RBF Neural Network equations with derivations

Tanmay Khandait
Viral Patel

November 9, 2018

Contents

1 RBFNN Model 1

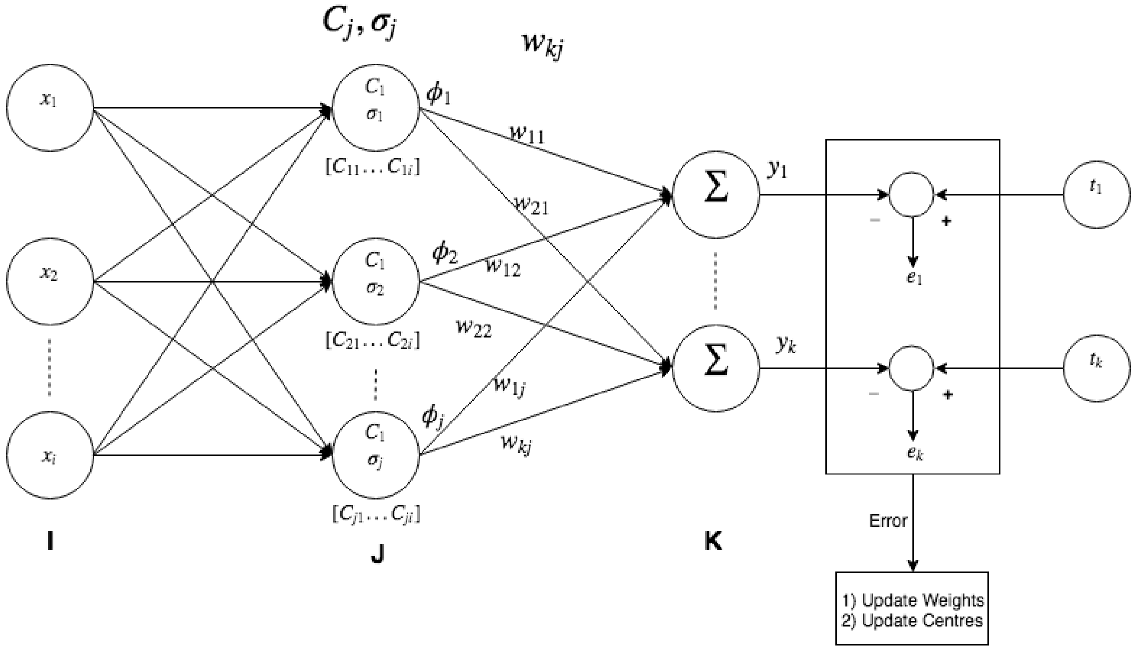

Figure 1: RBFNN with weight and center update

1.1 Key Equations

- $y_k = \sum_j w_{kj} \phi_j$
- $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$
- $z_j = ||X - C_j|| = \sqrt{\sum_i (x_i - c_{ji})^2}$
- Cost function: $E = \frac{1}{2} \sum_k (t_k - y_k)^2$

1.2 Gradient Descent Learning

- $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$
- $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$

1.3 Derivation of $\frac{\partial E}{\partial w_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} \tag{1}$$

From key equations, We know that cost function, $E = \frac{1}{2}\sum_k (t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum_k (t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{2}$$

From key equations, We know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial w_{kj}}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial(\sum_j w_{kj}\phi_j)}{\partial w_{kj}} \\ &= \phi_j\end{aligned} \tag{3}$$

Using equations (2) & (3) we can rewrite equation (1) as:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} \\ &= -(t_k - y_k)\phi_j\end{aligned} \tag{4}$$

1.4 Derivation of $\frac{\partial E}{\partial c_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{ji}}$ as:

$$\frac{\partial E}{\partial c_{ji}} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}} \tag{5}$$

From equation (2), we know that $\frac{\partial E}{\partial y_k} = -(t_k - y_k)$

From key equations, We know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial \phi_j} &= \frac{\partial(\sum_j w_{kj}\phi_j)}{\partial \phi_j} \\ &= w_{kj}\end{aligned} \tag{6}$$

From key equations, We know that $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$, so $\frac{\partial \phi_j}{\partial z_j}$ will be:

$$\begin{aligned}\frac{\partial \phi_j}{\partial z_j} &= \frac{\partial(e^{\frac{-z_j^2}{2\sigma_j^2}})}{\partial z_j} \\ &= -\frac{z_j \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^2} \\ &= -\frac{z_j \times \phi_j}{\sigma_j^2}\end{aligned} \tag{7}$$

