) al promedio de bits acertados respecto del concepto patrón al estabilizarse la red a lo largo del tiempo (en el caso determinista) o al transcurrir una cierta cantidad de ciclos (en el caso estocástico), siendo inicializada cerca (dentro de la cuenca de atracción) del concepto patrón (mediremos el porcentaje de reconocimiento o porcentaje de bits que coinciden entre el concepto patrón y el resultado de la red)

En el caso determinista, analizaremos el grado de reconocimiento de la red, respecto de la relación $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$, fijando la cantidad de neuronas y variando la cantidad de conceptos, a su vez tomando el promedio de 30 corridas distintas para cada cantidad de conceptos. Este procedimiento lo realizaremos para tres tamaños distintos de redes.

En el caso estocástico, analizaremos el grado de reconocimiento de la red, a medida que incrementamos el nivel de ruido. Esto lo haremos para 3 tamaños de redes y para 3 cantidades de conceptos distintos por tamaño, comparando los resultados para experimentos que tengan una misma relación $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$.

Detalles de la implementación

En la implementación del modelo de Hopfield usada para la experiencia, utilizamos todas las optimizaciones vistas en el teórico. Por lo tanto no usamos la matriz de pesos W . También se aplicó la optimización a nivel de bits lo cual disminuye el uso de memoria.

En pruebas adicionales notamos la marcada diferencia en tiempo de procesamiento para mismas cantidades de conceptos y distintos números de neuronas.

Red de Hopfield Determinista

Realizamos experimentos con redes neuronales de Hopfield para N neuronas y P conceptos con las siguientes relaciones:

- Para $N=416$ y $P \in [2; 208]$ con paso de a 2.
- Para $N=832$ y $P \in [4; 416]$ con paso de a 4.
- Para $N=1664$ y $P \in [8; 832]$ con paso de a 8.

A cada paso tomamos el promedio de 30 mediciones del grado de reconocimiento, para distintos conceptos aleatorios y graficamos el grado de reconocimiento ($\langle m \rangle$) respecto de $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$ obteniendo como resultados gráficos muy similares.

A continuación observamos el gráfico para 1664 neuronas, con P variando de 8 a 832 con barras de error.

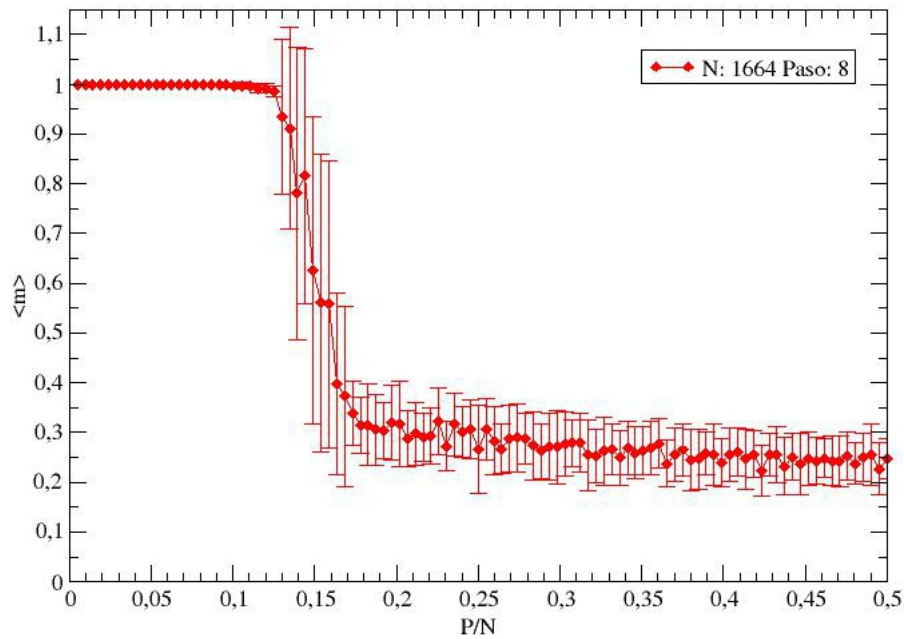


Figure 0.1: $\langle m \rangle$ (grado de reconocimiento) respecto de $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$ con barras de error.

