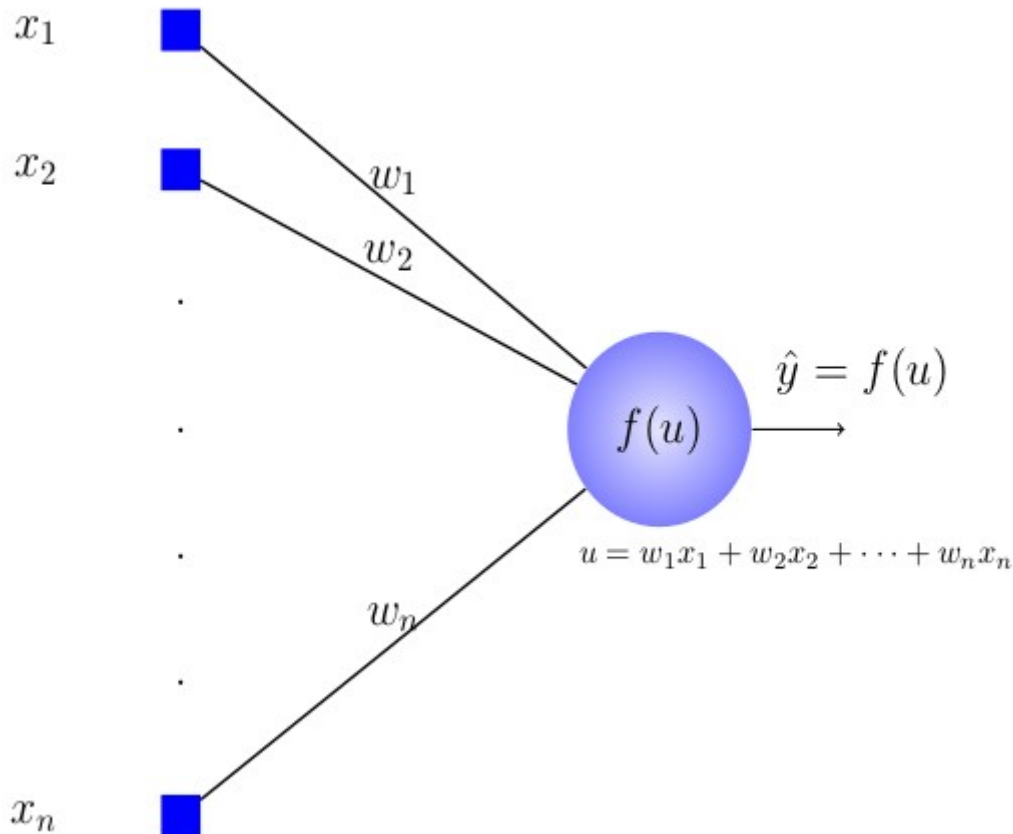


# RNA

## Extreme Learning Machines

# Introdução



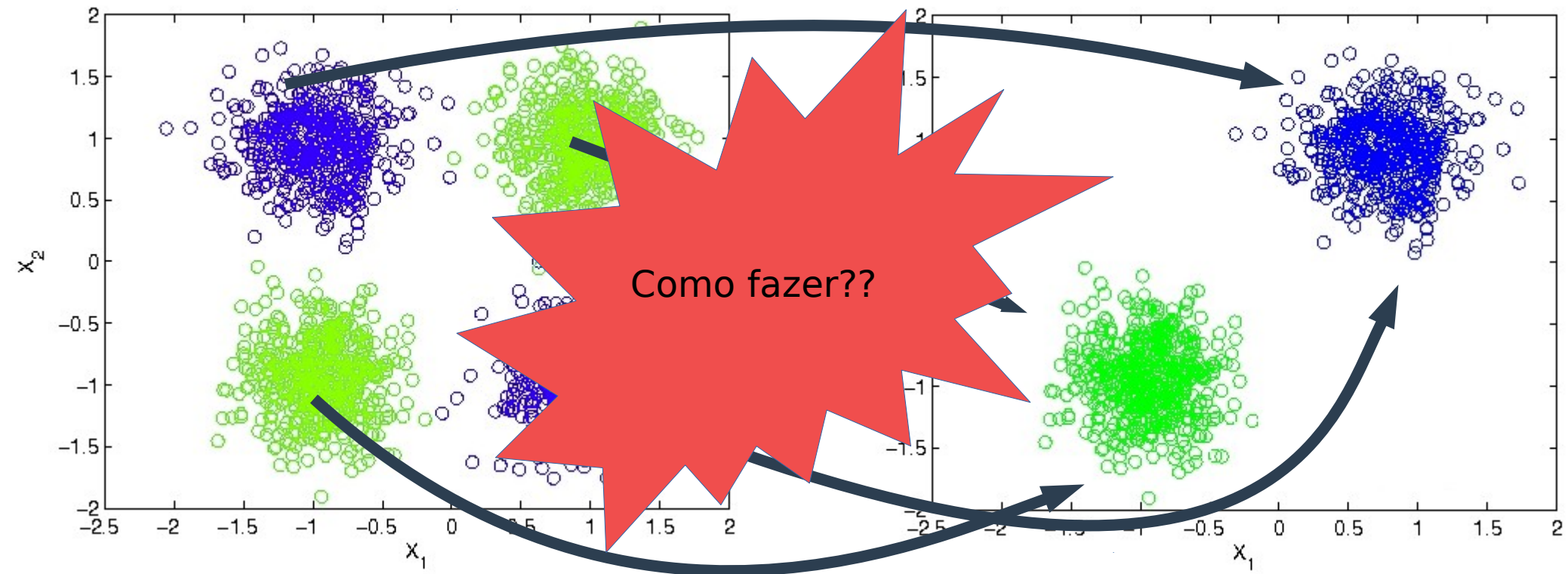
Até o momento  
trabalhamos com separadores  
e aproximadores lineares.

