

REDES RBF (RADIAL BASIS FUNCTION)

Prof. Dr. Ajalmar Rocha

Disciplina: Inteligência Computacional Aplicada (ICA)

Programa de Pós-Graduação em Eng. de Telecomunicações (PPGET)
Instituto Federal do Ceará (IFCE)

Setembro/2013

Introdução

- Uma rede RBF (*Radial Basis Function*) é uma rede neural artificial com apenas uma camada oculta de *funções de base* não lineares.
- A camada de saída da rede RBF normalmente possui funções de ativação não-lineares para resolução de problemas de classificações de padrões e saídas lineares para aproximação de função.
- A rede RBF, assim como a rede MLP, também é um aproximador universal de funções.
- O processo de treinamento da rede RBF é baseado em um problema de ajuste de curva em um espaço de alta dimensionalidade.

Arquitetura da rede RBF

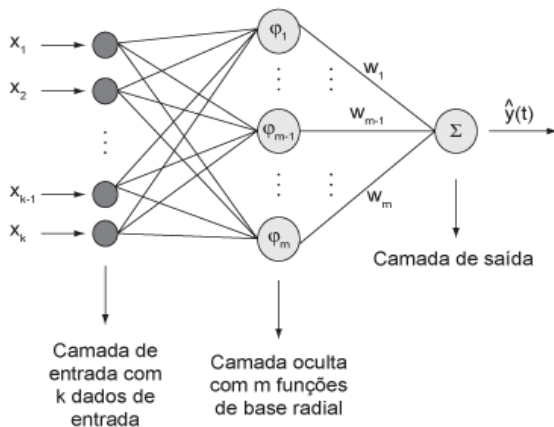


Figure: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2008000300016

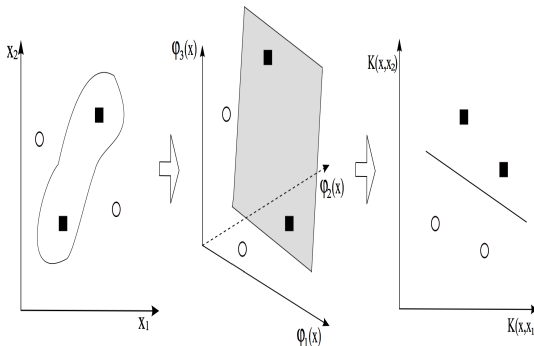
Arquitetura da rede RBF

- Uma rede RBF contém três camadas com papéis totalmente distintos.
 - 1 A camada de entrada é constituída por nós de fonte que conectam a rede ao ambiente (sensores);
 - 2 A camada escondida aplica uma transformação não-linear do espaço de entrada para espaço oculto (de alta dimensão, muitas vezes chamada espaço característico).
 - 3 A camada de saída é normalmente linear (atuador).
- A justificativa para uso de uma transformação não-linear remonta a um artigo de Cover (1965).

Cover, T.M. "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition", IEEE Transactions on Electronic Computers, 1965.

Cover (1965)

- No trabalho de Cover(1965) é mostrado que um problema de classificação de padrões disposto em um espaço de alta dimensionalidade tem maior probabilidade de ser linearmente separável do que em um espaço de baixa dimensionalidade.



Interpolação

- Como dito anteriormente, a rede RBF é projetada para realizar um mapeamento não-linear do espaço de entrada para o espaço oculto, $\mathbb{R}^p \rightarrow \mathbb{R}^q$; seguido de um mapeamento linear (ou não linear) para o espaço de saída, $\mathbb{R}^q \rightarrow \mathbb{R}^1$. Os termos p e q representam o tamanho do espaço de entrada e q o número de neurônios na camada escondida.
- De uma forma global, a rede representa um mapeamento de um espaço de dimensionalidade p para um espaço de saída unidimensional^a, a saber.

$$s : \mathbb{R}^p \rightarrow \mathbb{R}^1 \quad (1)$$

em que o mapa s é uma hipersuperfície $\Gamma \subset \mathbb{R}^{p+1}$.

^aAo se considerar apenas uma saída na rede.

