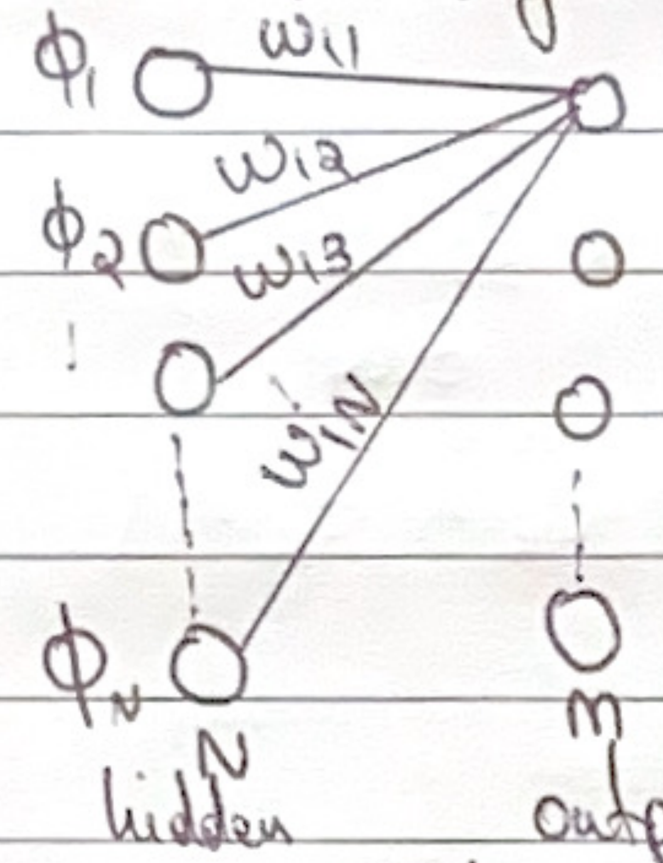
Answer-1

Stochastic Gradient Based Method for Training an RBF N.N

Consider an RBF neural network with 'N' neurons in the hidden layer. The hidden layer is fully connected to 'm' outputs.



Let ϕ be the outputs from the hidden layer. The weight matrix from hidden to output layer is given by

$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} & \cdots & W_{1N} \\ W_{21} & W_{22} & W_{23} & \cdots & W_{2N} \\ & & \vdots & & \\ W_{m1} & W_{m2} & W_{m3} & \cdots & W_{mN} \end{bmatrix}$$

where $W(n) = [W_{n1} \; W_{n2} \; W_{n3} \cdots \; W_{nN}]$

The output from hidden layer is a function of input, center and spread/radius. Thus the hidden layer output matrix is given by:

$$\psi(n) = \begin{bmatrix} \phi_1(x(1), c_1, \sigma_1) \\ \phi_2(x(2), c_2, \sigma_2) \\ \vdots \\ \phi_N(x(N), c_N, \sigma_N) \end{bmatrix}$$

The output of the RBF function is given by

$$y(n) = \sum_{K=1}^{N} W_K(n) \, \phi\{x(n), c_K, \sigma_K\} = W(n)\psi(n)$$

Let $y_d(n)$ be the desired output. The instantenous error cost function is given by

$$J(n) = \frac{1}{2}|e(n)|^2 = \frac{1}{2}\left[y_d(n) - y(n)\right]^2$$

FOR EDUCATIONAL USE

Weight Update Equation:

$$w(n+1) = w(n) - \mu_w \left.\frac{\partial J(n)}{\partial w}\right|_{w = w(n)}$$