From key equations, We know that $z_j = \sqrt{\sum_i (x_i - c_{ji})^2}$, so $\frac{\partial z_j}{\partial c_{ji}}$ will be:

$$
\begin{aligned}
\frac{\partial z_j}{\partial c_{ji}} &= \frac{\partial(\sqrt{\sum_i (x_i - c_{ji})^2})}{\partial c_{ji}} \\
&= -\frac{2 \times (x_i - c_{ji})}{2 \times \sqrt{\sum_i (x_i - c_{ji})^2}} \\
&= -\frac{(x_i - c_{ji})}{z_j}
\end{aligned}
\tag{8}
$$

Using equations (2), (6), (7) & (8) we can rewrite equation (5) as:

$$
\begin{aligned}
\frac{\partial E}{\partial c_{ji}} &= \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}} \\
&= \left[\sum_k -(t_k - y_k) w_{kj}\right] \times \frac{-(z_j \times \phi_j)}{\sigma_j^2} \times \frac{-(x_i - c_{ji})}{z_j} \\
&= \left[\sum_k -(t_k - y_k) w_{kj}\right] \times \frac{\phi_j}{\sigma_j^2} \times (x_i - c_{ji})
\end{aligned}
\tag{9}
$$

1.5 Final equations of $\frac{\partial E}{\partial w_{kj}}$ & $\frac{\partial E}{\partial c_{ji}}$

- $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$
- $\frac{\partial E}{\partial c_{ji}} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}} = \left[\sum_k -(t_k - y_k) w_{kj}\right] \times \frac{\phi_j}{\sigma_j^2} \times (x_i - c_{ji})$

2 RBFNN Model 2

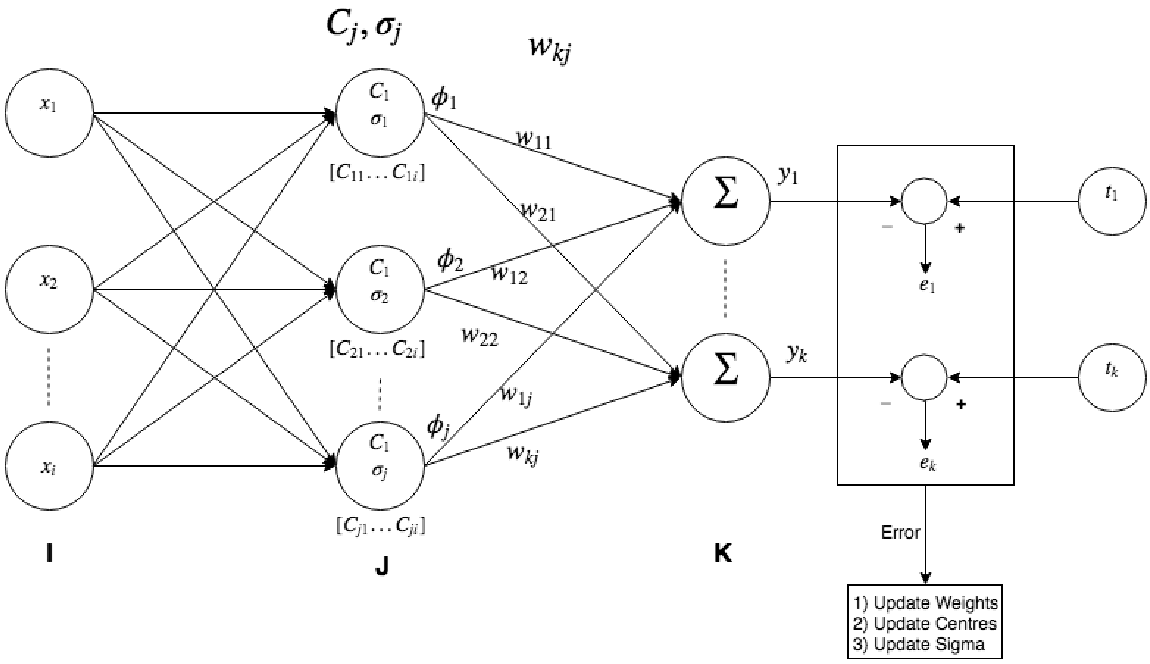

Figure 2: RBFNN with weight, center and sigma update

2.1 Key Equations

- $y_k = \sum\limits_j w_{kj}\phi_j$
- $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$
- $z_j = ||X - C_j|| = \sqrt{\sum\limits_i (x_i - c_{ji})^2}$
- Cost function: $E = \frac{1}{2}\sum\limits_k (t_k - y_k)^2$

2.2 Gradient Descent Learning

- $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$
- $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$
- $\sigma_j(t+1) = \sigma_j(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_j}$