Adaline -  $f(u) = u$

Perceptron -  $f(u) \begin{cases} 1 & \text{se } u \geq 0 \\ 0 & \text{se } u < 0 \end{cases}$

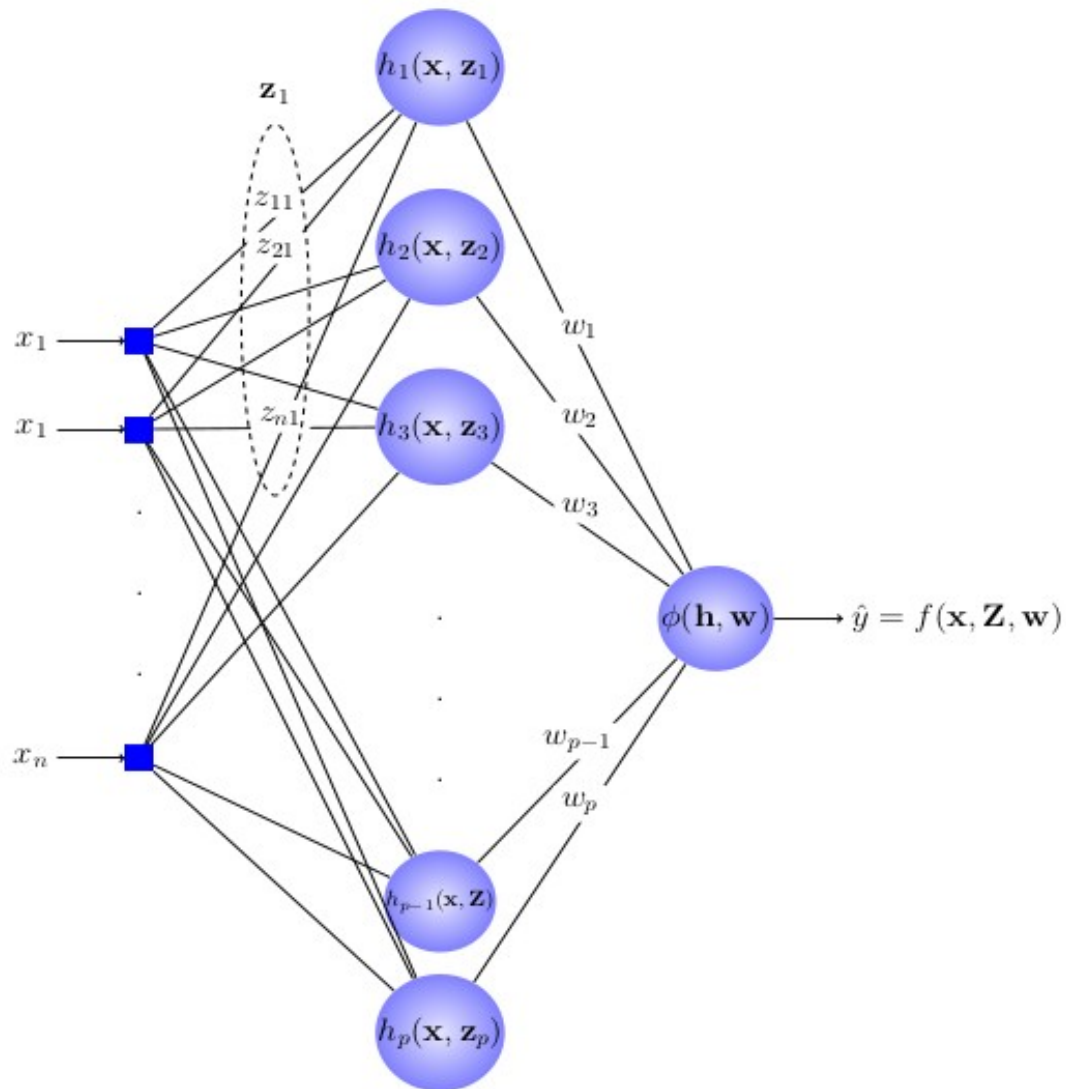
# Introdução

E quando o problema é não-linearmente separável?



