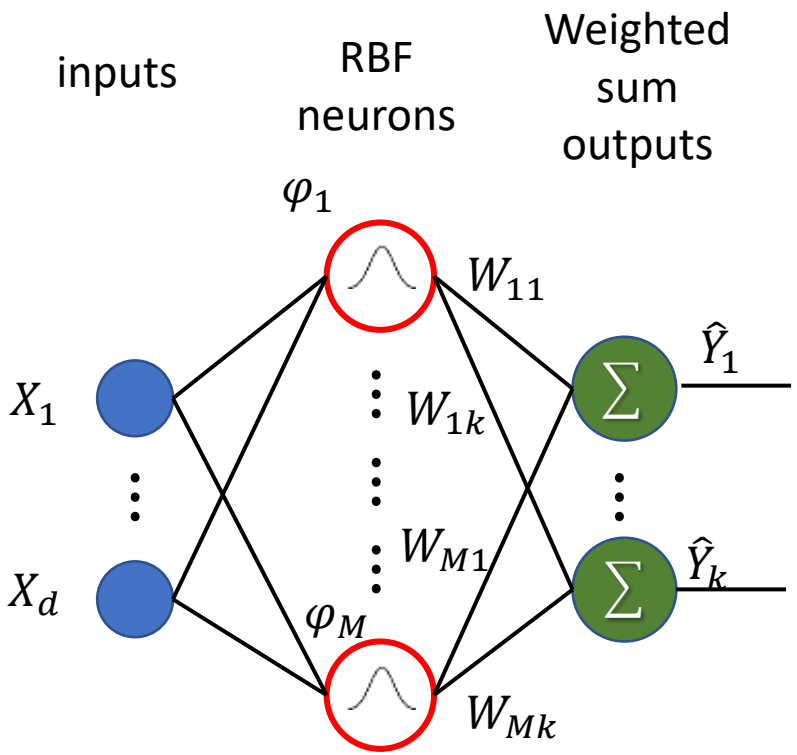
Lecture 9

Radial Basis Function Neural Network (RBF NN)

Radial Basis Function Neural Network

- Radial Basis Function Neural Network (RBF) NN is a type of neural network with only one hidden layer in addition to an input and output layer.

- Each node within the hidden layer uses a radial basis activation function.



Network Structure

1. Input:

$$\left\{(\mathbf{x}_i, y_i)\right\}_{i=1}^{n}, \text{ where } \mathbf{x}_i \subset \mathbb{R}^d$$

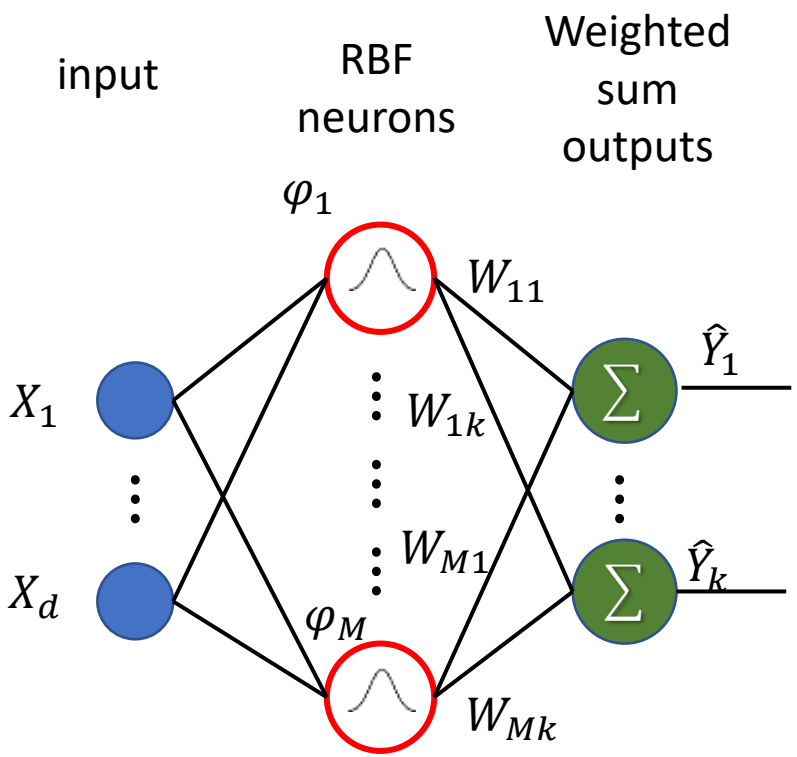
2. Basis functions:

   There are many choices for the basis function.
   The commonly used is radial basis.

$$\phi_j(\mathbf{x}_i) = e^{-|\mathbf{x}_i - \mu_j|^2}$$

3. Output:

$$y_k(\mathbf{x}) = \sum_{j=1}^{m} W_{jk}\phi_j(\mathbf{x})$$



Matrix Form Representation

- Alternatively we can use matrix-vectorial notation:

Inputs:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{d1} & \cdots & x_{dn} \end{bmatrix}_{d \times n} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ \vdots & \vdots & & \vdots \end{bmatrix}_{d \times n}$$

Weights:

$$W = \begin{bmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mk} \end{bmatrix}_{m \times k}$$

Radial Basis Functions:

$$\Phi = \begin{bmatrix} \phi_1(\mathbf{x}_1) & \cdots & \phi_1(\mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ \phi_m(\mathbf{x}_1) & \cdots & \phi_m(\mathbf{x}_n) \end{bmatrix}_{m \times n}$$

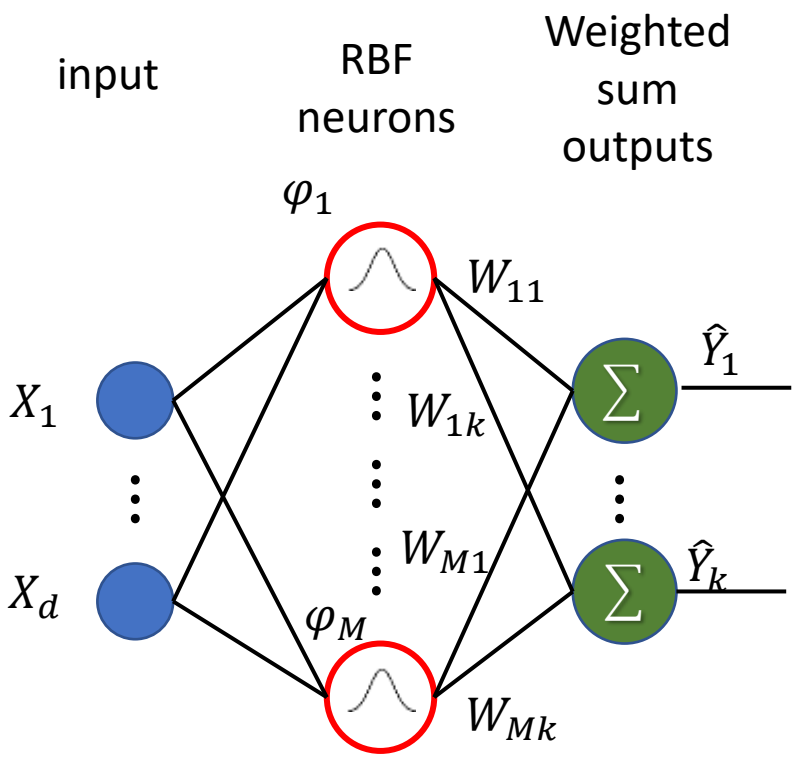
Outputs:

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{k1} & \cdots & y_{kn} \end{bmatrix}_{k \times n}$$

- The output will be:

$$Y = W^T \Phi$$

where Y and Φ are known while W is unknown.



Optimization

- List square objective function can be used for computing W

$$\psi = \left\| Y - W^T \Phi \right\|^2$$

- W can be computed by minimizing our objective function w.r.t W.

$$\min_W \left\| Y - W^T \Phi \right\|^2$$

Close Form Derivation

- This optimization problem can be solved in close form:

$$\frac{\partial}{\partial W}\left\|Y - W^T\Phi\right\|^2 = \frac{\partial}{\partial W} Tr[(Y - W^T\Phi)^T(Y - W^T\Phi)]$$

$$= \frac{\partial}{\partial W} Tr[Y^TY + \Phi^TWW^T\Phi - Y^TW^T\Phi - \Phi^TWY]$$

$$= 0 + 2\Phi\Phi^TW - 2\Phi Y^T = 0$$

$$\Rightarrow \Phi\Phi^TW = \Phi Y^T$$

$$\Rightarrow \boxed{W = (\Phi\Phi^T)^{-1}\Phi Y^T}$$

- Therefore the solution is:

$$W = (\Phi\Phi^T)^{-1}\Phi Y^T$$