2.3 Derivation of $\frac{\partial E}{\partial w_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{kj}}$ as:

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} \tag{10}$$

From key equations, We know that cost function, $E = \frac{1}{2}\sum_k (t_k - y_k)^2$, so $\frac{\partial E}{\partial y_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial y_k} &= \frac{\partial(\frac{1}{2}\sum_k (t_k - y_k)^2)}{\partial y_k} \\ &= -(t_k - y_k)\end{aligned} \tag{11}$$

From key equations, We know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial w_{kj}}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial(\sum_j w_{kj}\phi_j)}{\partial w_{kj}} \\ &= \phi_j\end{aligned} \tag{12}$$

Using equations (11) & (12) we can rewrite equation (10) as:

$$\begin{aligned}\frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} \\ &= -(t_k - y_k)\phi_j\end{aligned} \tag{13}$$

2.4 Derivation of $\frac{\partial E}{\partial c_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{ji}}$ as:

$$\frac{\partial E}{\partial c_{ji}} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}} \tag{14}$$

From equation (11), we know that $\frac{\partial E}{\partial y_k} = -(t_k - y_k)$

From key equations, We know that $y_k = \sum_j w_{kj}\phi_j$, so $\frac{\partial y_k}{\partial \phi_j}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial \phi_j} &= \frac{\partial(\sum_j w_{kj}\phi_j)}{\partial \phi_j} \\ &= w_{kj}\end{aligned} \tag{15}$$

From key equations, We know that $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$, so $\frac{\partial \phi_j}{\partial z_j}$ will be:

$$\begin{aligned}\frac{\partial \phi_j}{\partial z_j} &= \frac{\partial(e^{\frac{-z_j^2}{2\sigma_j^2}})}{\partial z_j} \\ &= -\frac{z_j \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^2} \\ &= -\frac{z_j \times \phi_j}{\sigma_j^2}\end{aligned} \tag{16}$$

From key equations, We know that $z_j = \sqrt{\sum_i (x_i - c_{ji})^2}$, so $\frac{\partial z_j}{\partial c_{ji}}$ will be:

$$\begin{aligned} \frac{\partial z_j}{\partial c_{ji}} &= \frac{\partial(\sqrt{\sum_i (x_i - c_{ji})^2})}{\partial c_{ji}} \\ &= -\frac{2 \times (x_i - c_{ji})}{2 \times \sqrt{\sum_i (x_i - c_{ji})^2}} \\ &= -\frac{(x_i - c_{ji})}{z_j} \end{aligned} \tag{17}$$

Using equations (11), (15), (16) & (17) we can rewrite equation (14) as:

$$\begin{aligned} \frac{\partial E}{\partial c_{ji}} &= \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j} \right] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}} \\ &= \left[\sum_k -(t_k - y_k) w_{kj} \right] \times \frac{-(z_j \times \phi_j)}{\sigma_j^2} \times \frac{-(x_i - c_{ji})}{z_j} \\ &= \left[\sum_k -(t_k - y_k) w_{kj} \right] \times \frac{\phi_j}{\sigma_j^2} \times (x_i - c_{ji}) \end{aligned} \tag{18}$$

2.5 Derivation of $\frac{\partial E}{\partial \sigma_j}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial \sigma_j}$ as:

$$\frac{\partial E}{\partial \sigma_j} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j} \right] \cdot \frac{\partial \phi_j}{\partial \sigma_j} \tag{19}$$

From equation (11), we know that $\frac{\partial E}{\partial y_k} = -(t_k - y_k)$

From equation (15), we know that $\frac{\partial y_k}{\partial \phi_j} = w_{kj}$

From key equations, We know that $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$, so $\frac{\partial \phi_j}{\partial \sigma_j}$ will be:

$$\begin{aligned} \frac{\partial \phi_j}{\partial \sigma_j} &= \frac{\partial(e^{\frac{-z_j^2}{2\sigma_j^2}})}{\partial \sigma_j} \\ &= \frac{z_j^2 \times e^{\frac{-z_j^2}{2\sigma_j^2}}}{\sigma_j^3} \\ &= \frac{z_j^2 \times \phi_j}{\sigma_j^3} \end{aligned} \tag{20}$$

Using equations (11), (15) & (20) we can rewrite equation (19) as:

$$\begin{aligned} \frac{\partial E}{\partial c_{ji}} &= \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j} \right] \cdot \frac{\partial \phi_j}{\partial \sigma_j} \\ &= \left[\sum_k -(t_k - y_k) w_{kj} \right] \times \frac{(z_j^2 \times \phi_j)}{\sigma_j^3} \end{aligned} \tag{21}$$