Fazemos um mapeamento não-linear

# Estrutura de uma RNA



Rede feed-forward

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T$$

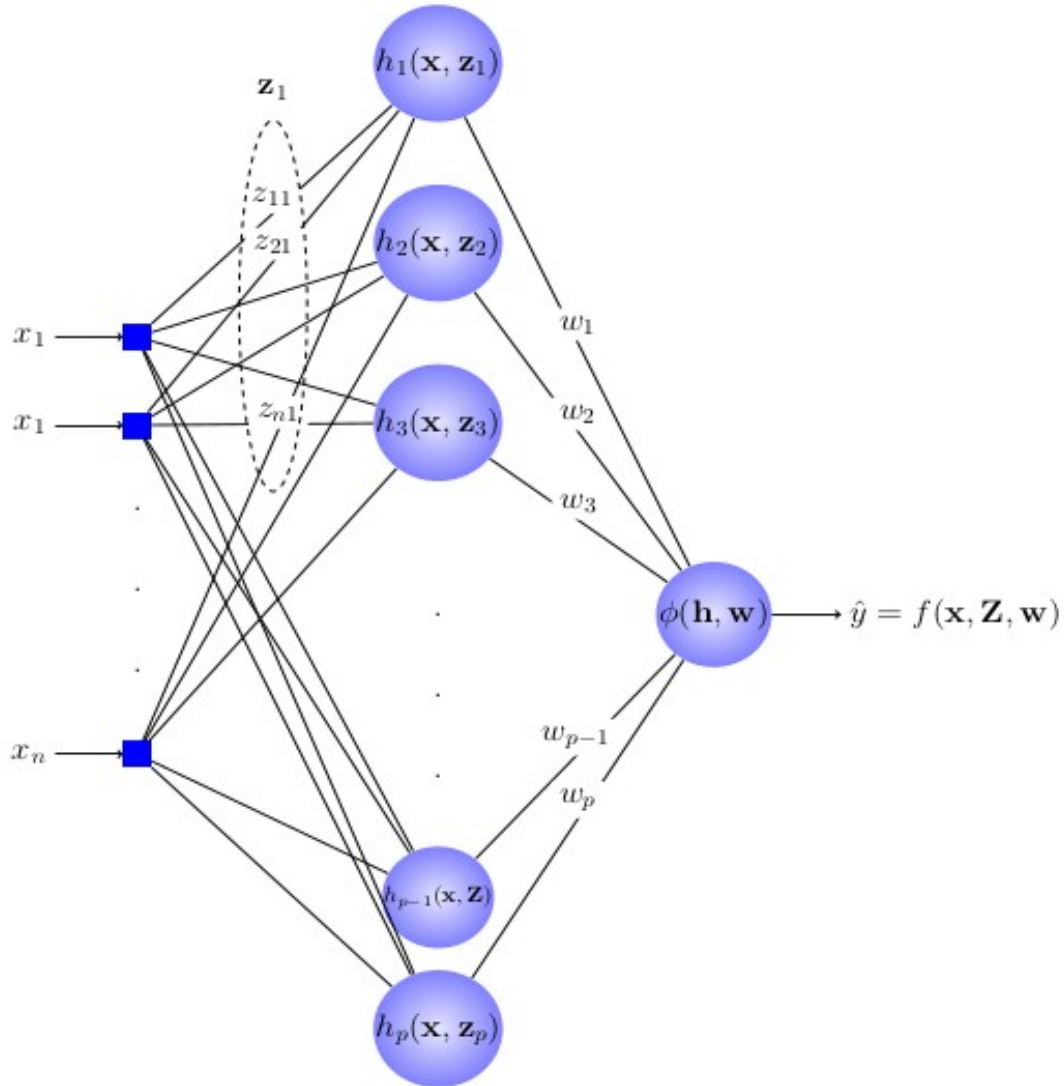
