

Memorias asociativas

Modelos de la computación (Redes Recurrentes y Autónomas)

Francisco Fernández Navarro

Departamento de Lenguajes y Ciencias de la Computación
Área: Ciencias de la Computación e Inteligencia Artificial



1 Memorias asociativas

- Introducción
- El asociador lineal
- Asociador no lineal simple (red de Hopfield)
- La Memoria Asociativa Bidireccional (BAM)

Definición de Memoria asociativa

Memoria asociativa

La memoria asociativa en el contexto de redes neuronales se refiere a la capacidad de una red neuronal para recordar patrones de entrada previamente aprendidos y recuperarlos cuando se presenta una entrada parcial o incompleta que se encuentra asociada con esos patrones.

La recuperación de la información se consigue según el grado de similitud entre el patrón de entrada (clave) y los patrones memorizados (referencias).



Evalando semejanza

- Usaremos la desviación cuadrática como medida de semejanza:

$$d(\mathbf{x}^*, \mathbf{x}_n) = \sum_{k=1}^K (x_k^* - x_{nk})^2,$$

donde $\mathbf{x}^* = (x_1^*, \dots, x_K^*) \in \mathbb{R}^K$ es el patrón clave,
 $\mathbf{x}_n = (x_{n1}, \dots, x_{nK}) \in \mathbb{R}^K$ es el n -ésimo patrón de referencia
almacenado, $n \in \{1, \dots, N\}$, y $d(\mathbf{x}^*, \mathbf{x}_n)$ es la semejanza existente
entre el patrón clave \mathbf{x}^* y el n -ésimo patrón de referencia, siendo
 K el número de componentes del patrón.

- Si las componentes de los vectores son binarias ($\{0, 1\}$), $d(\mathbf{x}^*, \mathbf{x}_n)$ coincide con la distancia de Hamming.
- Si los vectores son bipolares ($\{-1, 1\}$) la distancia de Hamming se calcula con $\frac{d(\mathbf{x}^*, \mathbf{x}_n)}{4}$.

Tipos de memorias

Heteroasociativa

Para un conjunto de pares de entradas y salidas

$\{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_n, \mathbf{y}_n)\}$, con $\mathbf{x}_N \in \mathbb{R}^K$ y $\mathbf{y}_N \in \mathbb{R}^J$, establece una correspondencia f de \mathbb{R}^K en \mathbb{R}^J , de tal manera que $f : \mathbf{x}_n \rightarrow \mathbf{y}_n$, $n \in \{1, \dots, N\}$. Además si \mathbf{x} está más próximo a \mathbf{x}_i que a cualquier otro \mathbf{x}_j entonces $f : \mathbf{x} \rightarrow \mathbf{y}_i$.

Autoasociativa

Establece la misma correspondencia que la memoria heteroasociativa pero siendo los patrones de entrada y de salida los mismos, es decir, $f : \mathbf{x}_n \rightarrow \mathbf{x}_n$.

1 Memorias asociativas

- Introducción
- El asociador lineal
- Asociador no lineal simple (red de Hopfield)
- La Memoria Asociativa Bidireccional (BAM)

El asociador lineal

Objetivo

Para un conjunto de pares de entradas y salidas $\{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_N, \mathbf{y}_N)\}$, con $\mathbf{x}_N \in \mathbb{R}^K$ y $\mathbf{y}_N \in \mathbb{R}^J$, queremos establecer una correspondencia f de \mathbb{R}^K en \mathbb{R}^J , de tal manera que $f : \mathbf{x}_n \rightarrow \mathbf{y}_n, n \in \{1, \dots, N\}$.

Asumimos que...

Los vectores de entrada (claves) son ortonormales, es decir, $(\mathbf{x}_k)' \mathbf{x}_k = 1$ y $(\mathbf{x}_k)' \mathbf{x}_i = 0, \forall k \neq i$.