μ_w - learning rate parameter

$$\frac{\partial J(n)}{\partial w} = \frac{\partial}{\partial w}\left(\frac{1}{2}(e(n))^2\right) = \frac{2}{2} \times e(n) \times \frac{\partial e(n)}{\partial w} \quad -①$$

$$e(n) = y_d(n) - \sum_{k=1}^{N} w_k(n)\phi\{x(n), c_k(n), \sigma_k(n)\}$$

$$\Rightarrow e(n) = y_d(n) - w(n)\psi(n)$$

$$\frac{\partial e(n)}{\partial w} = 0 - \frac{\partial (w(n)\psi(n))}{\partial w}$$

$$\frac{\partial e(n)}{\partial w} = -\psi(n)$$

Substitute in ① we get

$$\frac{\partial J(n)}{\partial w} = e(n) \times -\psi(n)$$

Thus the weight update equation is given by

$$w(n+1) = w(n) + \mu_w e(n)\psi(n)$$

Center Update Equation:

$$c_k(n+1) = c_k(n) - \mu_c \left.\frac{\partial J(n)}{\partial c_k}\right|_{c_k = c_k(n)}$$

$$\frac{\partial J(n)}{\partial c_k} = e(n)\frac{\partial e(n)}{\partial c_k} \quad -①$$

$$e(n) = y_d(n) - \sum_{k=1}^{N} w_k(n)\phi\{x(n), c_k(n), \sigma_k(n)\}$$

given that ϕ is gaussian Kernel

$$\Rightarrow \phi(n) = e^{\frac{-\|x(n) - c_k(n)\|^2}{2\sigma_k^2(n)}}$$

$$\frac{\partial \phi(n)}{\partial c_k} = \underbrace{\boxed{e^{\frac{-\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}}}}_{\phi(n)} \times -\frac{(x(n)-c_k(n))}{\sigma_k^2(n)} \times -1$$

$$\frac{\partial \phi(n)}{\partial c_k} = \frac{\phi(n)\,(x(n)-c_k(n))}{\sigma_k^2(n)}$$

$$\frac{\partial e(n)}{\partial c_k} = -\frac{w_k(n)\,\phi(n)}{\sigma_k^2(n)}\left[x(n)-c_k(n)\right]$$

Substitute in equation ① we get

$$\frac{\partial J(n)}{\partial c_k} = -\frac{e(n)\,w_k(n)\,\phi(n)}{\sigma_k^2(n)}\left[x(n)-c_k(n)\right]$$

Thus the center update fo equation is given by

$$c_k(n+1) = c_k(n) + \mu_c \frac{e(n)\,w_k(n)\,\phi(n)}{\sigma_k^2(n)}\left[x(n)-c_k(n)\right]$$

Spread Update Equation.

$$\sigma_k(n+1) = \sigma_k(n) - \mu_\sigma \left.\frac{\partial J(n)}{\partial \sigma_k}\right|_{\sigma_k = \sigma_k(n)}$$

$$\frac{\partial J(n)}{\partial \sigma_k} = e(n)\frac{\partial e(n)}{\partial \sigma_k} \quad - ①$$

$$e(n) = y_d(n) - \sum_{k=1}^{N} w_k(n)\, e^{\frac{-\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}}$$

$$\frac{\partial \phi(n)}{\partial \sigma_k} = \underbrace{\boxed{e^{\frac{-\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}}}}_{\phi(n)} \times -\|x(n)-c_k(n)\|^2 \times -\frac{2}{\sigma_k^3(n)}$$

$$= \frac{\phi(n)\,\|x(n)-c_k(n)\|^2}{\sigma_k^3(n)}$$

$$\frac{\partial e(n)}{\partial \sigma_K} = -W_K(n)\,\phi(n)\,\frac{\|x(n) - c_K(n)\|^2}{\sigma_K^3(n)}$$

Substituting in equation ① we get

$$\frac{\partial J(n)}{\partial \sigma_K} = -\frac{e(n)\,W_K(n)}{\sigma_K^3(n)}\,\phi\{x(n), c_K(n), \sigma_K(n)\}\,\|x(n) - c_K(n)\|^2$$

Thus the Spread update equation is given by

$$\sigma_K(n+1) = \sigma_K(n) + \mu_\sigma \frac{e(n)\,W_K(n)}{\sigma_K^3(n)}\,\phi\{x(n), c_K(n), \sigma_K(n)\}\,\|x(n) - c_K(n)\|^2$$