$$h_i(u_i)$$

$$u_i = \sum_i z_i x_i$$

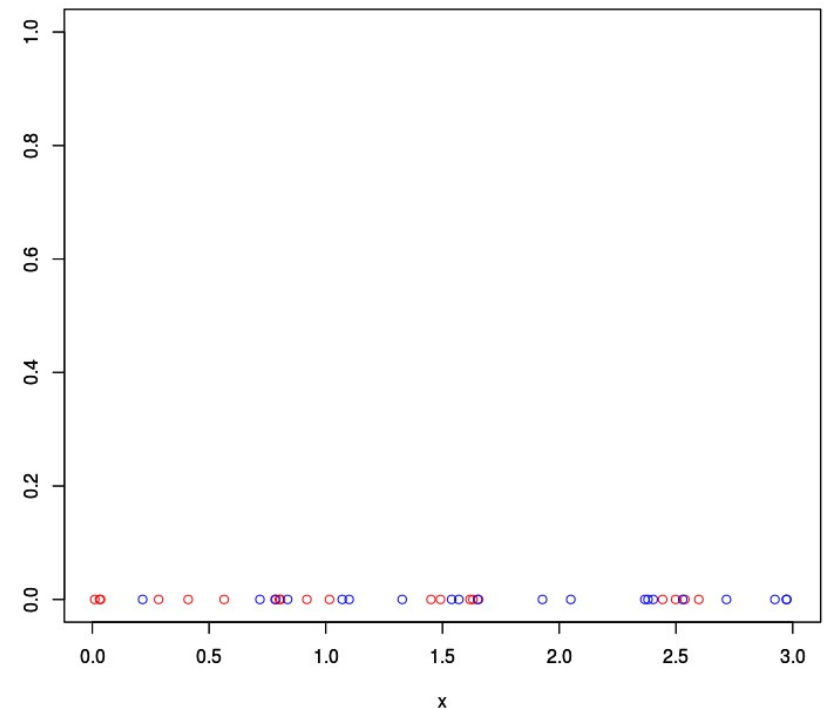
$$\Phi_o(\mathbf{h}, \mathbf{w})$$

$\mathbf{Z}$  é uma matriz  
