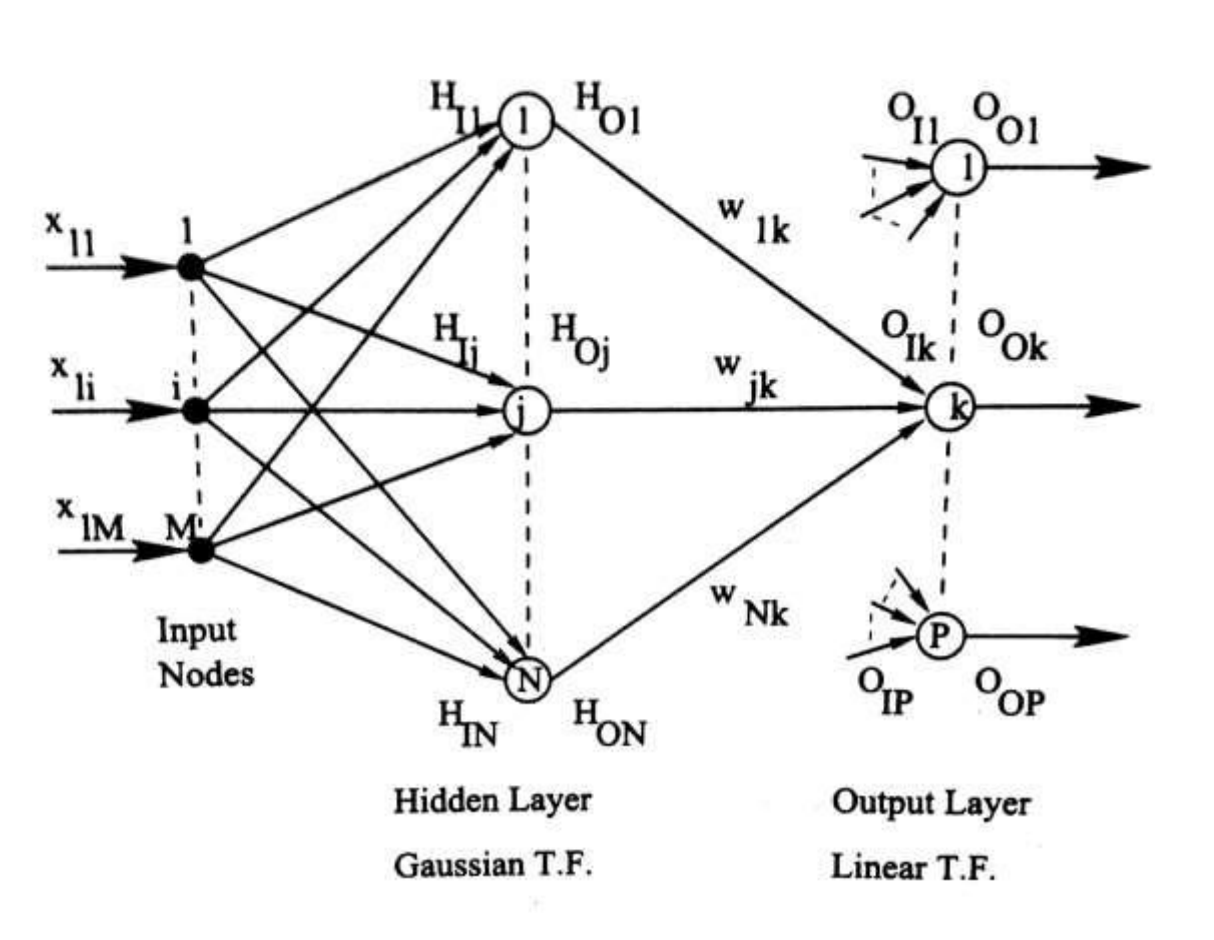
Radial Basis Function Neural Network (RBFNN)



Forward Calculations

- **Step 1: Determination of the outputs of input nodes**

  Let us consider L training scenarios.

  Input vector

$$
X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_l \\ \cdot \\ \cdot \\ x_L \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdot & \cdot & \cdot & \cdot & x_{1i} & \cdot & \cdot & \cdot & \cdot & x_{1M} \\ x_{21} & x_{22} & \cdot & \cdot & \cdot & \cdot & x_{2i} & \cdot & \cdot & \cdot & \cdot & x_{2M} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{l1} & x_{l2} & \cdot & \cdot & \cdot & \cdot & x_{li} & \cdot & \cdot & \cdot & \cdot & x_{lM} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{L1} & x_{L2} & \cdot & \cdot & \cdot & \cdot & x_{Li} & \cdot & \cdot & \cdot & \cdot & x_{LM} \end{bmatrix}
$$

Let us assume that *l*-th training scenario (that is, x_{l1}, x_{l2}, . . . , x_{li}, ..., x_{lM}) is passed through the network. Outputs of different nodes of the input layer are nothing but the inputs of corresponding nodes.