- Podemos ver claramente como la red tiene un grado de reconocimiento ($\langle m \rangle$) de 1 y su desviación estándar es 0 cuando se da que $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}} \leq 0.13$ indicando que la red se va al atractor o muy cerca de el.
- Cuando $0.13 \leq \frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}} \leq 0.17$ el $\langle m \rangle$ decrece abruptamente de 1 a 0.3 y su desviación estándar es máxima indicando una etapa de transición pronunciada. La red tiende a desplazarse del atractor quedando en algún radio dentro de la cuenca de atracción pero cada vez mas alejado a medida que crece la cantidad de conceptos; o bien aveces reconoce (quedándose cerca del atractor) y otras veces no reconoce (siendo atrapada en otra cuenca de otro atractor).
- Para $0.17 \leq \frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$ se puede observar que $\langle m \rangle$ se mantiene casi constante en el intervalo $[0.3; 0.25]$ decreciendo paulatinamente con una desviación estándar de 0.1 lo que indica que la red, prácticamente no reconoce y se mantiene lejos del patrón inicial, quizás en la cuenca de otro atractor vecino. Es decir para $0.17 \leq \frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$ la red tiende a salir de su cuenca de atracción y ser atraída por otros atractores.

A continuación comparamos los tres gráficos superpuestos sin las barras de error.

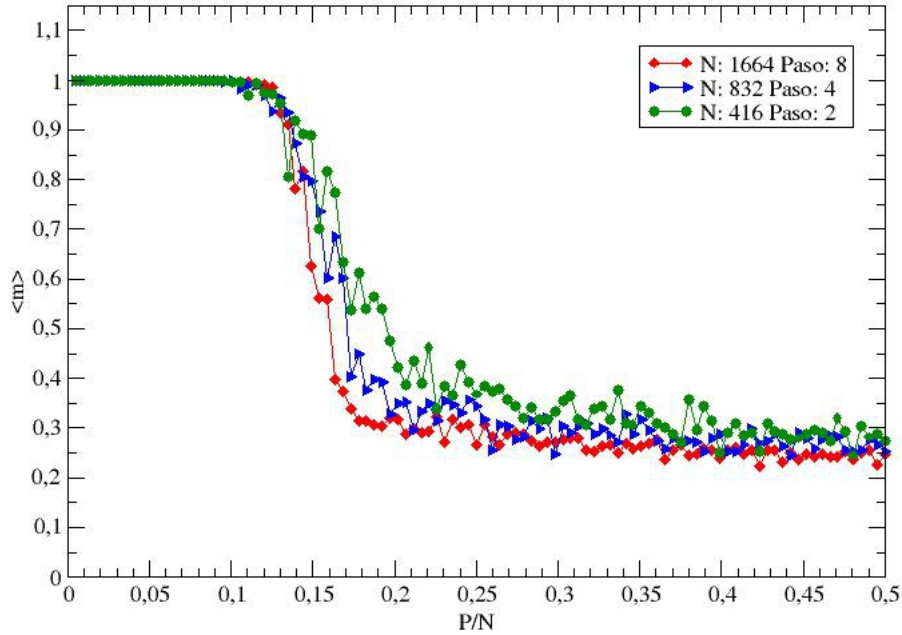


Figure 0.2: 3 curvas $\langle m \rangle$ (grado de reconocimiento) respecto de $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$ para distintas cantidades de neuronas.

- Como se dijo al principio, las 3 curvas son similares y a medida que se incrementa el numero de neuronas en la red, se obtienen resultados mas nítidos en las curvas.
- Notamos como, manteniendo la relación $0.17 \leq \frac{N^{\circ}conceptos}{N^{\circ}neuronas}$, el grado de reconocimiento toma valores cada vez menores a medida que se incrementa el numero de neuronas, lo cual se condice con la teoría que dice que para redes de gran cantidad de neurona, cuando $0.138 \leq \frac{N^{\circ}conceptos}{N^{\circ}neuronas}$, el grado de reconocimiento tiende a 0.
- Podemos ver también que cuando $0.13 \leq \frac{N^{\circ}conceptos}{N^{\circ}neuronas} \leq 0.17$, la pendiente de las curvas es cada vez mas pronunciada a medida que se aumenta el numero de neuronas.