 $\mathbf{W}$  é um vetor

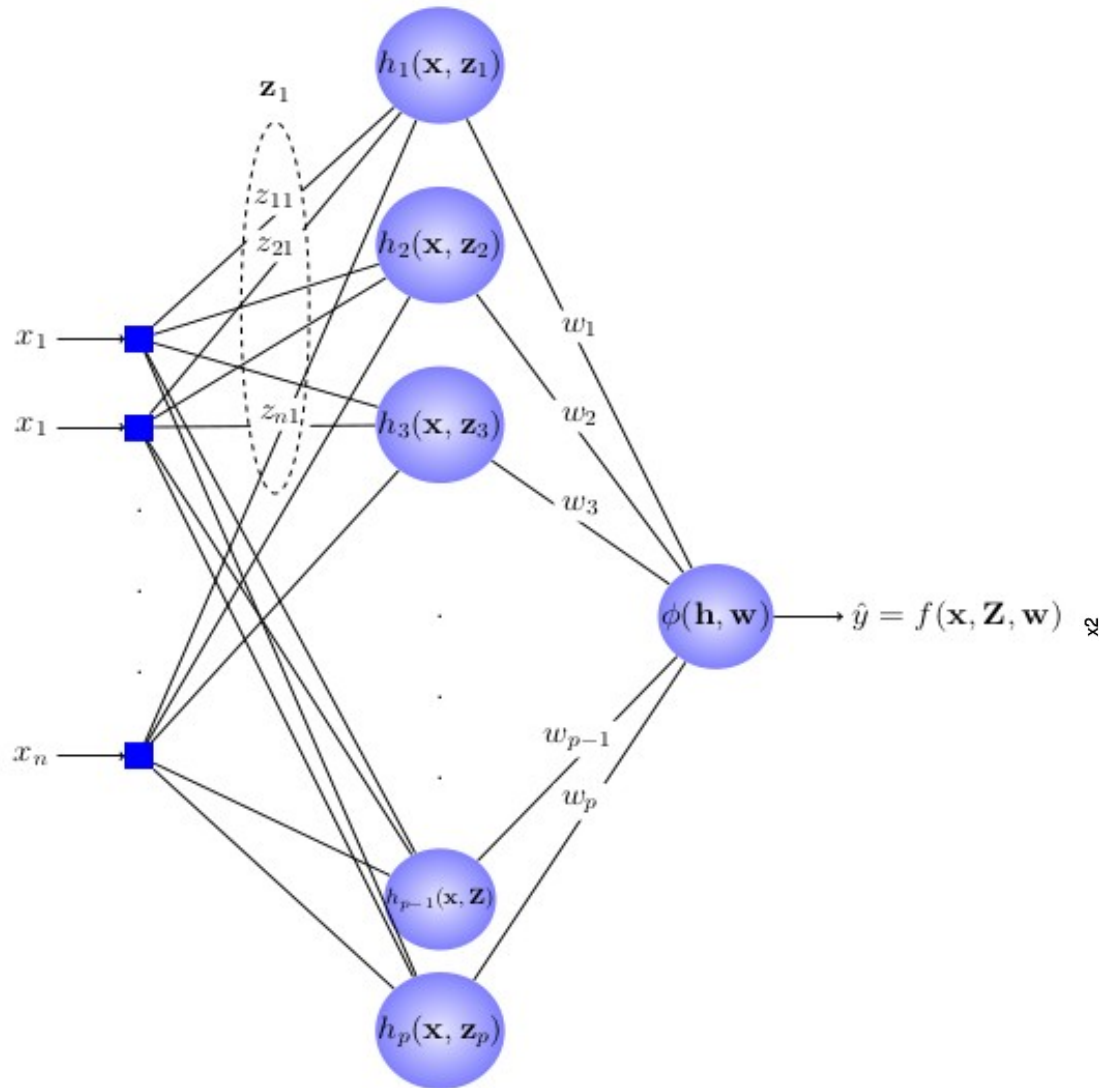
# Função da camada intermediária



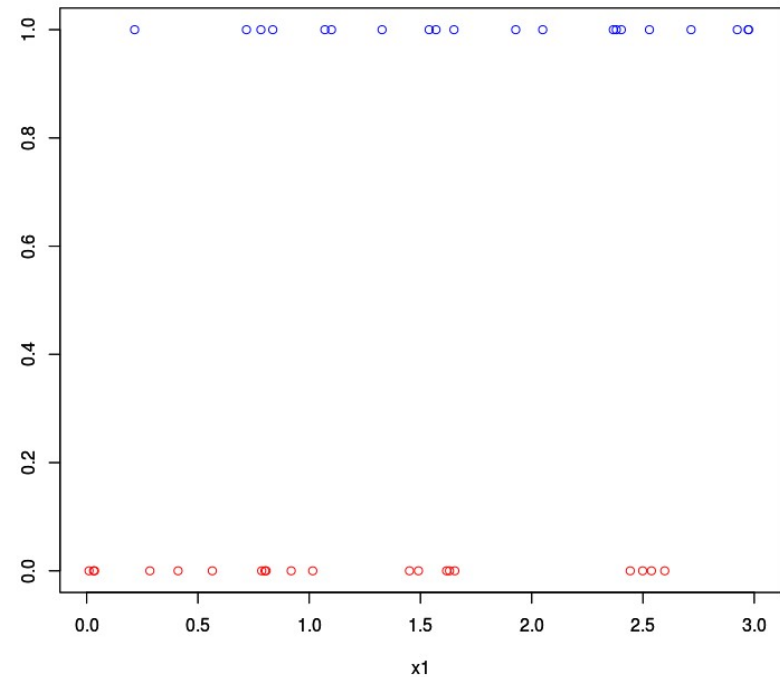
Teorema de Cover