- **Step 2: Determination of the outputs of hidden layer**
  Output of j-th hidden neuron

$$H_{Oj} = \exp\left[-\frac{\left\| x_l - \mu_j \right\|^2}{2\sigma_j^2}\right]$$

- **Step 3: Determination of the inputs of output layer**

  Input of k-th neuron lying on output layer

$$O_{Ik} = \sum_{j=1}^{N} W_{jk} H_{Oj}$$

- **Step 4: Determination of the outputs of the output layer**

$$O_{Ok} = O_{ik}$$

Error in prediction of k-th output neuron

$$E_k = ½ (T_{Ok} – O_{Ok})^2$$

Tuning of RBFNN Using BP Algorithm

Incremental Mode of Training

- **Step 1: Weight Updating**

$$w_{updated} = w_{previous} + \Delta w$$

Now, $\Delta w_{jk}(t) = -\eta \dfrac{\partial E_k}{\partial w_{jk}}(t) + \alpha' \Delta w_{jk}(t-1)$

Where $\dfrac{\partial E_k}{\partial w_{jk}} = \dfrac{\partial E_k}{\partial O_{Ok}} \cdot \dfrac{\partial O_{Ok}}{\partial O_{Ik}} \cdot \dfrac{\partial O_{Ik}}{\partial w_{jk}}$

Here $$\frac{\partial E_k}{\partial O_{Ok}} = -\left(T_{Ok} - O_{Ok}\right)$$

$$\frac{\partial O_{Ok}}{\partial O_{Ik}} = 1$$

$$\frac{\partial O_{Ik}}{\partial w_{jk}} = H_{Oj}$$

- **Step 2: Mean Updating**

$$\mu_{j,updated} = \mu_{j,previous} + \Delta\mu_j$$

Now, $$\Delta\mu_j(t) = -\eta\left\{\frac{\partial E}{\partial \mu_j}(t)\right\}_{av} + \alpha'\Delta\mu_j(t-1)$$

Where, $$\left\{\frac{\partial E}{\partial \mu_j}\right\}_{av} = \frac{1}{P}\sum_{k=1}^{P}\frac{\partial E_k}{\partial \mu_j}$$

Now, $$\frac{\partial E_k}{\partial \mu_j} = \frac{\partial E_k}{\partial O_{Ok}} \cdot \frac{\partial O_{Ok}}{\partial O_{Ik}} \cdot \frac{\partial O_{Ik}}{\partial H_{Oj}} \cdot \frac{\partial H_{Oj}}{\partial \mu_j}$$

Now, $$\frac{\partial \mathrm{E_k}}{\partial \mathrm{O_{Ok}}} \cdot \frac{\partial \mathrm{O_{Ok}}}{\partial \mathrm{O_{Ik}}} = -\left(\mathrm{T_{Ok}} - \mathrm{O_{Ok}}\right)$$

$$\frac{\partial \mathrm{O_{Ik}}}{\partial \mathrm{H_{Oj}}} = \mathrm{w_{jk}}$$

$$\frac{\partial H_{Oj}}{\partial \mu_j} = H_{Oj}\left\{\frac{\left(x_{l1} + x_{l2} + \ldots.. + x_{li} + \ldots. + x_{lM}\right) - M\mu_j}{\sigma_j^2}\right\}$$

- **Step 3: Standard Deviation Updating**

$$\boldsymbol{\sigma}_{\mathbf{j,updated}} = \boldsymbol{\sigma}_{\mathbf{j,previous}} + \boldsymbol{\Delta\sigma}_{\mathbf{j}}$$

Now, $$\Delta\sigma_j(t) = -\eta\left\{\frac{\partial E}{\partial \sigma_j}(t)\right\}_{av} + \alpha'\Delta\sigma_j(t-1)$$

where $$\left\{\frac{\partial E}{\partial \sigma_j}\right\}_{av} = \frac{1}{P}\sum_{k=1}^{P}\frac{\partial E_k}{\partial \sigma_j}$$

Now, $$\frac{\partial E_k}{\partial \sigma_j} = \frac{\partial E_k}{\partial O_{Ok}} \cdot \frac{\partial O_{Ok}}{\partial O_{Ik}} \cdot \frac{\partial O_{Ik}}{\partial H_{Oj}} \cdot \frac{\partial H_{Oj}}{\partial \sigma_j}$$

