Sieci neuronowe typu RBF do zadania interpolacji.

Niech

$$\varphi(x) = e^{-\left(\frac{x}{r}\right)^2}, \tag{1}$$

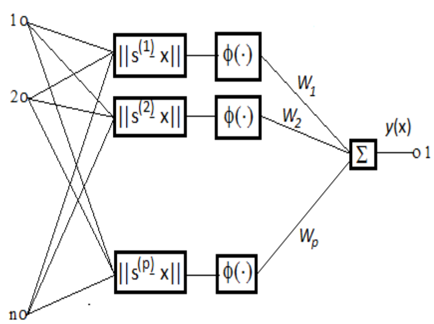
gdzie $r \in \mathbb{R}$ jest parametrem. Jest to funkcja o symetrii radialnej, która może być użyta do konstrukcji interpolatora p par wzorzec wejściowy-wzorzec wyjściowy

$$Z = \{(\mathbf{s}^{(k)}, f^{(k)})\}_{k=1}^{p}.$$

gdzie $\mathbf{s}^{(k)} \in \mathbf{R}^n$ i $f^{(k)} \in \mathbf{R}$ dla $k = 1, \dots, p$.
Rozpatrzmy sieć taką jak na rysunku 1. Wyjście takiej sieci wyraża się wzorem

$$y(\mathbf{x}) = \sum_{i=1}^{p} w_i \varphi(||\mathbf{s}^{(i)} - \mathbf{x}||).$$

dla $\mathbf{x} \in \mathbf{R}^n$.



Rysunek 1: Interpolacyjna sieć typu RBF (z radialnymi funkcjami aktywacji).

Rozpatrzmy teraz wzorzec $\mathbf{s}^{(k)}$ zapisując warunek interpolacyjny dla wyjścia

$$f^{(k)} = \sum_{i=1}^{p} w_i \phi(||\mathbf{s}^{(i)} - \mathbf{s}^{(k)}||),$$

dla $k = 1, \dots, p$. Tak więc mamy ich tyle ile elementów wektora wag $\mathbf{w} = [w_i] \in \mathbb{R}^p$.

Aby wyznaczyć wektor $\mathbf{w}$ tj. $[w_1, w_2, \dots, w_p]$ rozpatrujemy zestaw warunków: tzn.

$$\begin{cases} f^{(1)} & = & \sum_{i=1}^{p} w_i \varphi(||\mathbf{s}^{(i)} - \mathbf{s}^{(1)}||), \\ f^{(2)} & = & \sum_{i=1}^{p} w_i \varphi(||\mathbf{s}^{(i)} - \mathbf{s}^{(2)}||), \\ \dots & & \\ f^{(p)} & = & \sum_{i=1}^{p} w_i \varphi(||\mathbf{s}^{(i)} - \mathbf{s}^{(p)}||). \end{cases}$$

Z tych warunków wyznaczamy wartość składowych wektora $\mathbf{w}$ tj: $\mathbf{w}_{1:p}$. Ostatecznie więc mamy do rozwiązania układ równań z niewiadomą będącą wektorem $\mathbf{w}$

$$\mathbf{F} = \mathbf{\Phi} \mathbf{w},$$

gdzie

$$\mathbf{F} = [\mathbf{f}]^T = \begin{bmatrix} f^{(1)} \\ f^{(2)} \\ \dots \\ f^{(p)} \end{bmatrix},$$

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_p \end{bmatrix},$$

$$\mathbf{\Phi} = \begin{bmatrix} \varphi(||\mathbf{s}^{(1)} - \mathbf{s}^{(1)}||) & \varphi(||\mathbf{s}^{(1)} - \mathbf{s}^{(2)}||) & \dots & \varphi(||\mathbf{s}^{(1)} - \mathbf{s}^{(p)}||) \\ \varphi(||\mathbf{s}^{(2)} - \mathbf{s}^{(1)}||) & \varphi(||\mathbf{s}^{(2)} - \mathbf{s}^{(2)}||) & \dots & \varphi(||\mathbf{s}^{(2)} - \mathbf{s}^{(p)}||) \\ \dots & & & \\ \varphi(||\mathbf{s}^{(p)} - \mathbf{s}^{(1)}||) & \varphi(||\mathbf{s}^{(p)} - \mathbf{s}^{(2)}||) & \dots & \varphi(||\mathbf{s}^{(p)} - \mathbf{s}^{(p)}||) \end{bmatrix}.$$

Macierz $\mathbf{\Phi}$ jest nieosobliwa (chociaż zazwyczaj b. źle uwarunkowana) dzięki własnościom funkcji radialnej (1) (patrz np. [1] rozdz. 5.).

Jeśli chodzi o parametr r we wzorze (1), czyli promień gaussowskiej funkcji bazowej, to można przyjąć

$$r = \mathrm{diam}(S)/l, \quad \text{dla} \quad l = 1, 2, \dots$$

gdzie $S = \{\mathbf{s}^{(k)}\}_{k=1}^{p}$ a $\mathrm{diam}(S)$ jest odległością pomiędzy dwoma najbardziej od siebie oddalonymi wzorcami $\mathbf{s}^{(i)}$ i $\mathbf{s}^{(j)}$ w zbiorze S.

Literatura

[1] Ch. Bishop, Neural networks for pattern recognition, Clarendon Press, Oxford, 1995.