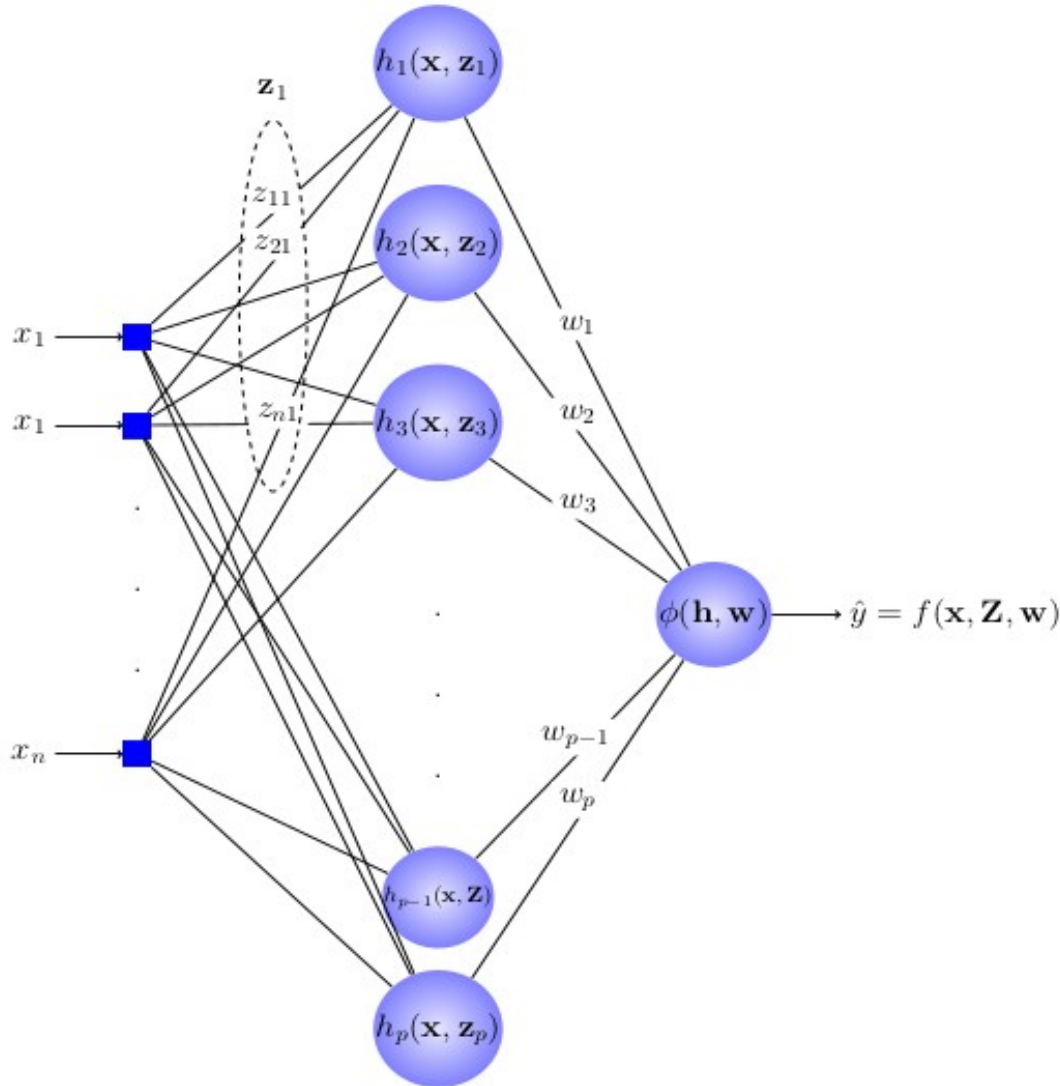
# Função da camada intermediária



Teorema de Cover

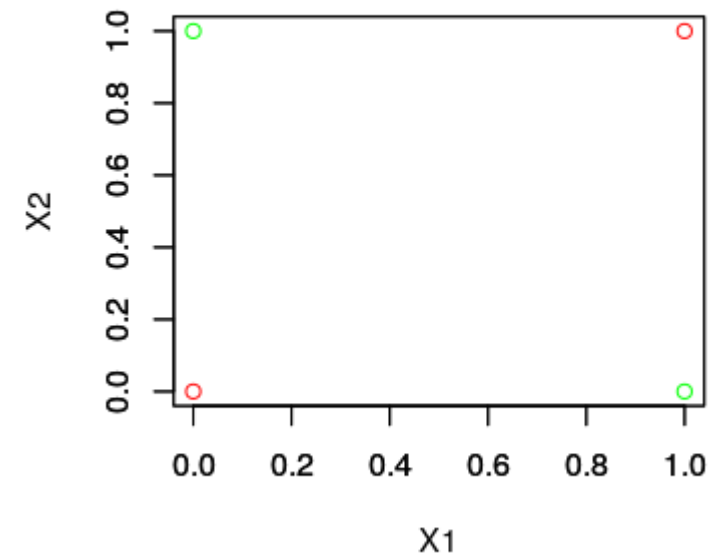
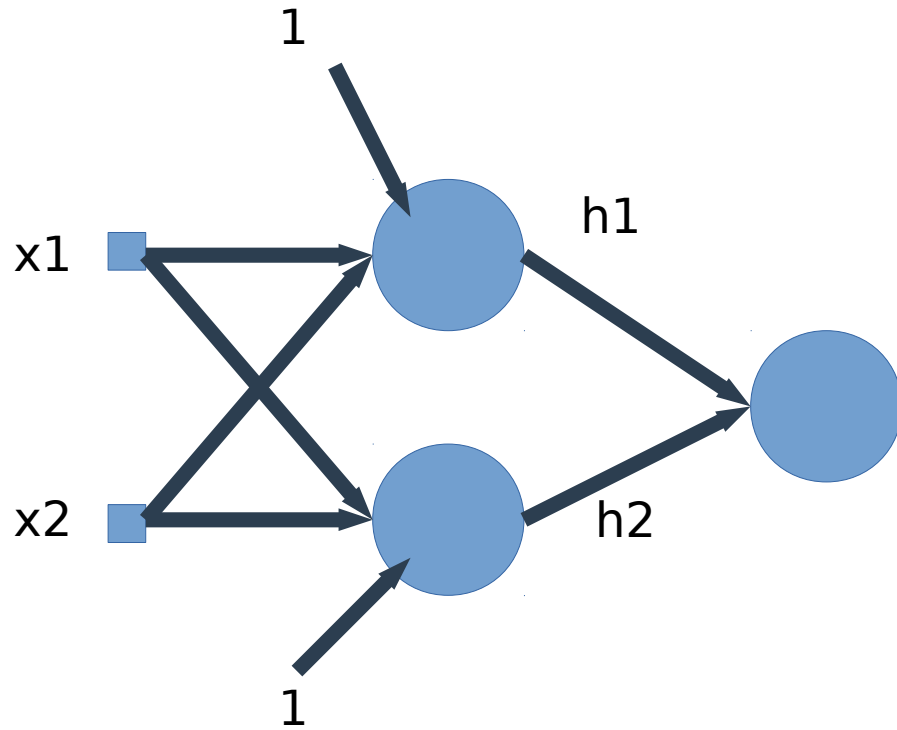


