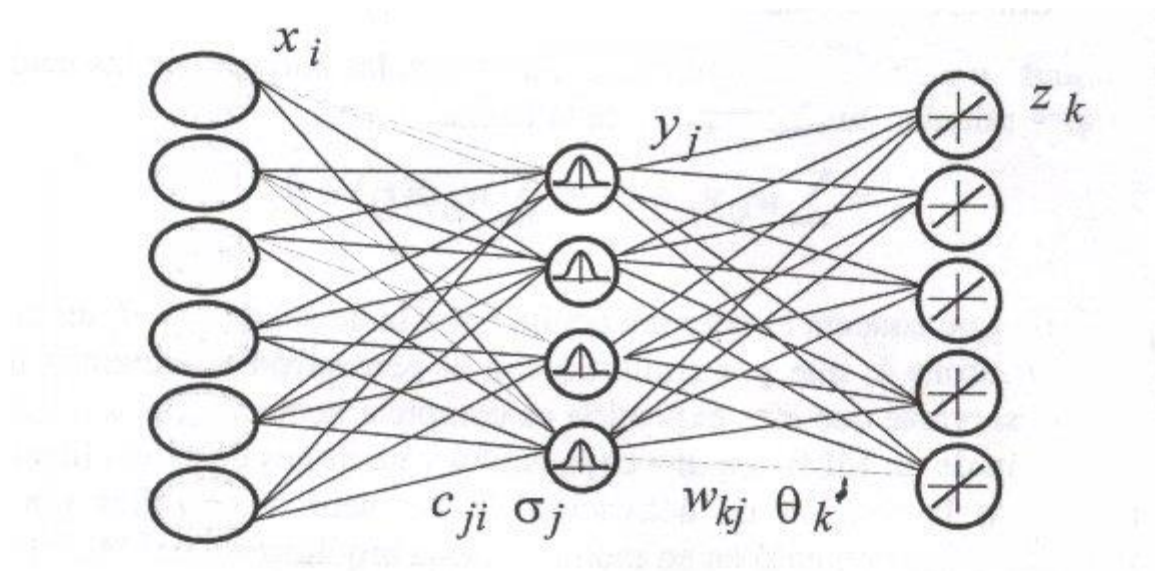


Redes de función de base radial (RBF)

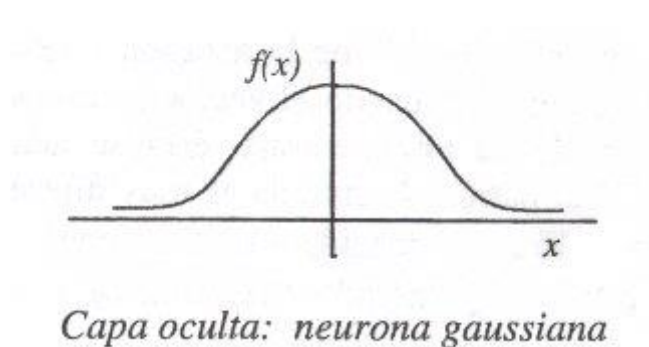
Este tipo de redes se caracteriza por tener un aprendizaje o entrenamiento híbrido. La arquitectura de estas redes se caracteriza por la presencia de tres capas: una de entrada, una única capa oculta y una capa de salida.

Figura 3-5. Arquitectura típica de una red de tipo RBF



Aunque la arquitectura pueda recordar a la de un MLP, la diferencia fundamental está en que las neuronas de la capa oculta en vez de calcular una suma ponderada de las entradas y aplicar una sigmoide, estas neuronas calculan la distancia euclídea entre el vector de pesos sinápticos (que recibe el nombre en este tipo de redes de centro o centroide) y la entrada (de manera casi analoga a como se hacía con los mapas SOM) y sobre esa distancia se aplica una función de tipo radial con forma gaussiana.