El asociador lineal

Forma funcional

Un asociador lineal es una aplicación lineal de la forma:

$$\hat{\mathbf{y}}_n = \mathbf{W}\mathbf{x}_n,$$

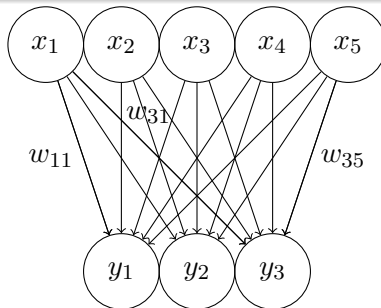
siendo $\mathbf{W} \in \mathbb{R}^{J \times K}$, la matriz de pesos que hace el mapping de \mathbb{R}^K a \mathbb{R}^J .

Arquitectura

- K unidades de entrada y J unidades de salida independientes.
- Cada entrada se conecta a todas las unidades de salida.

Ejemplo de asociador lineal

Ejemplo de asociador lineal con 5 atributos en la entrada y 3 en la salida $K = 5, J = 3$. $\mathbf{W} \in \mathbb{R}^{3 \times 5}$



Estimación de parámetros

- Ley de Aprendizaje:

$$\mathbf{W} = \sum_{n=1}^N \mathbf{y}_n (\mathbf{x}_n)'.$$

- Alternativamente:

$$\mathbf{W} = \mathbf{Y}\mathbf{X}^\dagger = \mathbf{Y}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}',$$

siendo $\mathbf{Y} = ((\mathbf{y}_1)'(\mathbf{y}_2)' \dots (\mathbf{y}_N)') \in \mathbb{R}^{J \times N}$ el conjunto de patrones de salida, $\mathbf{X} = ((\mathbf{x}_1)'(\mathbf{x}_2)' \dots (\mathbf{x}_N)') \in \mathbb{R}^{K \times N}$ el conjunto de patrones de entrada y $\mathbf{X}^\dagger \in \mathbb{R}^{N \times K}$ la matriz pseudoinversa de la matriz de patrones de entrada.

Asociador lineal

- ¿Se devuelven los patrones de referencia correctos?

$$f(\mathbf{x}_i) = \hat{\mathbf{y}}_i = \mathbf{W}\mathbf{x}_i = \sum_{n=1}^N \mathbf{y}_n(\mathbf{x}_n)' \mathbf{x}_i = \mathbf{y}_i,$$

ya que $(\mathbf{x}_n)' \mathbf{x}_i = 0$ cuando $i \neq n$, y $(\mathbf{x}_n)' \mathbf{x}_i = 1$ cuando $i = n$ (por la condición de ortonormalidad de los vectores).

- ¿Qué ocurre si perturbamos ligeramente la entrada?

$$\begin{aligned} f(\mathbf{x}_i + \boldsymbol{\epsilon}) &= \mathbf{W}(\mathbf{x}_i + \boldsymbol{\epsilon}) = \sum_{n=1}^N \mathbf{y}_n(\mathbf{x}_n)' (\mathbf{x}_i + \boldsymbol{\epsilon}) \\ &= \mathbf{y}_i + \sum_{n=1}^N \mathbf{y}_n(\mathbf{x}_n)' \boldsymbol{\epsilon} = \mathbf{y}_i + f(\boldsymbol{\epsilon}), \end{aligned}$$

donde $\boldsymbol{\epsilon}$ es el término de perturbación introducido. El asociador lineal interpola la salida.

Asociador lineal

Ventaja

Simplicidad.

Desventaja

La condición de ortogonalidad de los patrones clave

1 Memorias asociativas

- Introducción
- El asociador lineal
- Asociador no lineal simple (red de Hopfield)
- La Memoria Asociativa Bidireccional (BAM)

Asociador no lineal simple (red de Hopfield)

Objetivo

Deseamos memorizar N patrones $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N\}$, con $\mathbf{s}_n \in \mathbb{R}^K, n = 1, 2, \dots, N$, mediante una memoria asociativa dinámica no lineal, teniendo en cuenta que los patrones a memorizar no tienen por qué ser ortogonales.

Dinámica de computación

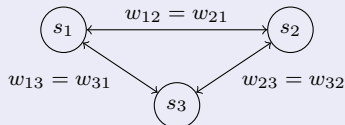
La salida del sistema se obtendrá tras un proceso de computación. Es decir, la red comienza con la configuración determinada por la entrada y los estados de las neuronas van cambiando, hasta que llegamos a una situación estable en la que podemos leer el resultado de la computación.

La red habrá memorizado un patrón si se estabiliza en él.