# Função da camada intermediária



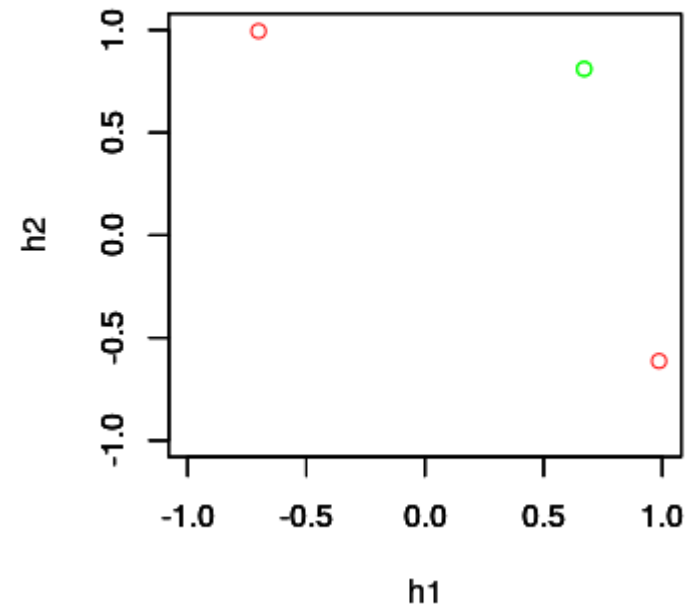
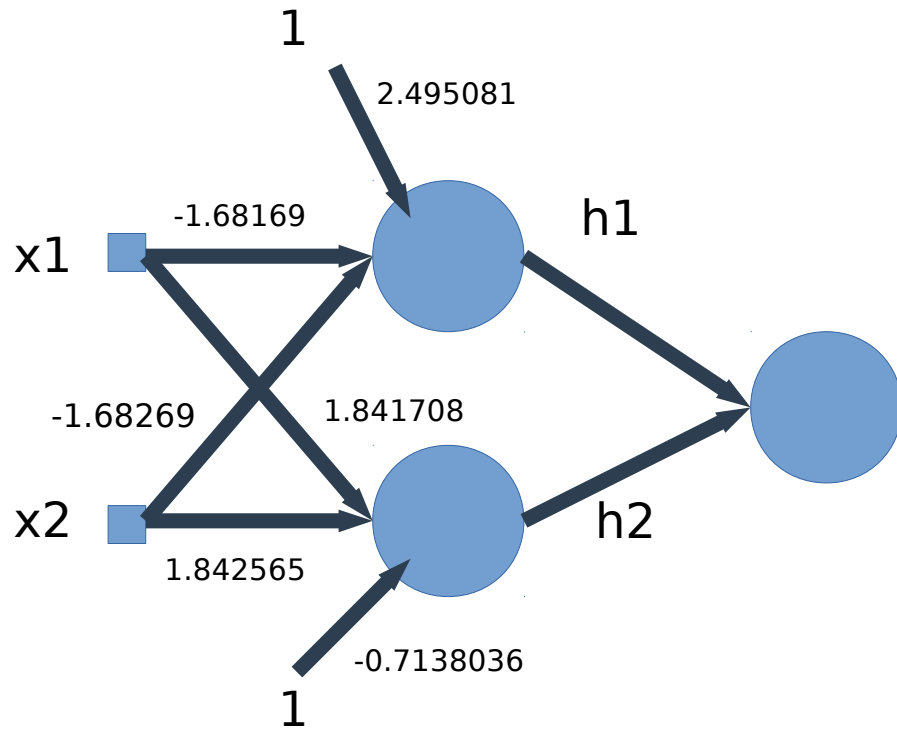
Resolvem o problema mapeando não-linearmente em um espaço onde o neurônio de saída consegue fazer a separação linear.