Red de Hopfield Estocástica

Para cada experimento con redes neuronales de Hopfield estocásticas, fijamos la cantidad de neuronas y de conceptos. Realizamos las distintas mediciones de grado de reconocimiento, variando el nivel de ruido de las neuronas de la red en el rango $[0.1; 1.2]$. Agrupamos las curvas resultantes según conservan la misma relación $\frac{N^{\circ}conceptos}{N^{\circ}neuronas}$ de la siguiente manera:

- $\frac{N^{\circ}conceptos}{N^{\circ}neuronas} = 0.012$ Con N=416; 832; 1664 y P=5;10;20 respectivamente.
- $\frac{N^{\circ}conceptos}{N^{\circ}neuronas} = 0.024$ Con N=416; 832; 1664 y P=10;20;40 respectivamente.
- $\frac{N^{\circ}conceptos}{N^{\circ}neuronas} = 0.036$ Con N=416; 832; 1664 y P=15;30;60 respectivamente.

Para cada nivel de ruido tomamos el promedio de 30 mediciones del grado de reconocimiento, para distintos conceptos aleatorios y graficamos el grado de reconocimiento ($\langle m \rangle$) respecto del nivel de ruido, obteniendo como resultados gráficos muy similares. (Para cada medición dejamos evolucionar previamente algunos ciclos a la red)

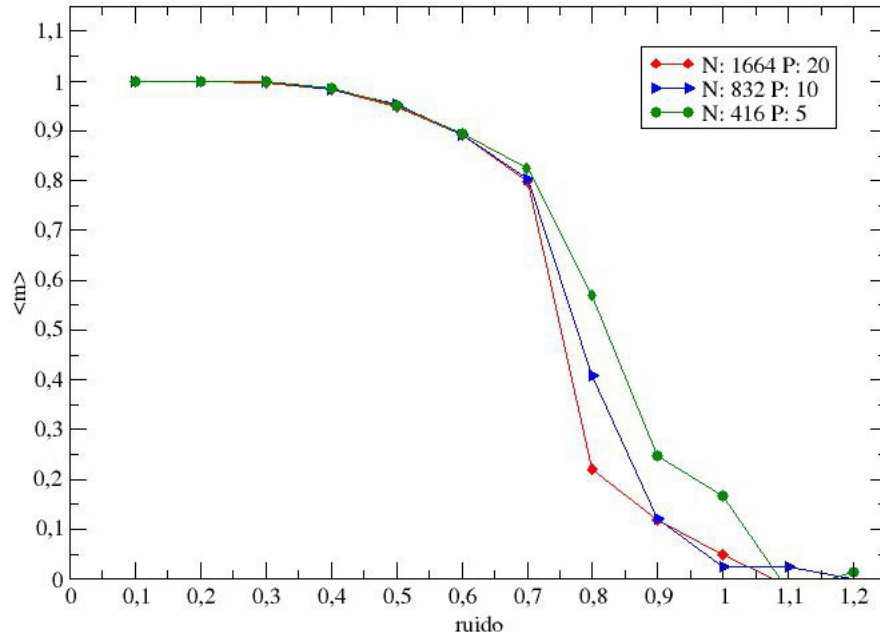


Figure 0.3: Grado de reconocimiento $\langle m \rangle$ respecto del nivel de ruido para la relación $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}} = 0.012$

- Podemos observar que a partir de un nivel de ruido de 0.3 a 0.7 las 3 curvas comienzan a decaer paulatinamente, lo que nos indica que el nivel de ruido hace que la red quede oscilando alrededor del atractor pero cada vez con un radio mayor.
- Luego, cuando el nivel de ruido esta entre 0.7 a 1 las 3 curvas tienen una pendiente pronunciada. Podemos pensar que en estos niveles de ruidos, la red esta tan perturbada que comienza a salirse de la cuenca del atractor, llegando a niveles de ruido mas altos, donde la red ya no reconoce mas nada y adopta valores aleatorios recorriendo todo el espacio de configuraciones.