Asociador no lineal simple

Arquitectura

- K unidades de proceso.
- Todas las neuronas conectadas con todas.



Estimación de parámetros

- Regla de Hebb:

$$w_{ij} = \frac{1}{K} \sum_{n=1}^N s_{ni} s_{nj},$$

donde w_{ij} es el peso asociado a la neurona i con la j , s_{ni} es la componente i -ésima del patrón n -ésimo y K es el número de atributos de los patrones a memorizar.

Dinámica de computación

- Denotaremos a $s_k(i)$ al estado de la neurona k en el instante i de ejecución. Será 1 si la neurona está activa y -1 si no lo está. La actualización se realiza en tiempo discreto.
- La entrada $\mathbf{s} = (s_1, s_2, \dots, s_K)$, sirve para inicializar el estado de las neuronas, es decir, $s_k(0) = s_k, \forall k \in \{1, 2, \dots, K\}$.
- Cuando $s_k(i) = s_k(i+1) = s_k, i \geq 1, \forall k \in \{1, 2, \dots, K\}$, diremos que la red se ha estabilizado:

$$s_k(i+1) = \operatorname{sgn} \left(\sum_{l=1}^K w_{kl} s_l(i) \right)$$

Caso de un único patrón a memorizar

- Para entender el funcionamiento de nuestro modelo neuronal consideremos el caso particular en el que queremos memorizar sólo un patrón, $\mathbf{s} = (s_1, s_2, \dots, s_K)$.
- La regla de aprendizaje se particulariza a: $w_{ij} = \frac{1}{K} s_i s_j$.
- Con ese entrenamiento, ¿se memoriza el patrón?

$$\begin{aligned} s_k(1) &= \operatorname{sgn} \left(\sum_{l=1}^K w_{kl} s_l(0) \right) = \operatorname{sgn} \left(\sum_{l=1}^K w_{kl} s_l \right) \\ &= \operatorname{sgn} \left(\sum_{l=1}^K \frac{1}{K} (s_k s_l) s_l \right) = \operatorname{sgn} (s_k) = s_k, \end{aligned}$$

ya que $s_l \cdot s_l = 1$ (con una codificación bipolar), y $\frac{1}{K} > 0$, por lo que el signo dependería únicamente del signo de s_k .

- Las autoconexiones, w_{ii} , bajo la formulación expuesta valdrían $w_{ii} = \frac{N}{K}$, ya que $s_i \cdot s_i = 1$.
- Sin embargo, se suele tomar $w_{ii} = 0$ ya que no supone diferencia apreciable en la estabilidad de la red.

Patrones no memorizados

¿Qué ocurre ante entradas no memorizadas que difieren en k componentes del patrón memorizado?

$$r_i = \begin{cases} -s_i & \text{si } i = \{1, 2, \dots, k\} \\ s_i & \text{si } i = \{k+1, k+2, \dots, K\} \end{cases} \quad (1)$$

$$\begin{aligned} s_k(1) &= \operatorname{sgn} \left(\sum_{l=1}^K w_{kl} s_l(0) \right) = \operatorname{sgn} \left(\sum_{l=1}^K w_{kl} r_l \right) \\ &= \operatorname{sgn} \left(\sum_{l=1}^k w_{kl} (-s_l) + \sum_{l=k+1}^K w_{kl} s_l \right) \\ &= \operatorname{sgn} \left(\sum_{l=1}^k \frac{1}{K} (s_k s_l) (-s_l) + \sum_{l=k+1}^K \frac{1}{K} (s_k s_l) s_l \right) \\ &= \operatorname{sgn} \left(\frac{-1}{K} \sum_{l=1}^k s_k + \frac{1}{K} \sum_{l=k+1}^K s_k \right) = \operatorname{sgn} \left(\frac{-k}{K} s_k + \frac{K-k}{K} s_k \right) = \operatorname{sgn} \left(\left(1 - \frac{2k}{K} \right) s_k \right) \end{aligned}$$

¿Cuándo cambiará de signo?