Interpolação

- Em uma situação real, a superfície Γ é desconhecida e os dados de treinamento estão normalmente contaminados com ruído.
- A fase de treinamento e generalização do processo de aprendizagem podem ser vistos como:
 - 1 A fase de treinamento constitui a otimização de um procedimento de ajuste para a superfície Γ , baseada nos padrões de treinamento.
 - 2 A fase de generalização é o sinônimo da **interpolação** entre os pontos de dados, a fim de aproximar a superfície Γ .
- Na interpolação estrita a superfície Γ deve passar por todos os pontos dos dados de treinamento.

Problema de Interpolação

Dado um conjunto de N pontos diferentes $\{\mathbf{x}_i \in \mathbb{R}^p\}_{i=1}^N$ e um conjunto correspondente de N números reais $\{d_i \in \mathbb{R}^1\}_{i=1}^N$, encontre uma função $F : \mathbb{R}^p \rightarrow \mathbb{R}^1$ que satisfaça a condição de interpolação:

$$F(\mathbf{x}_i) = d_i, \quad i = 1, 2, \dots, N \quad (2)$$

Funções de Base Radial (RBFs)

Funções de Base Radial

A técnica de funções de base radial consiste em escolher uma função F que tenha a seguinte forma

$$F(\mathbf{x}) = \sum_{i=1}^N w_i \phi(\|\mathbf{x} - \mathbf{x}_i\|). \quad (3)$$

em que $\phi_i(\|\mathbf{x} - \mathbf{x}_i\|)$ representa uma função não-linear arbitrária, chamada de função de base radial, tal que $\|\cdot\|$ representa uma norma que usualmente é a euclidiana. Neste contexto, os pontos de dados \mathbf{x}_i são tomados como centros das funções de base radial.

Funções de Base Radial (RBFs)

Funções de Base Radial - Gaussiana

- A variável de ativação do i -ésimo neurônio oculto da rede RBF é dada por

$$u_i = \|\mathbf{x} - \mathbf{x}_i\| = \sqrt{x_1x_{i1} + x_2x_{i2} + \cdots + x_px_{ip}} \quad (4)$$

- Enquanto a variável de saída pode ser calculada como

$$\phi(u_i) = \phi(\|\mathbf{x} - \mathbf{x}_i\|) \quad (5)$$

$$\phi(u_i) = \exp\left(\frac{u_i^2}{2\sigma^2}\right)$$

$$\phi(u_i) = \exp\left(\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2\sigma^2}\right)$$

em que σ é o raio (*spread*) de abertura da função de ativação.

RBF Gaussian Unidimensional

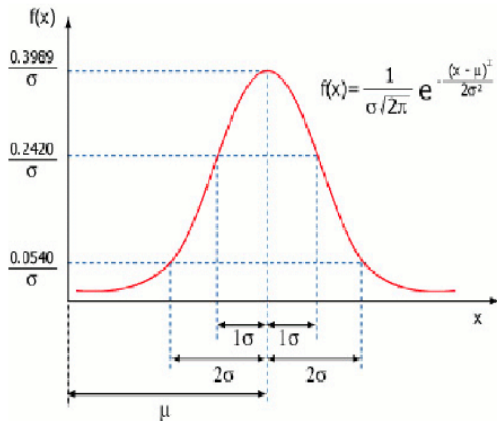


Figure: [http : // www.deti.ufc.br/ guilherme](http://www.deti.ufc.br/guilherme)

RBF Gaussian Bidimensional

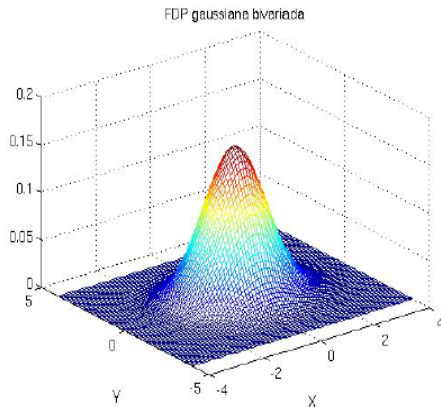


Figure: [http : //www.deti.ufc.b/ guilherme](http://www.deti.ufc.br/guilherme)

Funções de Base Radial (RBFs)

Funções de Base Radial - Outras

- **Função Multiquadrática:**

$$\phi(u_i) = \sqrt{\left(\frac{1}{2\sigma^2}\right)^2 + \|\mathbf{x} - \mathbf{x}_i\|^2} \quad (6)$$