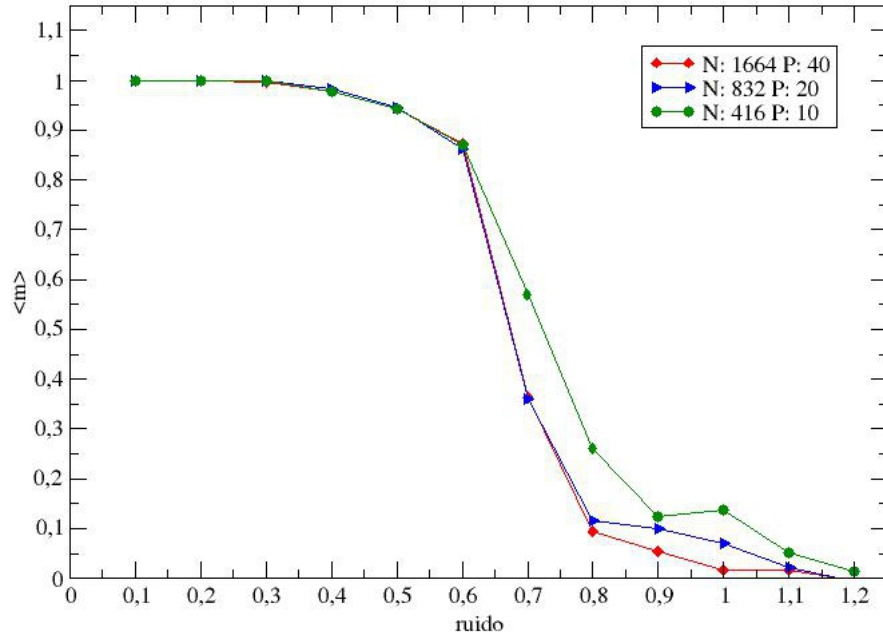


Figure 0.4: Grado de reconocimiento $\langle m \rangle$ respecto del nivel de ruido para la relación $\frac{N^{\circ}conceptos}{N^{\circ}neuronas} = 0.024$

- Aquí podemos observar un comportamiento de las redes con relación $\frac{N^{\circ}conceptos}{N^{\circ}neuronas} = 0.024$ muy similar al anterior donde la relación era $\frac{N^{\circ}conceptos}{N^{\circ}neuronas} = 0.012$
- En comparación con el gráfico anterior, podemos destacar que la caída abrupta se produce para un nivel de ruido menor
 - $\frac{N^{\circ}conceptos}{N^{\circ}neuronas} = 0.012$ la caída se produce con $0.7 \leq ruido \leq 0.9$
 - $\frac{N^{\circ}conceptos}{N^{\circ}neuronas} = 0.024$ la caída se produce con $0.6 \leq ruido \leq 0.8$

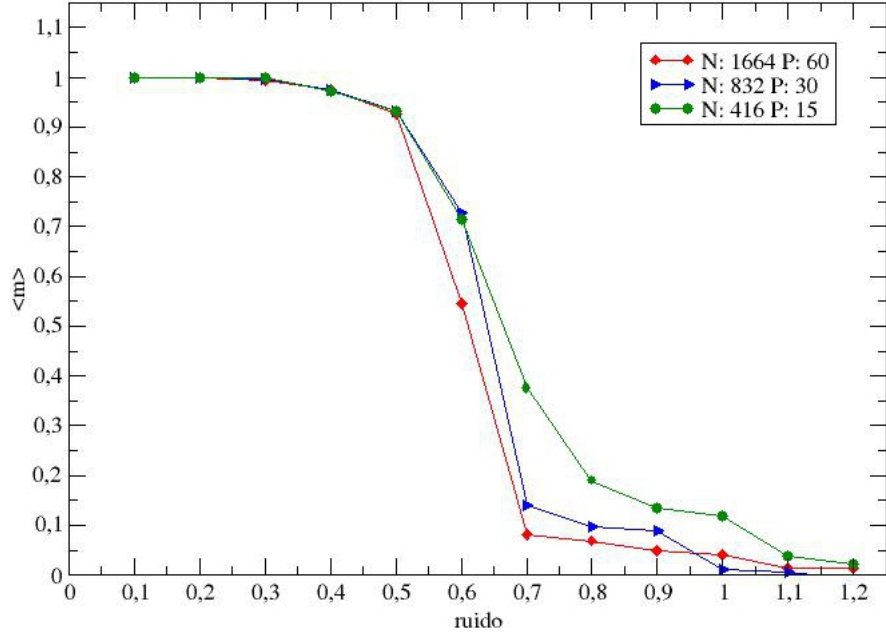


Figure 0.5: Grado de reconocimiento $\langle m \rangle$ respecto del nivel de ruido para la relación $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}} = 0.036$

