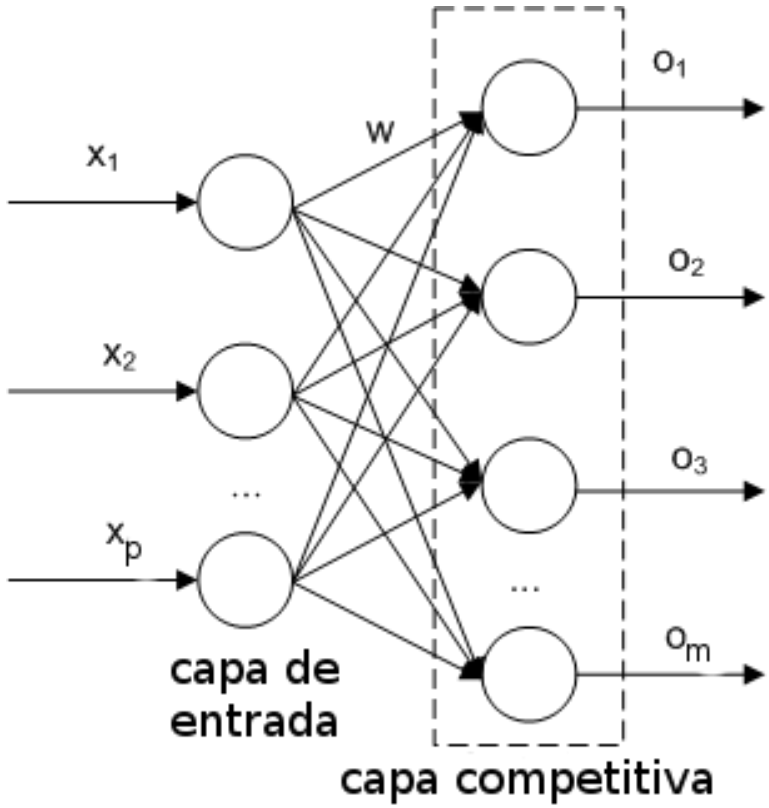
Redes Neuronales Artificiales Competitivas No Supervisadas

Carleos Artime, C.; Corral Blanco, N.

15 de diciembre de 2017

Red Neuronal Competitiva No Supervisada

- Las neuronas de salida compiten entre ellas para activarse
- Sólo se activa la que alcanza un mayor *potencial sináptico*
- Las neuronas se especializan en detectar la presencia de ciertos patrones de entrada



Redes Neuronales Artificiales Competitivas No Supervisadas

Los elementos básicos de estas redes son:

- La capa de entrada
- La capa de salida que tiene naturaleza competitiva
- Las neuronas de la capa de salida son de naturaleza binaria y sólo pueden activarse una con cada patrón de entrada.
- Para un vector de entrada $\vec{x} = (x_1, \ldots, x_p)$, el potencial sináptico de la neurona de salida k viene dado por la expresión:

$$h_k = \sum_{r=1}^{p} w_{rk} x_r - \theta_k = \sum_{r=1}^{p} w_{rk} x_r - \frac{1}{2} \sum_{r=1}^{p} w_{rk}^2$$

- La salida k vale: $o_k = \begin{cases} 1 & \text{si } h_k = \text{máx}\{h_1, \ldots, h_m\} \\ 0 & \text{en otro caso} \end{cases}$
- La neurona mayor potencial sináptico es la que tiene la menor distancia euclídea entre su vector de pesos $\vec{w}$ y $\vec{x}$.

El entrenamiento de la red pretende alcanzar los siguientes objetivos:

- Activar la neurona cuyos pesos sean los más parecidos al patrón de entrada
- Conseguir que los pesos sean una buena representación de los datos de entrenamiento