- **Função Multiquadrática Inversa:**

$$\phi(u_i) = \frac{1}{\sqrt{\left(\frac{1}{2\sigma^2}\right)^2 + \|\mathbf{x} - \mathbf{x}_i\|^2}} \quad (7)$$

Funções de Base Radial (RBFs)

- Adicionando as condições de interpolação apresentadas na Eq. (2) na Eq. (3) pode-se obter o seguinte conjunto de equações lineares simultâneas para os pesos desconhecidos w_i , a saber:

$$\begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2N} \\ & \vdots & \vdots & \vdots \\ \phi_{N1} & \phi_{N2} & \dots & \phi_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix}$$

em que $\phi_{ij}(\|\mathbf{x} - \mathbf{x}_i\|)$ e $(i, j) = 1, 2, \dots, N$.

Funções de Base Radial (RBFs)

- Considere a matriz, bem como os dois vetores

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2N} \\ & \vdots & \vdots & \vdots \\ \phi_{N1} & \phi_{N2} & \dots & \phi_{NN} \end{bmatrix}, \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} \text{ e } \mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix}$$

que são, respectivamente, a matriz de interpolação, o vetor de pesos e o vetor de respostas desejadas.

- O sistema linear apresentado anteriormente pode ser rescrito como segue.

$$\Phi \mathbf{w} = \mathbf{d}. \quad (8)$$

E, assim, a matriz de pesos \mathbf{w} pode ser obtida por $\mathbf{w} = \Phi^{-1} \mathbf{d}$.

Funções de Base Radial (RBFs)

- No entanto, resolver o sistema linear da Eq. (8) em muitos casos é inviável em termos computacionais.
- Além disto, requer a inversão de uma matriz $N \times N$ que cresce de modo polinomial com N .

Rede RBF Generalizada)

- A rede RBF pode ser generalizada para realizar uma busca em um espaço de menor dimensionalidade que se aproxime da solução apresentada na Eq. (3).
- Para tal pode-se considerar uma base finita para obtenção desta solução aproximada $F(\mathbf{x})$, tal que

$$F^*(\mathbf{x}) = \sum_{i=1}^q w_i \phi_i(||\mathbf{x} - C_i||). \quad (9)$$

em que $q < N$ e C_i é o i -ésimo centro em que $i = 1, \dots, q$.

Funções de Base Radial (RBFs)

- Os centros podem ser escolhidos do conjunto de padrões de treinamento.
- O novo sistema é apresentado como segue.

$$\hat{\Phi} = \begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1q} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2q} \\ & \vdots & \vdots & \vdots \\ \phi_{n1} & \phi_{n2} & \dots & \phi_{nq} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_q \end{bmatrix} \quad e \quad \mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

- Neste contexto, o sistema linear pode ser resolvido pela pseudo-inversa:

$$\mathbf{w} = (\hat{\Phi}^T \hat{\Phi})^{-1} \hat{\Phi}^T \mathbf{d} \quad (10)$$

Arquitetura da rede RBF

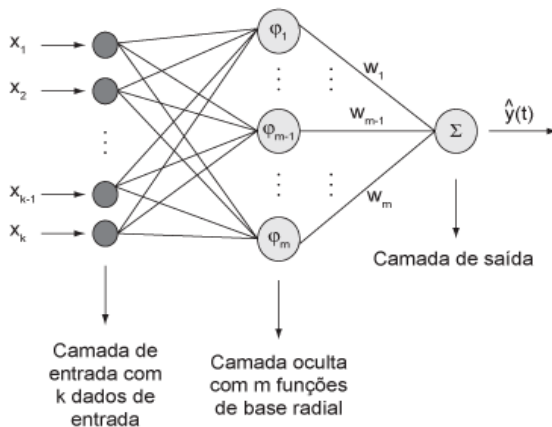


Figure: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2008000300016

Considerações

- Os pesos da camada de saída podem também ser obtidos utilizando a **regra delta**.
- A abertura da gaussiana σ é um parâmetro da rede RBF.
- Os neurônios ocultos podem ser determinados por seleção aleatório de padrões de treinamento para representar o centro do neurônio.

OBRIGADO