Cuando $\frac{2k}{K} > 1$, es decir: $s_i(1) = \begin{cases} -s_i & \text{si } k > K/2 \\ s_i & \text{si } k \leq K/2 \end{cases}$

- Junto al patrón que le enseñamos, la red ha aprendido por cuenta propia el patrón opuesto.
- Siguiendo el criterio de semejanza planteado el patrón referencia que se recupera será uno u otro

Capacidad de almacenamiento

Indicadores de la capacidad de una red:

- Considerando patrones almacenados (N) y total de neuronas (K):

$$C = \frac{N}{K}.$$

- En base a los patrones (N) y al número de conexiones (K_w):

$$C_w = \frac{N}{K_w}.$$

Teorema

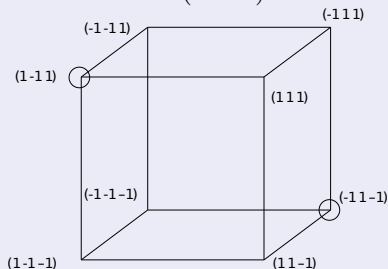
La capacidad máxima de una red de Hopfield está acotada por:

$$C \leq \frac{1}{4 \ln K}$$

Ejemplo

Definición del problema

Diseño de un asociador no lineal que memorice los patrones $(1 \ -1 \ 1)$ y $(-1 \ 1 \ -1)$. Probarla con la entrada $(1 \ 1 \ 1)$.



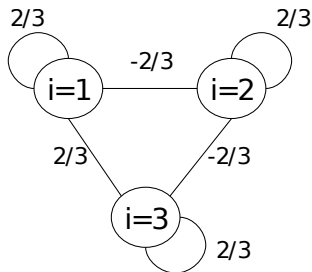
Ejemplo

- Nuestra red tendrá 3 unidades de proceso. Cada una de ellas se encargará de procesar una componente del vector de entrada.
- La dinámica de la computación es la definida.
- Empleando la regla de Hebb, calculamos la interacción entre las neuronas:

$$w_{12} = \frac{1}{3}(-1 - 1) = \frac{-2}{3}, w_{13} = \frac{1}{3}(1 + 1) = \frac{2}{3}, w_{23} = \frac{1}{3}(-1 - 1) = \frac{-2}{3}$$
$$w_{11} = w_{22} = w_{33} = \frac{2}{3}$$

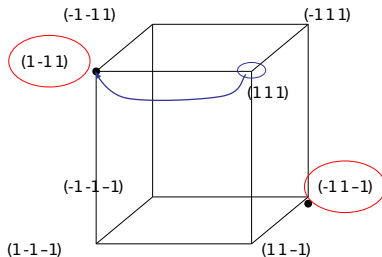
Ejemplo

- Estado inicial: $(1 \ 1 \ 1)$
- Computamos hasta que se estabiliza la red.



Ejemplo

- Estado inicial: $(1\ 1\ 1)$
- Computamos hasta que se estabiliza la red.
- Estado final: $(1\ -1\ 1)$.



1 Memorias asociativas

- Introducción
- El asociador lineal
- Asociador no lineal simple (red de Hopfield)
- La Memoria Asociativa Bidireccional (BAM)

Definición

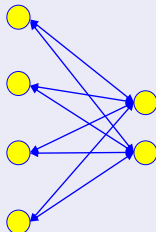
Se trata de una memoria heteroasociativa que asocia vectores bipolares (binarios) de distinta dimensión. Es decir, puede asociar a un código binario de 10 bits una firma digitalizada de 10.000 bits, o a una imagen 140.000 bits una imagen comprimida de 7.000 bits.

Memoria Asociativa Bidireccional

Arquitectura

La BAM consta de dos capas de unidades de proceso, K unidades en la primera capa y J en la segunda, estando conectadas entre sí solamente las unidades de la primera capa con las unidades de la segunda.

Representaremos por w_{ij} el valor del peso sináptico de la conexión de la unidad i de la primera capa con la unidad j de la segunda, $i = 1, 2, \dots, K$, $j = 1, 2, \dots, J$, siendo estas conexiones bidireccionales.



Memoria Asociativa Bidireccional

Dinámica de computación

La red bidireccional comienza con una configuración inicial y va actualizando simultáneamente en cada iteración todas las unidades de proceso de una capa y a continuación las de la otra capa, y así sucesivamente hasta que la red se estabilice (alcance una configuración de equilibrio).