2.6 Final equations of $\frac{\partial E}{\partial w_{kj}}$, $\frac{\partial E}{\partial c_{ji}}$ & $\frac{\partial E}{\partial \sigma_j}$

- $\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$

- $\frac{\partial E}{\partial c_{ji}} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}} = \left[\sum_k -(t_k - y_k)w_{kj}\right] \times \frac{\phi_j}{\sigma_j^2} \times (x_i - c_{ji})$

- $\frac{\partial E}{\partial \sigma_j} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial \sigma_j} = \left[\sum_k -(t_k - y_k)w_{kj}\right] \times \frac{z_j^2 \times \phi_j}{\sigma_j^3}$

3 RBFNN Model 3

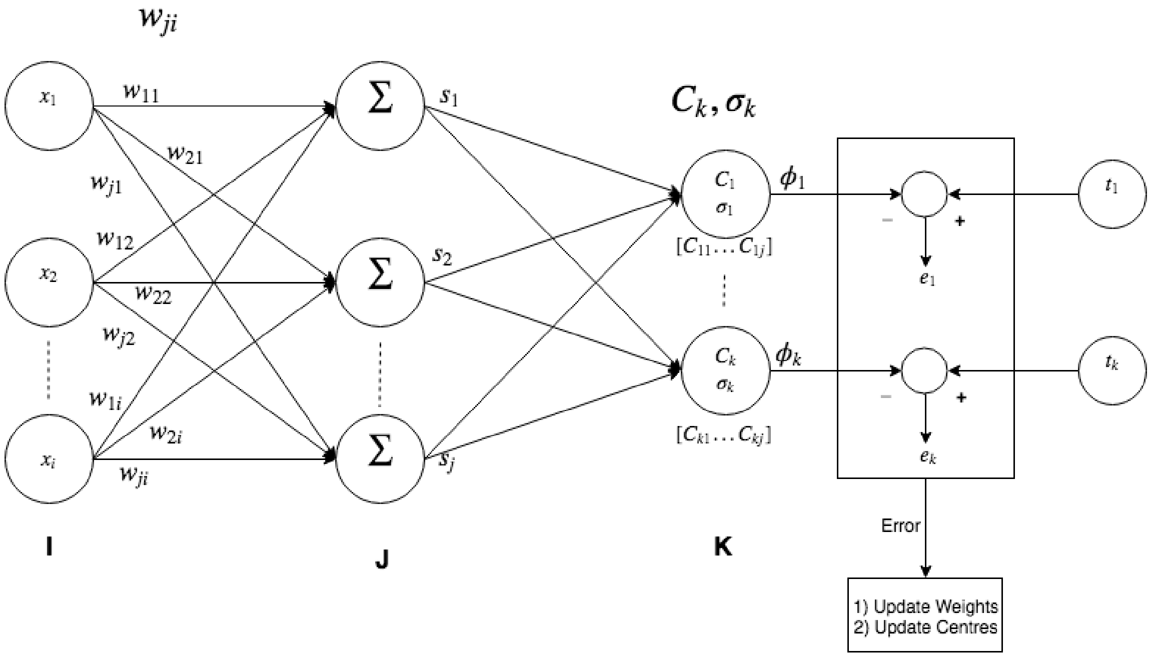

Figure 3: Reversed RBFNN with weight and center update

3.1 Key Equations

- $s_j = \sum_i w_{ji} x_i$
- $\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$
- $z_k = ||S - C_k|| = \sqrt{\sum_j (s_j - c_{kj})^2}$
- Cost function: $E = \frac{1}{2} \sum_k (t_k - \phi_k)^2$

3.2 Gradient Descent Learning

- $w_{ji}(t+1) = w_{ji}(t) - \eta_w \frac{\partial E}{\partial w_{ji}}$
- $c_{kj}(t+1) = c_{kj}(t) - \eta_c \frac{\partial E}{\partial c_{kj}}$

3.3 Derivation of $\frac{\partial E}{\partial c_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{kj}}$ as:

$$\frac{\partial E}{\partial c_{kj}} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}} \tag{22}$$

From key equations, We know that cost function, $E = \frac{1}{2} \sum_k (t_k - \phi_k)^2$, so $\frac{\partial E}{\partial \phi_k}$ will be:

$$\begin{aligned} \frac{\partial E}{\partial \phi_k} &= \frac{\partial (\frac{1}{2} \sum_k (t_k - \phi_k)^2)}{\partial \phi_k} \\ &= -(t_k - \phi_k) \end{aligned} \tag{23}$$