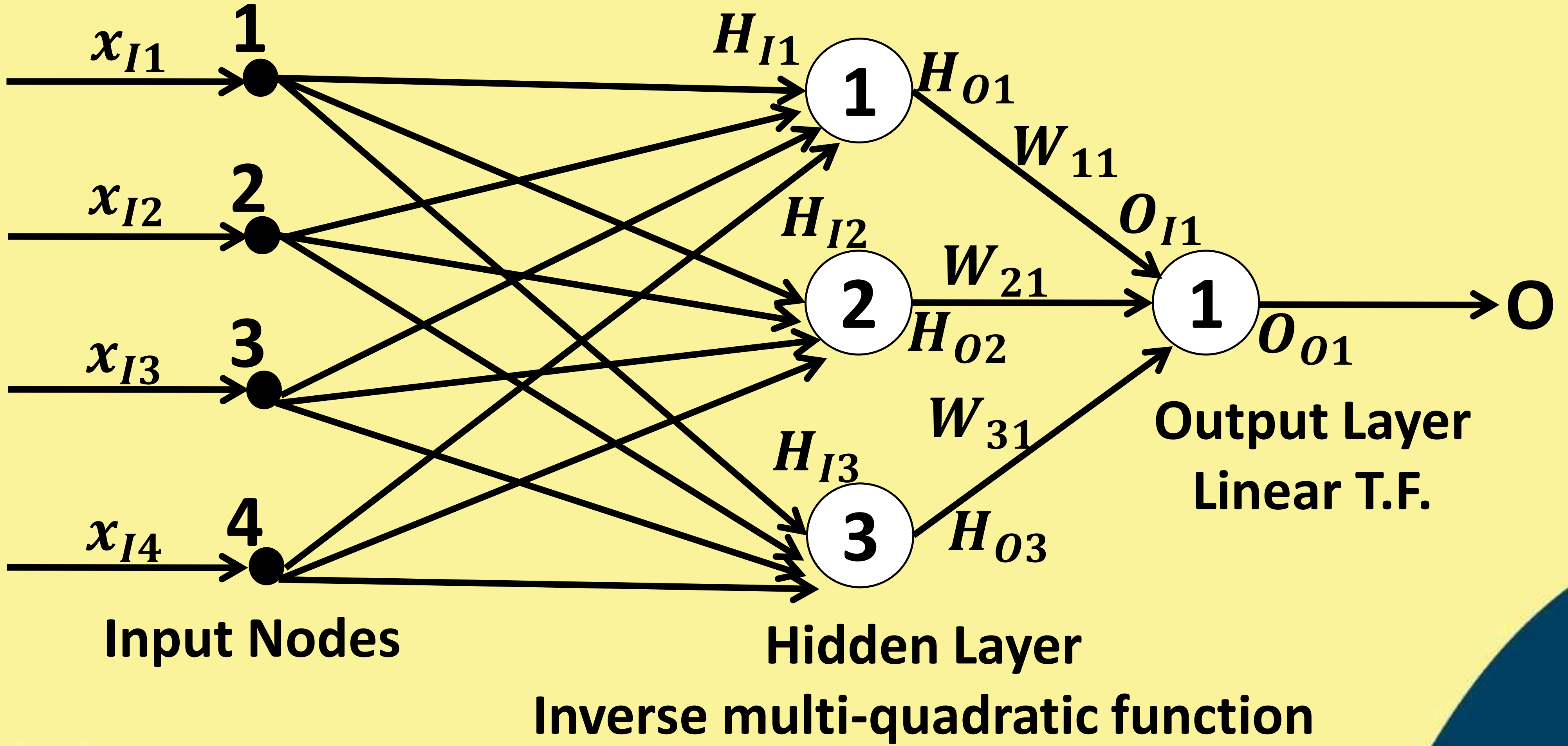
$$\frac{\partial E_k}{\partial O_{Ok}} \cdot \frac{\partial O_{Ok}}{\partial O_{Ik}} = -\left(T_{Ok} - O_{Ok}\right)$$

$$\frac{\partial O_{Ik}}{\partial H_{Oj}} = w_{jk}$$

$$\frac{\partial H_{Oj}}{\partial \sigma_j} = H_{Oj} \left\{ \frac{\left(x_{l1} - \mu_j\right)^2 + \left(x_{l2} - \mu_j\right)^2 + \ldots. + \left(x_{lM} - \mu_j\right)^2}{\sigma_j^3} \right\}$$

Numerical Example

- A radial basis function neural network (RBFNN) is to be used to model input-output relationships of an engineering process having four inputs and one output, as shown in Figure.


x_{I1}
x_{I2}
x_{I3}
x_{I4}
1
2
3
4
H_{I1}
H_{I2}
H_{I3}
H_{O1}
H_{O2}
H_{O3}
W_{11}
W_{21}
W_{31}
O_{I1}
O_{O1}
O
Input Nodes
Hidden Layer
Inverse multi-quadratic function
Output Layer
Linear T.F.

