

Redes Neuronales de Base Radial

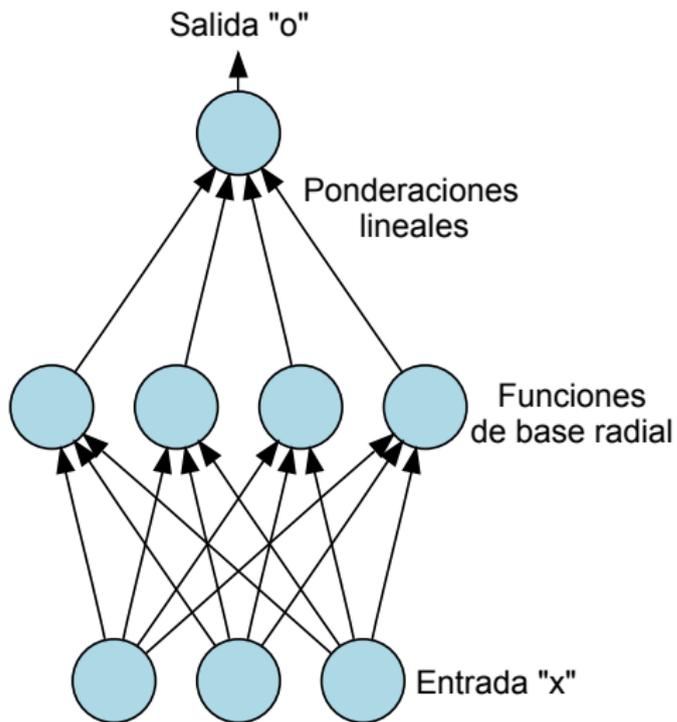
Carleos Artime, C.; Corral Blanco, N.

13 de diciembre de 2018

Características de las Redes de Neuronas de Base Radial (RNBR)

- ▶ Redes multicapa con conexiones hacia delante
- ▶ Tienen una sola capa oculta
 - ▶ Las neuronas de la capa oculta actúan localmente
 - ▶ Las neuronas de la capa oculta se activan en regiones diferentes del espacio de entradas
- ▶ Las conexiones de la capa de entrada y la capa oculta no tienen ponderaciones
- ▶ Las funciones de salida son combinaciones lineales de funciones locales no lineales

Estructura de una red neuronal radial



Funciones de base radial: la información de salida de las neuronas ocultas es una transformación no lineal del vector de entrada

$$\phi_j(\vec{x}) = \phi\left(\frac{\|\vec{x} - \vec{\mu}_j\|}{\sigma_j}\right)$$

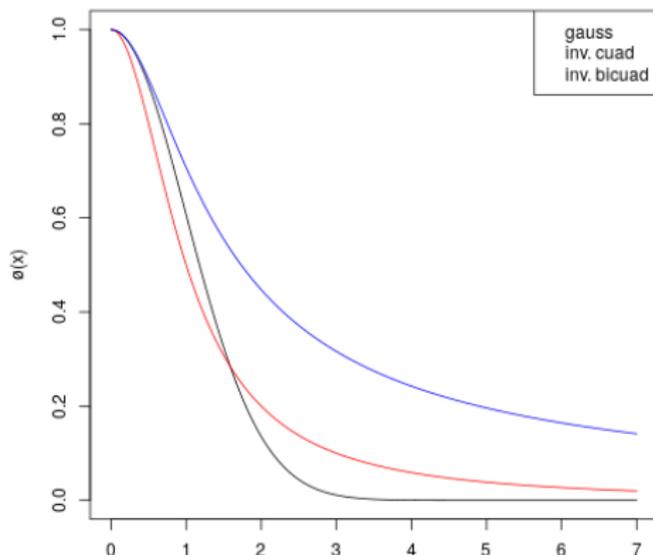
donde

- ▶ $\vec{\mu}_j = (\mu_{j1}, \dots, \mu_{jp})$ centros de las funciones de base radial
- ▶ σ_j desviaciones de las funciones;
- ▶ $\|\vec{x} - \vec{\mu}_j\|$ es la distancia euclídea desde el vector de entrada al centro de la función

Redes Neuronales de Base Radial

Funciones de base radial más habituales:

- ▶ Gaussiana $\phi(x) = \exp(-x^2/2)$
- ▶ Inversa Cuadrática: $\phi(x) = \frac{1}{1+x^2}$
- ▶ Inversa multicuadrática: $\phi(x) = \sqrt{\frac{1}{1+x^2}}$



