Introdução às Redes Neurais

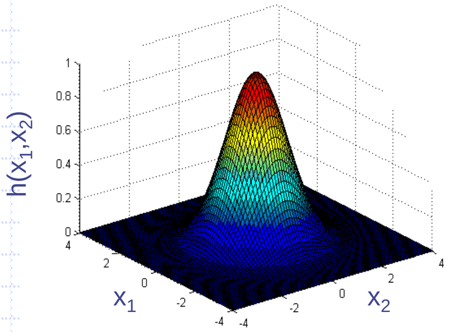
Redes RBF:
Parte II

Prof. Ricardo J. G. B. Campello

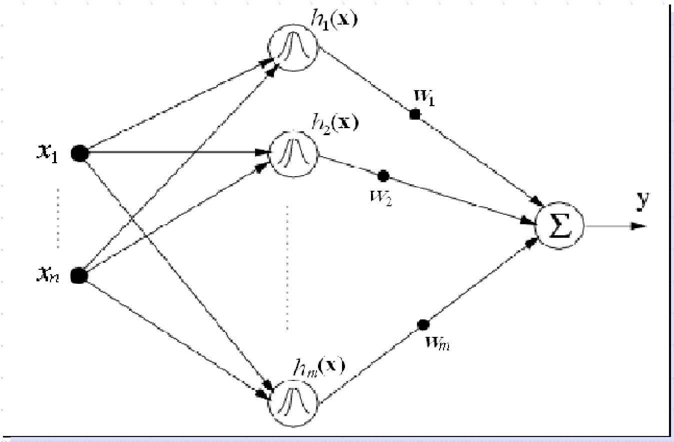
Modelo RBF (revisão)

- Função Radial para n Entradas (Gaussiana):

$$h(x_1, \ldots, x_n) = \exp\left(-\frac{(x_1 - c_1)^2}{\sigma_1^{\,2}}\right) \times \cdots \times \exp\left(-\frac{(x_n - c_n)^2}{\sigma_n^{\,2}}\right)$$



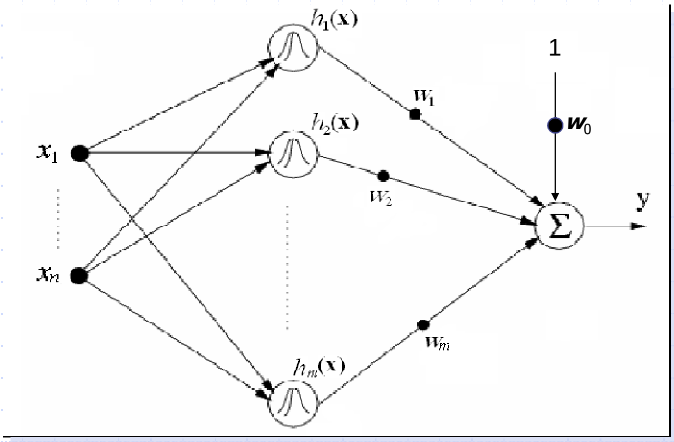
Modelo RBF (revisão)



$$y = \sum_{i=1}^{m} w_i \cdot h_i(x_1, \cdots, x_n) \quad = \quad \mathbf{h}^{\mathrm{T}} \cdot \mathbf{w}$$

3

Modelo RBF (revisão)



$$y \;=\; w_0 + \sum_{i=1}^{m} w_i \cdot h_i(x_1, \cdots, x_n) \;=\; \begin{bmatrix} 1 & \mathbf{h}^{\mathrm{T}} \end{bmatrix} \cdot \begin{bmatrix} w_0 \\ \mathbf{w} \end{bmatrix}$$

4

Treinamento de Redes RBF

- Problemas:
  - Dadas as funções de base radial, como determinar o melhor vetor de pesos **w** ?
  - Como determinar bons parâmetros para as funções de base radial (centros e aberturas) ?
    - reduzir no. de funções para dada precisão, ou
    - aumentar precisão para dado no. de funções

5

Treinamento de Redes RBF

- Problemas:
  - **Dadas as funções de base radial, como determinar o melhor vetor de pesos w ?**
  - Como determinar bons parâmetros para as funções de base radial (centros e aberturas) ?
    - reduzir no. de funções para dada precisão, ou
    - aumentar precisão para dado no. de funções

6

Determinação dos Pesos

- Queremos determinar os pesos de uma rede RBF com funções radiais conhecidas para aproximar a função (mapeamento) entre um conjunto de N padrões de entrada $\mathbf{x}_k = [x_{1k} \ldots x_{nk}]^T$ e saída $y(\mathbf{x}_k)$ para k=1,...,N
  - Por simplicidade, assume-se aqui que a saída é única
- Exemplo:
  - para N = 100 clientes de um banco, descritos por n = 5 variáveis (salário, idade, tempo de relacionamento com o banco, sexo, estado civil), queremos obter uma rede que, dado um cliente $\mathbf{x}_k$, responda se este cliente possui maior risco de ser caloteiro ($y(\mathbf{x}_k) = 1$) ou não ($y(\mathbf{x}_k) = 0$)

7

Determinação dos Pesos

- Como as m funções radiais são conhecidas, podemos calcular o valor dessas funções para cada padrão $\mathbf{x}_k$
- Representando esses valores em uma matriz tem-se:

$$\mathbf{H} = \begin{bmatrix} h_1(\mathbf{x}_1) & \cdots & h_m(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ h_1(\mathbf{x}_N) & \cdots & h_m(\mathbf{x}_N) \end{bmatrix}$$