From key equations, We know that $\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$, so $\frac{\partial \phi_k}{\partial z_k}$ will be:

$$\begin{aligned}\frac{\partial \phi_k}{\partial z_k} &= \frac{\partial (e^{\frac{-z_k^2}{2\sigma_k^2}})}{\partial z_k} \\ &= -\frac{z_k \times e^{\frac{-z_k^2}{2\sigma_k^2}}}{\sigma_k^2} \\ &= -\frac{z_k \times \phi_k}{\sigma_k^2}\end{aligned} \tag{24}$$

From key equations, We know that $z_k = \sqrt{\sum_j (s_j - c_{kj})^2}$, so $\frac{\partial z_k}{\partial c_{kj}}$ will be:

$$\begin{aligned}\frac{\partial z_k}{\partial c_{kj}} &= \frac{\partial (\sqrt{\sum_j (s_j - c_{kj})^2})}{\partial c_{kj}} \\ &= -\frac{2 \times (s_j - c_{kj})}{2 \times \sqrt{\sum_j (s_j - c_{kj})^2}} \\ &= -\frac{s_j - c_{kj}}{z_k}\end{aligned} \tag{25}$$

Using equations (23), (24) & (25) we can rewrite equation (22) as:

$$\begin{aligned}\frac{\partial E}{\partial c_{kj}} &= \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}} \\ &= -(t_k - \phi_k) \times \frac{-(z_k \times \phi_k)}{\sigma_k^2} \times \frac{-(s_j - c_{kj})}{z_k} \\ &= -(t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})\end{aligned} \tag{26}$$

3.4 Derivation of $\frac{\partial E}{\partial w_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{ji}}$ as:

$$\frac{\partial E}{\partial w_{ji}} = \left[\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j}\right] \cdot \frac{\partial s_j}{\partial w_{ji}} \tag{27}$$

From equation (23) we know that $\frac{\partial E}{\partial \phi_k} = -(t_k - \phi_k)$

From equation (24) we know that $\frac{\partial \phi_k}{\partial z_k} = -\frac{z_k \times \phi_k}{\sigma_k^2}$

From key equations, We know that $z_k = \sqrt{\sum_j (s_j - c_{kj})^2}$, so $\frac{\partial z_k}{\partial s_j}$ will be:

$$\begin{aligned}\frac{\partial z_k}{\partial s_j} &= \frac{\partial (\sqrt{\sum_j (s_j - c_{kj})^2})}{\partial s_j} \\ &= \frac{2 \times (s_j - c_{kj})}{2 \times \sqrt{\sum_j (s_j - c_{kj})^2}} \\ &= \frac{s_j - c_{kj}}{z_k}\end{aligned} \tag{28}$$

From key equations, We know that $s_j = \sum_i w_{ji} x_i$, so $\frac{\partial s_j}{\partial w_{ji}}$ will be:

$$\begin{aligned}\frac{\partial s_j}{\partial w_{ji}} &= \frac{\partial(\sum_i w_{ji} x_i)}{\partial w_{ji}} \\ &= x_i\end{aligned} \tag{29}$$

Using equations (23), (24), (28) & (29), we can rewrite equation (27) as:

$$\begin{aligned}\frac{\partial E}{\partial w_{ji}} &= \left[\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j}\right] \cdot \frac{\partial s_j}{\partial w_{ji}} \\ &= \Big[\sum_k -(t_k - \phi_k) \times \frac{-(z_k \times \phi_k)}{\sigma_k^2} \times \frac{(s_j - c_{kj})}{z_k}\Big] \times x_i \\ &= \Big[\sum_k (t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})\Big] \times x_i\end{aligned} \tag{30}$$

3.5 Final equations of $\frac{\partial E}{\partial w_{ji}}$ & $\frac{\partial E}{\partial c_{kj}}$

- $\frac{\partial E}{\partial c_{kj}} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}} = -(t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})$
- $\frac{\partial E}{\partial w_{ji}} = \left[\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j}\right] \cdot \frac{\partial s_j}{\partial w_{ji}} = \Big[\sum_k (t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})\Big] \times x_i$