Figura 3-6. Forma funcional de una función tipo Gaussiana



Para el aprendizaje de la capa oculta, hay varios métodos, siendo uno de los más conocidos el algoritmo denominado k-medias (k-means) que es un algoritmo no supervisado de clustering. k es el número de grupos que se desea encontrar, y se corresponde con el número de neuronas de la capa oculta, que es un parámetro que hay que decidir de antemano. El algoritmo se plantea como sigue:

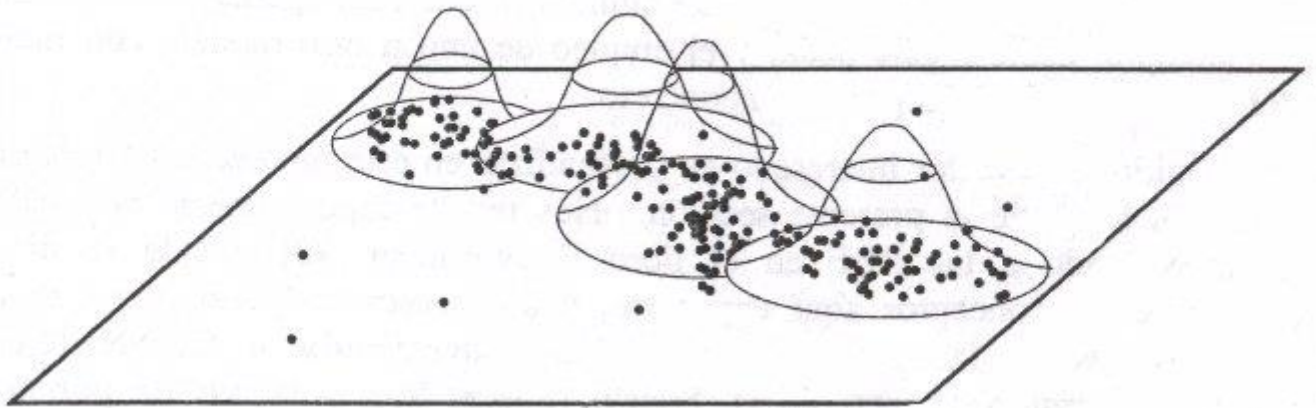
1. Inicializar los pesos (los centros) en el instante inicial. Una inicialización típica es la denominada k-primeras mediante la cual los k centros se hacen iguales a las k primeras muestras del conjunto de datos de entrenamiento $\{\mathbf{x}_p\}_{p=1..N}$

$$\mathbf{c}_1 = \mathbf{x}_1, \mathbf{c}_2 = \mathbf{x}_2, \dots, \mathbf{c}_N = \mathbf{x}_N,$$

2. En cada iteración, se calculan los dominios, es decir, se reparten las muestras entre los k centros. Esto se hace de la siguiente manera: Dada una muestra \mathbf{x}_j se calcula las distancias a cada uno de los centros \mathbf{c}_k . La muestra pertenecerá al dominio del centro cuya distancia calculada sea la menor
3. Se calculan los nuevos centros como los promedios de los patrones de aprendizaje pertenecientes a sus dominios. Viene a ser como calcular el centro de masas de la distribución de patrones, tomando que todos pesan igual.
4. Si los valores de los centros varían respecto a la iteración anterior se vuelve al paso 2, si no, es que se alcanzó la convergencia y se finaliza el aprendizaje

Una vez fijados los valores de los centros, sólo resta ajustar las anchuras de cada neurona. Las anchuras son los parámetros sigma que aparecen en cada una de las funciones gaussianas y reciben ese nombre por su interpretación geométrica, dan una medida de cuando una muestra activa una neurona oculta para que de una salida significativa

Figura 3-7. Centros en el espacio de las entradas



normalmente se toma el criterio de que para cada neurona se toma como valor sigma la distancia al centro mas cercano.

Finalmente, se entrena la capa de salida. El entrenamiento de esta capa se suele usar un algoritmo parecido al que se usa para la capa de salida del MLP. La actualizacion de los pesos viene dada por la expresi3n:

Ecuaci3n 3-5. Actualizaci3n de los pesos de la capa de salida en una red RBF

$$z_k = \sum_j w_{kj} \phi(r_j) + \theta_k$$

$$\delta w_{kj}(t) = \epsilon (d_k - z_k) \phi(r_j)$$

Con este fin se suele presentar todos los patrones de la muestra de entrenamiento varias veces. Cada una de estas veces recibe el nombre de epoca.