# Função da camada intermediária





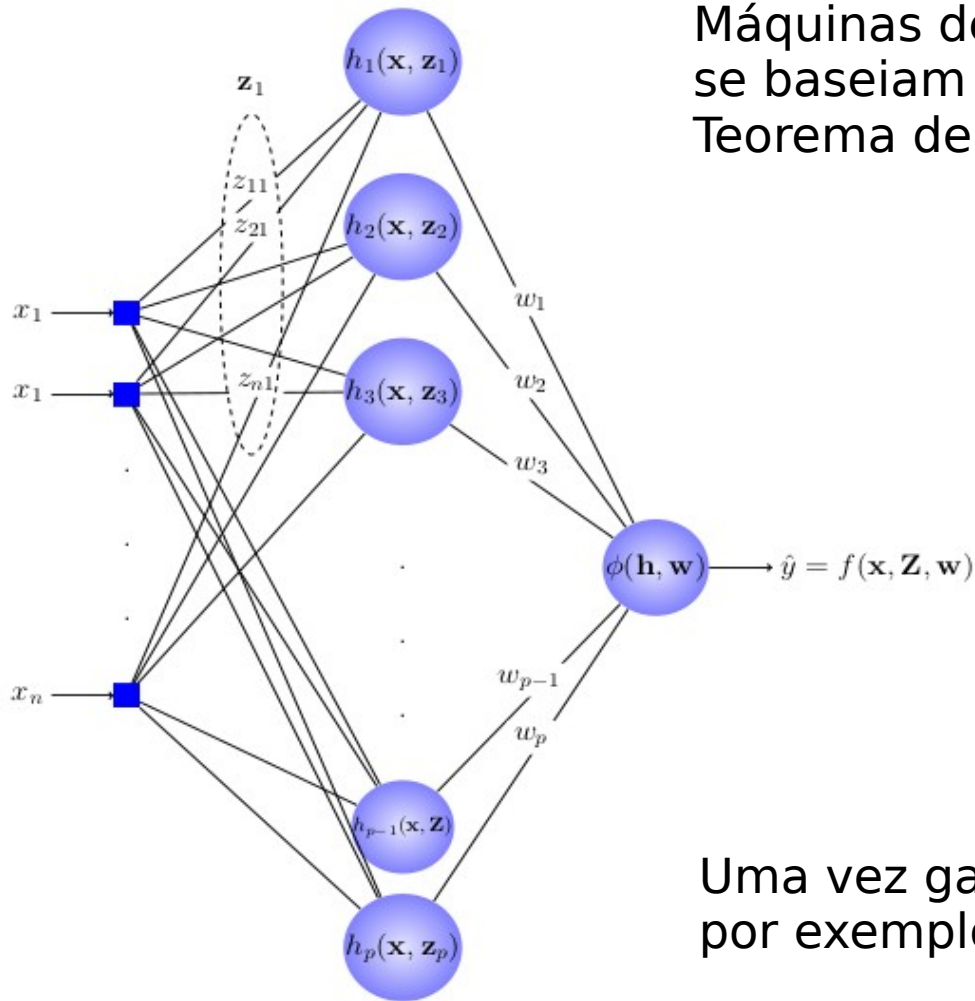
# Função da camada intermediária



$$h_1 = 2.495081 * 1 + (-1.681699) * x_1 + (-1.68269) * x_2$$
$$h_2 = -0.7138036 * 1 + 1.841708 * x_1 + 1.842565 * x_2$$

# ELM

As ELMs (Extreme Learning Machines) ou Máquinas de Aprendizado Extremo, se baseiam em uma interpretação particular do Teorema de Cover.



Como no Teorema não há nenhuma restrição sobre a forma do mapeamento  $h(\mathbf{x}_i, \mathbf{z}_i)$ , o princípio das ELMs é que a matriz  $\mathbf{Z}$  seja selecionada de forma aleatória e que o número de funções  $h_i(\mathbf{x}, \mathbf{z}_i)$  seja suficientemente grande para garantir a separabilidade no espaço da camada intermediária.

Uma vez garantida a separação linear  $\rightarrow$  pseudoinversa por exemplo

Alguns princípios básicos devem ser satisfeitos para que a solução do problema através de ELMs seja plausível:

- A escolha aleatória da matriz de pesos  $Z$  deve garantir a separação linear no espaço da camada intermediária.
- Aumento da dimensão do espaço da camada intermediária para garantir a separabilidade não afeta diretamente o sobre-ajuste do modelo aos dados.
- Uma vez que a separabilidade esteja garantida pela expansão da camada intermediária, uma solução direta de erro mínimo e margem máxima pode ser obtida.