4 RBFNN Model 4

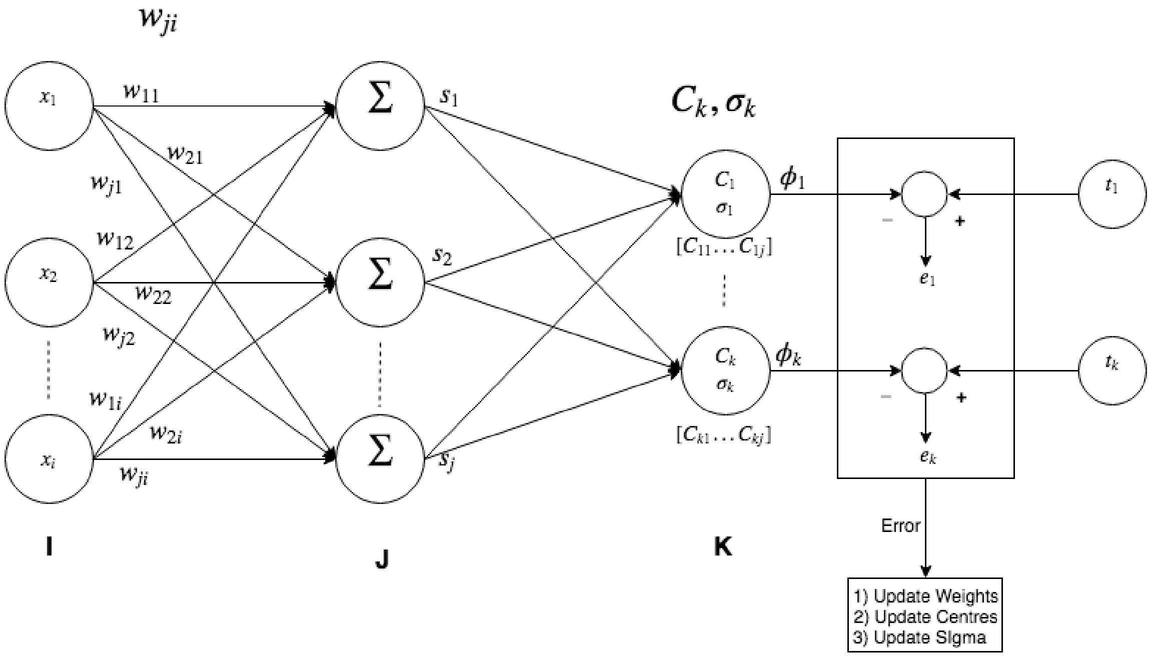

Figure 4: Reversed RBFNN with weight, center and sigma update

4.1 Key Equations

- $s_j = \sum_i w_{ji} x_i$
- $\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$
- $z_k = ||S - C_k|| = \sqrt{\sum_j (s_j - c_{kj})^2}$
- Cost function: $E = \frac{1}{2} \sum_k (t_k - \phi_k)^2$

4.2 Gradient Descent Learning

- $w_{ji}(t+1) = w_{ji}(t) - \eta_w \frac{\partial E}{\partial w_{ji}}$
- $c_{kj}(t+1) = c_{kj}(t) - \eta_c \frac{\partial E}{\partial c_{kj}}$
- $\sigma_k(t+1) = \sigma_k(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_k}$

4.3 Derivation of $\frac{\partial E}{\partial c_{kj}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial c_{kj}}$ as:

$$\frac{\partial E}{\partial c_{kj}} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}} \tag{31}$$

From key equations, We know that cost function, $E = \frac{1}{2}\sum_k (t_k - \phi_k)^2$, so $\frac{\partial E}{\partial \phi_k}$ will be:

$$\begin{aligned}\frac{\partial E}{\partial \phi_k} &= \frac{\partial(\frac{1}{2}\sum_k (t_k - \phi_k)^2)}{\partial \phi_k} \\ &= -(t_k - \phi_k)\end{aligned} \tag{32}$$

From key equations, We know that $\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$, so $\frac{\partial \phi_k}{\partial z_k}$ will be:

$$\begin{aligned}\frac{\partial \phi_k}{\partial z_k} &= \frac{\partial(e^{\frac{-z_k^2}{2\sigma_k^2}})}{\partial z_k} \\ &= -\frac{z_k \times e^{\frac{-z_k^2}{2\sigma_k^2}}}{\sigma_k^2} \\ &= -\frac{z_k \times \phi_k}{\sigma_k^2}\end{aligned} \tag{33}$$

From key equations, We know that $z_k = \sqrt{\sum_j (s_j - c_{kj})^2}$, so $\frac{\partial z_k}{\partial c_{kj}}$ will be:

$$\begin{aligned}\frac{\partial z_k}{\partial c_{kj}} &= \frac{\partial(\sqrt{\sum_j (s_j - c_{kj})^2})}{\partial c_{kj}} \\ &= -\frac{2 \times (s_j - c_{kj})}{2 \times \sqrt{\sum_j (s_j - c_{kj})^2}} \\ &= -\frac{s_j - c_{kj}}{z_k}\end{aligned} \tag{34}$$