- Nuevamente podemos ver en comparación con los gráficos anteriores que la caída abrupta se produce para un nivel de ruido menor
 - $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}} = 0.012$ la caída se produce con $0.7 \leq \text{ruido} \leq 0.9$
 - $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}} = 0.024$ la caída se produce con $0.6 \leq \text{ruido} \leq 0.8$
 - $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}} = 0.036$ la caída se produce con $0.55 \leq \text{ruido} \leq 0.75$

De la comparación entre los 3 gráficos podemos observar que:

- Para redes que conserven la misma relación $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$, se comportan de manera muy similar ante distintos niveles de ruido.
- A medida que la relación $\frac{N^{\circ} \text{conceptos}}{N^{\circ} \text{neuronas}}$ aumenta, el nivel de ruido al cual la red deja de reconocer es cada vez menor.

Este ultimo echo lo podemos observar en el siguiente gráfico donde comparamos el grado de reconocimiento de una red de 1664 neuronas para 3 cantidades de conceptos distintos, respecto del nivel de ruido.

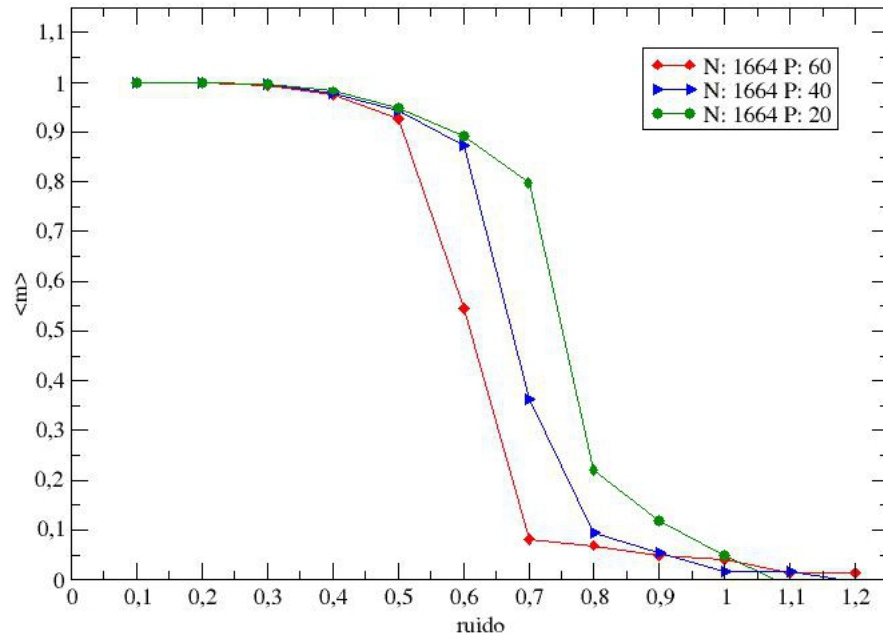


Figure 0.6: $\langle m \rangle$ para 3 cantidades de conceptos en una misma red, respecto del nivel de ruido.

Podemos notar claramente, como en una misma red neuronal, a medida que incrementamos la cantidad de conceptos, la misma deja de reconocer a niveles de ruido cada vez menores.

Conclusiones

Redes deterministas

1. Pudimos ver claramente en los gráficos de las redes neuronales deterministas, que cuando $0.138 \leq \frac{N^{\circ}conceptos}{N^{\circ}neuronas}$, la red deja de reconocer.
2. A medida que incrementamos el numero de neuronas en la red, los gráficos resultantes se asemejan a los vistos en la teoría.

Redes estocásticas

1. Redes que conservan la misma relación $\frac{N^{\circ}conceptos}{N^{\circ}neuronas}$, se comportan de manera muy similar para cada valor de ruido.
2. A medida que aumenta el nivel de ruido, la red tiende a salir del atractor correcto, pasando por otros atractores. Para niveles de ruido muy alto, la red deja de comportarse como memoria asociativa y comienza a recorrer el espacio de configuraciones aleatoriamente.
3. A medida que aumenta relación $\frac{N^{\circ}conceptos}{N^{\circ}neuronas}$, el nivele de ruido al cual la red deja de reconocer es cada vez menor.