Regla de actualización

$$x_k(i+1) = \begin{cases} 1 & \text{si } \sum_{j=1}^J w_{kj} y_j(i) > \theta_k \\ x_k(i) & \text{si } \sum_{j=1}^J w_{kj} y_j(i) = \theta_k \\ -1 & \text{si } \sum_{j=1}^J w_{kj} y_j(i) < \theta_k \end{cases} \quad (2)$$

$$y_j(i+1) = \begin{cases} 1 & \text{si } \sum_{k=1}^K w_{kj} x_k(i) > \eta_j \\ y_j(i) & \text{si } \sum_{k=1}^K w_{kj} x_k(i) = \eta_j \\ -1 & \text{si } \sum_{k=1}^K w_{kj} x_k(i) < \eta_j \end{cases} \quad (3)$$

Memoria Asociativa Bidireccional

Función de energía

$$E(i) = - \sum_{j=1}^J \sum_{k=1}^K w_{kj} x_k(i) y_j(i) + \sum_{k=1}^K \theta_k x_k(i) + \sum_{j=1}^J \eta_j y_j(i).$$

Estimación de parámetros: Regla de Hebb

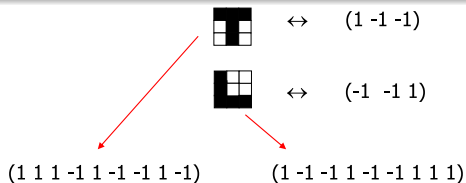
Para un conjunto de pares $\{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_N, \mathbf{y}_N)\}$, con $\mathbf{x}_n \in \mathbb{R}^K$ y $\mathbf{y}_n \in \mathbb{R}^J$, queremos definir una función de mapping $f: \mathbf{x}_n \rightarrow \mathbf{y}_n, n \in \{1, \dots, N\}$. La matriz de pesos \mathbf{W} :

$$\mathbf{W} = \sum_{n=1}^N \mathbf{x}_n (\mathbf{y}_n)' \in \mathbb{R}^{K \times J},$$

donde el elemento w_{ij} de la matriz \mathbf{W} nos da el peso sináptico de la conexión entre la unidad i de la primera capa y la unidad j de la segunda capa.

Ejemplo BAM

Memorizar los patrones y códigos siguientes:



Ejemplo BAM

$$W = \begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & -2 \\ 2 & 0 & -2 \\ -2 & 0 & 2 \\ 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 2 \\ 0 & -2 & 0 \\ -2 & 0 & 2 \end{pmatrix}$$

Actualización:
(-1 -1 1)



Patrón a reconocer:



(-1 -1 -1 1 1 -1 -1 1 1 -1)

$$\begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & -2 \\ 2 & 0 & -2 \\ -2 & 0 & 2 \\ 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 2 \\ 0 & -2 & 0 \\ -2 & 0 & 2 \end{pmatrix}$$

= (-8 -2 8)

Ejemplo BAM

Patrón a reconocer:



Actualización:

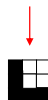
$(-1 \ -1 \ 1)$

$$\begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & -2 \\ 2 & 0 & -2 \\ -2 & 0 & 2 \\ 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 2 \\ 0 & -2 & 0 \\ -2 & 0 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ -4 \\ 4 \\ -4 \\ -2 \\ 4 \\ 2 \\ 4 \end{pmatrix}$$





Actualización:


$(1 \ -1 \ -1 \ 1 \ -1 \ -1 \ 1 \ 1 \ 1)$



Ejemplo BAM

$(1 \ -1 \ -1 \ 1 \ -1 \ -1 \ 1 \ 1 \ 1)$


 Actualización:
 $(-1 \ -1 \ 1)$

$(1 \ -1 \ -1 \ 1 \ -1 \ -1 \ 1 \ 1 \ 1)$

$$\begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & -2 \\ 2 & 0 & -2 \\ -2 & 0 & 2 \\ 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 2 \\ 0 & -2 & 0 \\ -2 & 0 & 2 \end{pmatrix} = (-12 \ -6 \ 12)$$


Resultado final:
 Patrón a reconocer:
 

 $(-1 \ -1 \ 1)$

¡Gracias por vuestra atención!