- As saídas desejadas (conhecidas) também podem ser representadas em um vetor:

$$\mathbf{y} = [y(\mathbf{x}_1) \quad \cdots \quad y(\mathbf{x}_N)]^T$$

8

Determinação dos Pesos

- Para uma representação ideal da saída, deseja-se que a rede RBF possua pesos **w** tais que:

$$\begin{bmatrix} h_1(\mathbf{x}_1) & \cdots & h_m(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ h_1(\mathbf{x}_N) & \cdots & h_m(\mathbf{x}_N) \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} y(\mathbf{x}_1) \\ \vdots \\ y(\mathbf{x}_N) \end{bmatrix}$$

ou seja: $\mathbf{H} \cdot \mathbf{w} = \mathbf{y}$

- Se **H** fosse quadrada, poderíamos multiplicar por $\mathbf{H}^{-1}$ em ambos os lados e resolver o problema como:

$$\mathbf{w} = \mathbf{H}^{-1} \cdot \mathbf{y}$$

9

Determinação dos Pesos

- Podemos obter uma matriz quadrada multiplicando ambos os lados da equação por $\mathbf{H}^T$:

$$\mathbf{H}^T \cdot \mathbf{H} \cdot \mathbf{w} = \mathbf{H}^T \cdot \mathbf{y}$$

- Agora sim podemos multiplicar ambos os lados da equação pela inversa $(\mathbf{H}^T\mathbf{H})^{-1}$ e obter a solução como:

$$\boxed{\mathbf{w} = \left(\mathbf{H}^T \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^T \cdot \mathbf{y}}$$

10

Determinação dos Pesos

- Se houver o termo aditivo w_0 na saída, é preciso somente redefinir **H** e **w** de tal forma que:

$$\underbrace{\begin{bmatrix} 1 & h_1(\mathbf{x}_1) & \cdots & h_m(\mathbf{x}_1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & h_1(\mathbf{x}_N) & \cdots & h_m(\mathbf{x}_N) \end{bmatrix}}_{\mathbf{H}} \cdot \underbrace{\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_m \end{bmatrix}}_{\mathbf{w}} = \underbrace{\begin{bmatrix} y(\mathbf{x}_1) \\ \vdots \\ y(\mathbf{x}_N) \end{bmatrix}}_{\mathbf{y}}$$

o que não afeta a solução:

$$\boxed{\mathbf{w} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{y}}$$

Determinação dos Pesos

- Algoritmo:
  - Para um conjunto de k = 1, ..., N padrões de entrada e saída, $\{\mathbf{x}_k , y(\mathbf{x}_k)\}$, calcule a matriz **H**, construa o vetor **y** e calcule os pesos da rede segundo a equação:

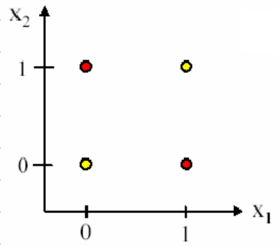
$$\boxed{\mathbf{w} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{y}}$$

- Nota: É possível demonstrar que, se a representação não for exata, a solução acima é aquela que minimiza o erro quadrático entre a saída da rede e a saída desejada (**Mínimos Quadrados**) !

Exercício

- Sabemos que o problema XOR consiste em classificar N = 4 padrões (descritos por n = 2 variáveis) em 2 classes diferentes:

k	x_1	x_2	y (classe)
1	0	0	0
2	0	1	1
3	1	0	1
4	1	1	0



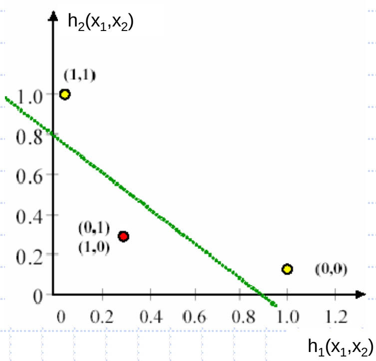
Note que esse problema não é linearmente separável, ou seja, não existe uma função linear que separe as classes no espaço $x_1 \times x_2$

13

Exercício (cont.)

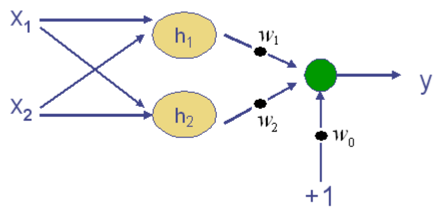
- No entanto, os padrões podem passar a ser linearmente separáveis após serem processados por funções de base radial, desde que essas sejam escolhidas de forma apropriada. P. ex., escolhendo 2 funções Gaussianas h_1 e h_2 com centros dados por [0,0] e [1,1], respectivamente, e desvios padrão unitários para ambas, tem-se:

x_1	x_2	$h_1(x_1,x_2)$	$h_2(x_1,x_2)$
0	0	1	0,1353
0	1	0,3679	0,3679
1	0	0,3679	0,3679
1	1	0,1353	1



Exercício (cont.)

- Considere então a rede RBF com essas funções de base radial e pesos w_1 e w_2, além do peso adicional w_0 (aditivo na saída), conforme abaixo:



- Obtenha os valores dos pesos w_0, w_1 e w_2 para que a rede consiga indicar precisamente em sua saída a classe correta (0 ou 1) de cada um dos 4 padrões de entrada do problema XOR.
  - Apresente a solução passo a passo, em detalhes e de forma justificada !

15

Bibliografia

- Braga, et al., "Redes Neurais Artificiais: Teoria e Aplicações", LTC, 2ª Edição, 2007
- Haykin, "Neural Networks", Prentice Hall, 2nd Edition, 1999
- Kovács, "Redes Neurais Artificiais: Fundamentos e Aplicações", Collegium Cognitio, 2ª Edição, 1996

16