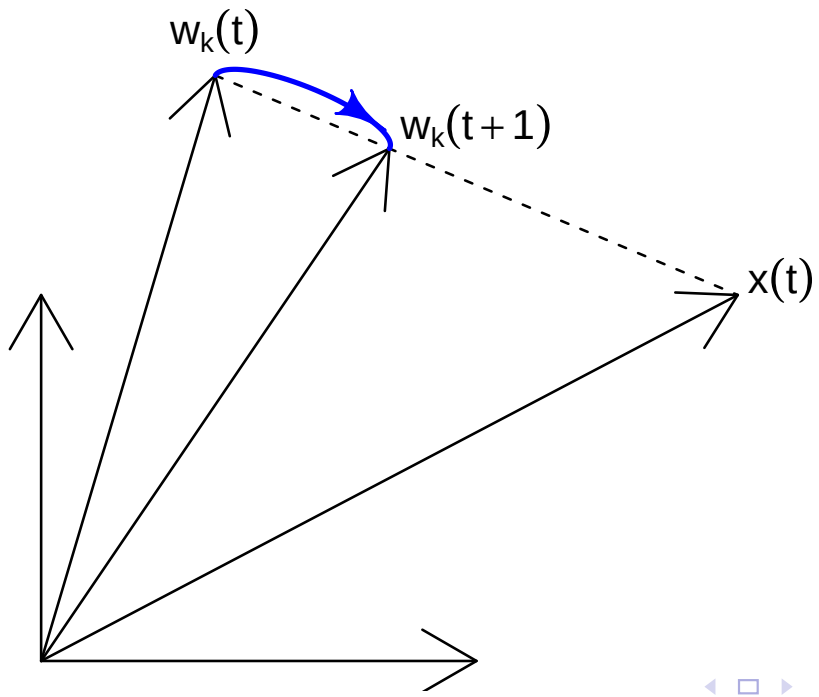
Para ello se define el siguiente error:

$$E(\vec{x}) = \sum_{k=1}^{m} a_k \|\vec{x} - \vec{w}_k\|^2 = \sum_{k=1}^{m} a_k \sum_{r=1}^{p} (x_r - w_{rk})^2$$

$$\text{donde } a_k = \begin{cases} 1 & \text{si } \|\vec{x} - \vec{w}_k\|^2 = \min\limits_{i=1,\ldots,m} \|\vec{x} - \vec{w}_i\|^2 \\ 0 & \text{en el resto} \end{cases}$$

Los pesos sinápticos se estiman, de manera iterativa, usando el descenso del gradiente para minimizar el error

$$w_{jk}(t+1) = w_{jk}(t) - \eta_t \frac{\partial E}{\partial w_{jk}} = w_{jk}(t) + \eta_t \, a_k \sum_{r=1}^{p} (x_r - w_{rk})$$



Los pesos sinápticos se estiman, de manera iterativa, usando el descenso del gradiente para minimizar el error

$$w_{jk}(t+1) = w_{jk}(t) - \eta_t \frac{\partial E}{\partial w_{jk}} = w_{jk}(t) + \eta_t \, a_k \sum_{r=1}^{p} (x_r - w_{rk})$$

En cada iteración sólo cambian los pesos de la neurona ganadora.

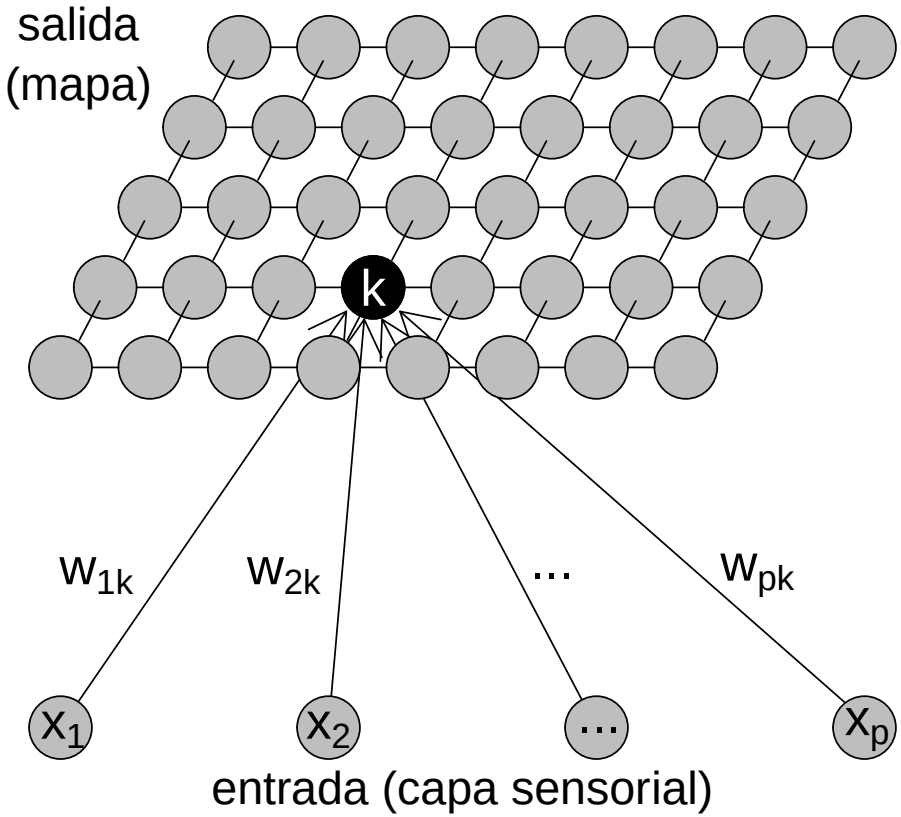
La tasa de aprendizaje debe disminuir con cada iteración del proceso de entrenamiento. Algunos criterios habituales son:

- $\eta_t = \eta_0 e^{-t/\lambda}$
- $\eta_t = 1 - \frac{t}{T}$ con $T =$ número máximo de iteraciones

Mapas de Kohonen

Naturaleza de los mapas de Kohonen (Self Organizing Map. SOM)

- Son redes de aprendizaje no supervisado
- Tienen naturaleza competitiva
- Se construyen teniendo en cuenta las posiciones que ocupan las neuronas.



Mapas de Kohonen

Los elementos básicos de una red de Kohonen son:

- La capa de salida está formada por una rejilla rectangular, de neuronas binarias, de tamaño $m_1 \times m_2$. La rejilla también puede ser hexagonal.
- Cada neurona k de la capa de salida tiene asociado un vector de posición en el plano $\vec{p}_k = (p_{k1}, p_{k2})$
- La proximidad entre las neuronas de la capa de salida. Las función más usual es

$$T_{g,k} = exp\left(-\frac{d^2(\vec{p}_g, \vec{p}_k)}{2\sigma^2}\right)$$

donde

$$d(p_g, p_k) = \left(|p_{g1} - p_{k1}|^\alpha + |p_{g2} - p_{k2}|^\alpha\right)^{\frac{1}{\alpha}}$$

  - $\alpha = 1 \Longrightarrow d$ es la distancia de Manhattan
  - $\alpha = 2 \Longrightarrow d$ es la distancia euclídea

Mapas de Kohonen

Etapas del algoritmo de una red de Kohonen:

1. Calcular el valor de activación de cada neurona de salida

$$h_k(\vec{x}) = exp(-d^2(\vec{x}, \vec{w}_k))$$

   La neurona ganadora es la que tiene un vector de pesos más cercano al patrón de entrada

2. Medir la proximidad de las neuronas, respecto de la ganadora g, teniendo en cuenta el radio de vecindad, σ

$$T_{g,k}(t) = exp\left(-\frac{d^2(\vec{p}_g, \vec{p}_k)}{2\sigma_t^2}\right)$$

   El valor de σ_t va decreciendo con cada iteración. Un criterio habitual es $\sigma_t = \sigma_0 exp(-t/\tau)$

3. Los pesos de las neuronas se modifican teniendo en cuenta su cercanía a la ganadora

$$\bigtriangledown w_{rk}(t+1) = \eta_t T_{g,k}(t)(x_r - w_{rk})$$

Mapas de Kohonen

En el entrenamiento de una red de Kohonen hay dos etapas

1. Auto-organización: Organización topológica de los vectores de pesos.
   Inicialmente se consideran una tasa de aprendizaje y radio de vecindad altos. Por ejemplo, $\eta_0 = 1$ y $\sigma_0 =$ diámetro del mapa.
2. Convergencia: Está dirigida a obtener valores estables de los pesos sinápticos. Se suele emplear una tasa de aprendizaje pequeña (0,05) y un radio igual a 1.