Using equations (32), (33) & (34) we can rewrite equation (31) as:

$$\begin{aligned}\frac{\partial E}{\partial c_{kj}} &= \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}} \\ &= -(t_k - \phi_k) \times \frac{-(z_k \times \phi_k)}{\sigma_k^2} \times \frac{-(s_j - c_{kj})}{z_k} \\ &= -(t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})\end{aligned} \tag{35}$$

4.4 Derivation of $\frac{\partial E}{\partial w_{ji}}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial w_{ji}}$ as:

$$\frac{\partial E}{\partial w_{ji}} = \left[\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j}\right] \cdot \frac{\partial s_j}{\partial w_{ji}} \tag{36}$$

From equation (32) we know that $\frac{\partial E}{\partial \phi_k} = -(t_k - \phi_k)$

From equation (33) we know that $\frac{\partial \phi_k}{\partial z_k} = -\frac{z_k \times \phi_k}{\sigma_k^2}$

From key equations, We know that $z_k = \sqrt{\sum_j (s_j - c_{kj})^2}$, so $\frac{\partial z_k}{\partial s_j}$ will be:

$$\begin{aligned} \frac{\partial z_k}{\partial s_j} &= \frac{\partial(\sqrt{\sum_j (s_j - c_{kj})^2})}{\partial s_j} \\ &= \frac{2 \times (s_j - c_{kj})}{2 \times \sqrt{\sum_j (s_j - c_{kj})^2}} \\ &= \frac{s_j - c_{kj}}{z_k} \end{aligned} \tag{37}$$

From key equations, We know that $s_j = \sum_i w_{ji} x_i$, so $\frac{\partial s_j}{\partial w_{ji}}$ will be:

$$\begin{aligned} \frac{\partial s_j}{\partial w_{ji}} &= \frac{\partial(\sum_i w_{ji} x_i)}{\partial w_{ji}} \\ &= x_i \end{aligned} \tag{38}$$

Using equations (32), (33), (37) & (38), we can rewrite equation (36) as:

$$\begin{aligned} \frac{\partial E}{\partial w_{ji}} &= \left[\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j} \right] \cdot \frac{\partial s_j}{\partial w_{ji}} \\ &= \Big[\sum_k -(t_k - \phi_k) \times \frac{-(z_k \times \phi_k)}{\sigma_k^2} \times \frac{(s_j - c_{kj})}{z_k} \Big] \times x_i \\ &= \Big[\sum_k (t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj}) \Big] \times x_i \end{aligned} \tag{39}$$

4.5 Derivation of $\frac{\partial E}{\partial \sigma_k}$

Using chain rule, we can rewrite $\frac{\partial E}{\partial \sigma_k}$ as:

$$\frac{\partial E}{\partial \sigma_k} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial \sigma_k} \tag{40}$$

From equation (32) we know that $\frac{\partial E}{\partial \phi_k} = -(t_k - \phi_k)$

From key equations, We know that $\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$, so $\frac{\partial \phi_k}{\partial \sigma_k}$ will be:

$$\begin{aligned} \frac{\partial \phi_k}{\partial \sigma_k} &= \frac{\partial(e^{\frac{-z_k^2}{2\sigma_k^2}})}{\partial \sigma_k} \\ &= \frac{z_k^2 \times e^{\frac{-z_k^2}{2\sigma_k^2}}}{\sigma_k^3} \\ &= \frac{z_k^2 \times \phi_k}{\sigma_k^3} \end{aligned} \tag{41}$$

Using equations (32) & (41), we can rewrite equation (40) as:

$$\begin{aligned} \frac{\partial E}{\partial \sigma_k} &= \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial \sigma_k} \\ &= -(t_k - \phi_k) \times \frac{z_k^2 \times \phi_k}{\sigma_k^3} \end{aligned} \tag{42}$$

4.6 Final equations of $\frac{\partial E}{\partial w_{ji}}$, $\frac{\partial E}{\partial c_{kj}}$ & $\frac{\partial E}{\partial \sigma_k}$

- $\frac{\partial E}{\partial c_{kj}} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}} = -(t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})$
- $\frac{\partial E}{\partial w_{ji}} = \left[\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j} \right] \cdot \frac{\partial s_j}{\partial w_{ji}} = \left[\sum_k (t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj}) \right] \times x_i$
- $\frac{\partial E}{\partial \sigma_k} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial \sigma_k} = -(t_k - \phi_k) \times \frac{\phi_k \times z_k^2}{\sigma_k^3}$