Se dispone de N elementos de entrenamiento (\vec{x}, \vec{y}) , donde

- ▶ $\vec{x} = (x_1, \dots, x_p)$ son las variables predictoras
- ▶ $\vec{y} = (y_1, \dots, y_m)$ las variables objetivo

La salida de las neuronas viene dado por la expresión:

$$o_k(\vec{x}) = \sum_{j=1}^L w_{jk} \phi_j(\vec{x}) + \theta_k = \sum_{j=1}^{L+1} w_{jk} \phi_j(\vec{x})$$

Aprendizaje de la red:

Consiste en la determinación de todos los parámetros

- ▶ Estimación de los centros $\vec{\mu}_j$
- ▶ Estimación de las desviaciones σ_j
- ▶ Cálculo de los pesos de la capa oculta a la capa de salida w_{jk}

Métodos de aprendizaje más usuales

- ▶ **Método híbrido.** Es el más frecuente. Tiene dos etapas:
 - ▶ Fase no supervisada: se estiman los centros y desviaciones típicas de las funciones radiales
 - ▶ Fase supervisada: se calculan los pesos y umbrales de la capa de salida
- ▶ **Método supervisado:** hace una estimación supervisada, simultánea, de todos los elementos de la red

Método híbrido. Fase no supervisada

1. Determinar los centros mediante un procedimiento de análisis clúster
 - ▶ Hay que establecer de antemano el número de grupos.
 - ▶ El número de neuronas ocultas coincide con el de grupos
 - ▶ Un procedimiento muy utilizado es el de las k-medias.
 - ▶ El centro de cada neurona será el del grupo correspondiente
2. Cálculo de las desviaciones
 - ▶ Las desviaciones deben tratar de que la región de influencia de cada neurona se solape lo menos posible con las demás.
 - ▶ Un criterio habitual es calcular la desviación como la media geométrica de cada centro a sus dos vecinos más próximos

$$\sigma_k = \sqrt{\|\vec{\mu}_{i_1} - \vec{\mu}_k\| \|\vec{\mu}_{i_2} - \vec{\mu}_k\|}$$

Método híbrido. Fase Supervisada

Se buscan, de forma iterativa, las ponderaciones que hagan mínimo el error cuadrático:

$$E = \sum_{k=1}^m (y_k - o_k(\vec{x}))^2 = \sum_{k=1}^m \left(y_k - \sum_{j=1}^{L+1} w_{jk} \phi_j(\vec{x}) \right)^2$$

El cambio en las ponderaciones viene dado por:

$$w_{jk}(t+1) = w_{jk}(t) - \eta \frac{\partial E}{\partial w_{jk}} = w_{jk}(t) + \eta (y_k - o_k) \phi_j$$

Método supervisado

Se utiliza el criterio de mínimos cuadrados para estimar todos los elementos desconocidos de la red.

Se pueden perder las características locales de las RNBR.

En el ajuste de los parámetros se emplea el descenso del gradiente.

$$w_{jk}(t+1) = w_{jk}(t) - \eta_1 \frac{\partial E}{\partial w_{jk}} = w_{jk}(t) + \eta_1 (y_k - o_k) \phi_j$$

$$\mu_{jr}(t+1) = \mu_{jr}(t) - \eta_2 \frac{\partial E}{\partial \mu_{jr}} = \mu_{jr}(t) + \eta_2 \left(\sum_{k=1}^m (y_k - o_k) w_{jk} \right) \phi_j \frac{x_j - \mu_{jr}}{\sigma_j}$$

$$\sigma_j(t+1) = \sigma_j(t) - \eta_3 \frac{\partial E}{\partial \sigma_j} = \sigma_j(t) + \eta_3 \left(\sum_{k=1}^m (y_k - o_k) w_{jk} \right) \phi_j \frac{\|\vec{x} - \vec{\mu}_j\|^2}{\sigma_j^3}$$

Método supervisado

- ▶ En la etapa inicial es necesario inicializar todos los elementos de la red.
- ▶ Una buena alternativa es emplear primero el método híbrido y usar los resultados como inicio de la etapa supervisada.
- ▶ El objetivo es minimizar el error cuadrático sin tener en cuenta si se solapan las zonas de influencia de las neuronas. Por ello se pueden perder las propiedades locales de la red.

Librerías de R

- ▶ RSNNS
- ▶ neural