There are three neurons on the hidden layer, which are assumed to have inverse multi-quadratic function of the form $y = f(x) = \frac{1}{\sqrt{x^2+\sigma^2}}$. Take σ_1, σ_2 and σ_3 for the first, second and third hidden neurons as 0.2, 3.0 and 4.0, respectively. Assume initial weights: $w_{11} = 0.2, w_{21} = 0.4$, $w_{31} = 0.5$. Use incremental training scheme with the help of a scenario: $x_{I1} = 1.5, x_{I2} = 2.0, x_{I3} = 1.7, x_{I4} = 2.5$, and output $O = 0.14$. Use back-propagation algorithm with a learning rate of $\eta = 0.2$. Calculate the updated values of w_{11} , σ_1. Show only one iteration.

Solution

Given $x_{I1} = 1.5, x_{I2} = 2.0, x_{I3} = 1.7, x_{I4} = 2.5$

- $H_{I1} = H_{I2} = H_{I3} = 1.5 + 2.0 + 1.7 + 2.5 = 7.7$
- $H_{O1} = \frac{1}{\sqrt{x^2 + {\sigma_1}^2}} = \frac{1}{\sqrt{7.7^2 + (0.2)^2}} = 0.129$
- $H_{O2} = \frac{1}{\sqrt{x^2 + {\sigma_2}^2}} = \frac{1}{\sqrt{7.7^2 + (3.0)^2}} = 0.121$

- $H_{O3} = \frac{1}{\sqrt{x^2+{\sigma_3}^2}} = \frac{1}{\sqrt{7.7^2+(4.0)^2}} = 0.115$

- $O_{I1} = H_{O1} \times w_{11} + H_{O2} \times w_{21} + H_{O3} \times w_{31}$

$= 0.129 \times 0.2 + 0.121 \times 0.4 + 0.115 \times 0.5 = 0.1317$

$\therefore\ O_{O1} = 0.1317$

Now, w_{11} (updated) $= w_{11}$ (previous) + Δw_{11}

- $\Delta w_{11} = -\eta \frac{\partial E}{\partial w_{11}}$

Now, $\frac{\partial E}{\partial w_{11}} = \frac{\partial E}{\partial O_{O1}} \times \frac{\partial O_{O1}}{\partial O_{I1}} \times \frac{\partial O_{I1}}{\partial w_{11}}$

$= -(T_{O1} - O_{O1}) \times 1 \times H_{O1}$

$= -(0.14 - 0.1317) \times 1 \times 0.129 = -0.00107$

$\therefore\ \Delta \mathbf{w}_{11} = -\mathbf{0.2} \times (-\mathbf{0.00107}) = \mathbf{0.000214}$

$\mathbf{w}_{11}$ **(updated)** $= \mathbf{0.2} + \mathbf{0.000214} = \mathbf{0.200214}$

Similarly, $\mathbf{w}_{21}$ **(updated) and** $\mathbf{w}_{31}$ **(updated) can be determined.**

σ_1 (updated) $=\sigma_1$ (previous) + $\Delta\sigma_1$

- $\Delta\sigma_1 = -\eta \frac{\partial E}{\partial \sigma_1}$

Now, $\frac{\partial E}{\partial \sigma_1} == \frac{\partial E}{\partial O_{O1}} \times \frac{\partial O_{O1}}{\partial O_{I1}} \times \frac{\partial O_{I1}}{\partial H_{O1}} \times \frac{\partial H_{O1}}{\partial \sigma_1}$

$$\frac{\partial H_{O1}}{\partial \sigma_1} = \frac{\partial\left(\frac{1}{\sqrt{x^2+\sigma_1{}^2}}\right)}{\partial \sigma_1} = -\frac{1}{2}\left(x^2+\sigma_1{}^2\right)^{-\frac{3}{2}} \times 2\sigma_1$$

$$= -\left(x^2+\sigma_1{}^2\right)^{-\frac{3}{2}} \times \sigma_1$$

Now, $\frac{\partial E}{\partial \sigma_1} = -(T_{O1} - O_{O1}) \times 1 \times w_{11} \times \left\{-\left({H_{I1}}^2 + {\sigma_1}^2\right)^{-\frac{3}{2}} \times \sigma_1\right\}$

$= -(0.14 - 0.1317) \times 1 \times 0.2 \times \left\{-\left(7.7^2 + 0.2^2\right)^{-\frac{3}{2}} \times 0.2\right\}$

$= 7.26 \times 10^{-7}$

$\therefore\ \Delta\sigma_1 = -\eta\frac{\partial E}{\partial \sigma_1} = -0.2 \times 7.26 \times 10^{-7} = -1.45 \times 10^{-7}$

$\therefore$ σ_1 **(updated) =**σ_1 **(previous) +** $\Delta\sigma_1$

$= \mathbf{0.2} - \mathbf{1.45 \times 10^{-7}}$

$= \mathbf{0.199999}$

Similarly, the updated values of σ_2 **and** σ_3 **can be determined.**