5 Table of equations

Table 1: Key equation, and back-propagation derivation results of all RBFNN models

	Key Equations	Gradient Descent Learning	Computation of $\frac{\partial E}{\partial w}$, $\frac{\partial E}{\partial c}$ & $\frac{\partial E}{\partial \sigma}$
Model-1	$y_k = \sum_j w_{kj}\phi_j$ $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$ $z_j = \|\|X - C_j\|\| = \sqrt{\sum_i (x_i - c_{ji})^2}$ Cost function: $E = \frac{1}{2}\sum_k (t_k - y_k)^2$	Learning of weights: $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$ Learning of centers: $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$	$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$ $\frac{\partial E}{\partial c_{ji}} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}}$ $= \left[\sum_k -(t_k - y_k)w_{kj}\right] \times \frac{\phi_j}{\sigma_j^2} \times (x_i - c_{ji})$
Model-2	$y_k = \sum_j w_{kj}\phi_j$ $\phi_j = e^{\frac{-z_j^2}{2\sigma_j^2}}$ $z_j = \|\|X - C_j\|\| = \sqrt{\sum_i (x_i - c_{ji})^2}$ Cost function: $E = \frac{1}{2}\sum_k (t_k - y_k)^2$	Learning of weights: $w_{kj}(t+1) = w_{kj}(t) - \eta_w \frac{\partial E}{\partial w_{kj}}$ Learning of centers: $c_{ji}(t+1) = c_{ji}(t) - \eta_c \frac{\partial E}{\partial c_{ji}}$ Learning of sigma: $\sigma_j(t+1) = \sigma_j(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_j}$	$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial w_{kj}} = -(t_k - y_k)\phi_j$ $\frac{\partial E}{\partial c_{ji}} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial c_{ji}}$ $= \left[\sum_k -(t_k - y_k)w_{kj}\right] \times \frac{\phi_j}{\sigma_j^2} \times (x_i - c_{ji})$ $\frac{\partial E}{\partial \sigma_j} = \left[\sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \phi_j}\right] \cdot \frac{\partial \phi_j}{\partial \sigma_j}$ $= \left[\sum_k -(t_k - y_k)w_{kj}\right] \times \frac{\phi_j \times z_j^2}{\sigma_j^3}$
Model-3	$s_j = \sum_i w_{ji}x_i$ $\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$ $z_k = \|\|S - C_k\|\| = \sqrt{\sum_j (s_j - c_{kj})^2}$ Cost function: $E = \frac{1}{2}\sum_k (t_k - \phi_k)^2$	Learning of weights: $w_{ji}(t+1) = w_{ji}(t) - \eta_w \frac{\partial E}{\partial w_{ji}}$ Learning of centers: $c_{kj}(t+1) = c_{kj}(t) - \eta_c \frac{\partial E}{\partial c_{kj}}$	$\frac{\partial E}{\partial w_{ji}} = \left[\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j}\right] \cdot \frac{\partial s_j}{\partial w_{ji}}$ $= \left[\sum_k (t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})\right] \times x_i$ $\frac{\partial E}{\partial c_{kj}} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}}$ $= -(t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})$
Model-4	$s_j = \sum_i w_{ji}x_i$ $\phi_k = e^{\frac{-z_k^2}{2\sigma_k^2}}$ $z_k = \|\|S - C_k\|\| = \sqrt{\sum_j (s_j - c_{kj})^2}$ Cost function: $E = \frac{1}{2}\sum_k (t_k - \phi_k)^2$	Learning of weights: $w_{ji}(t+1) = w_{ji}(t) - \eta_w \frac{\partial E}{\partial w_{ji}}$ Learning of centers: $c_{kj}(t+1) = c_{kj}(t) - \eta_c \frac{\partial E}{\partial c_{kj}}$ Learning of sigma: $\sigma_k(t+1) = \sigma_k(t) - \eta_\sigma \frac{\partial E}{\partial \sigma_k}$	$\frac{\partial E}{\partial w_{ji}} = \left[\sum_k \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial s_j}\right] \cdot \frac{\partial s_j}{\partial w_{ji}}$ $= \left[\sum_k (t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})\right] \times x_i$ $\frac{\partial E}{\partial c_{kj}} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial c_{kj}}$ $= -(t_k - \phi_k) \times \frac{\phi_k}{\sigma_k^2} \times (s_j - c_{kj})$ $\frac{\partial E}{\partial \sigma_k} = \frac{\partial E}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial \sigma_k}$ $= -(t_k - \phi_k) \times \frac{\phi_k \times z_k^2}{\sigma_k^3}$