Universidade Federal do Espírito Santo
Centro de Ciências Agrárias – CCENS UFES
Departamento de Computação

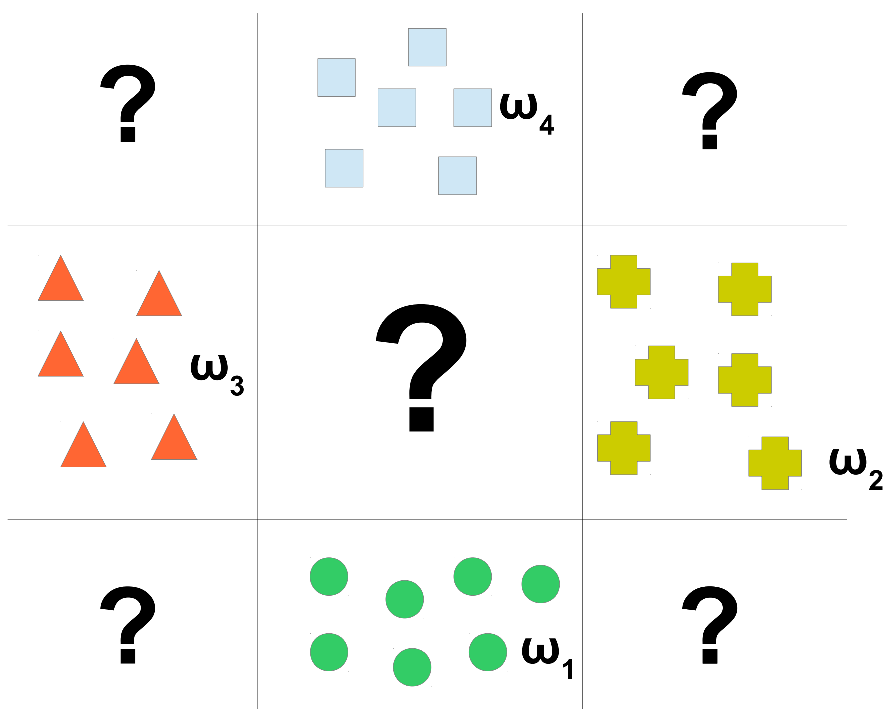
Máquina Linear

Redes Neurais Artificiais

Site: http://jeiks.net
E-mail: jacsonrcsilva@gmail.com

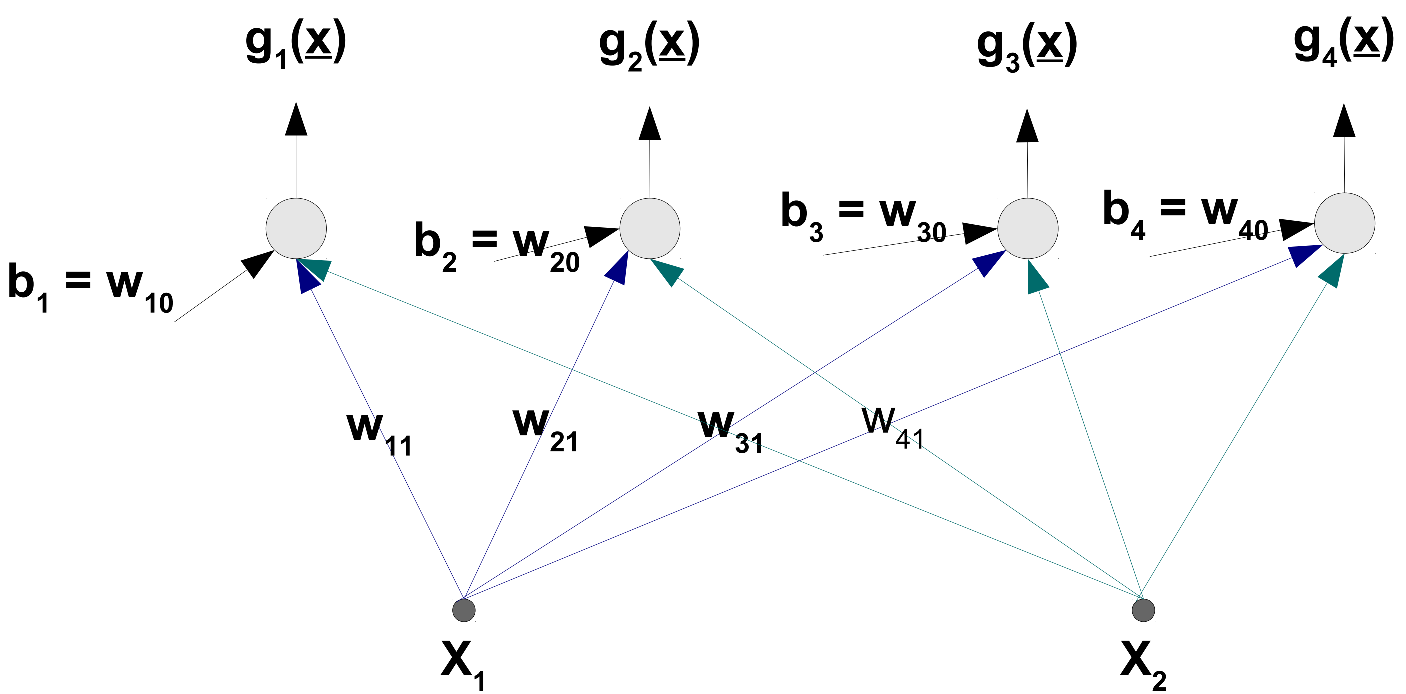
Máquina Linear

- Caso multi-classes: c classes $\omega_1, \ldots, \omega_c$.
  - Para cada classe ω_i,
    deve-se agrupar todos os padrões $\{\phi^{(k)}\}$, $1 \leq k \leq n$
    com $\phi^{(k)} \notin \omega_i$ em uma classe "não-ω_i" e
    gerar c-1 planos que separam ω_i de "não-ω_i", i=1, …, c.
  - Problema:
    existirão regiões ambíguas, sem classes definidas ("?")



Máquina Linear

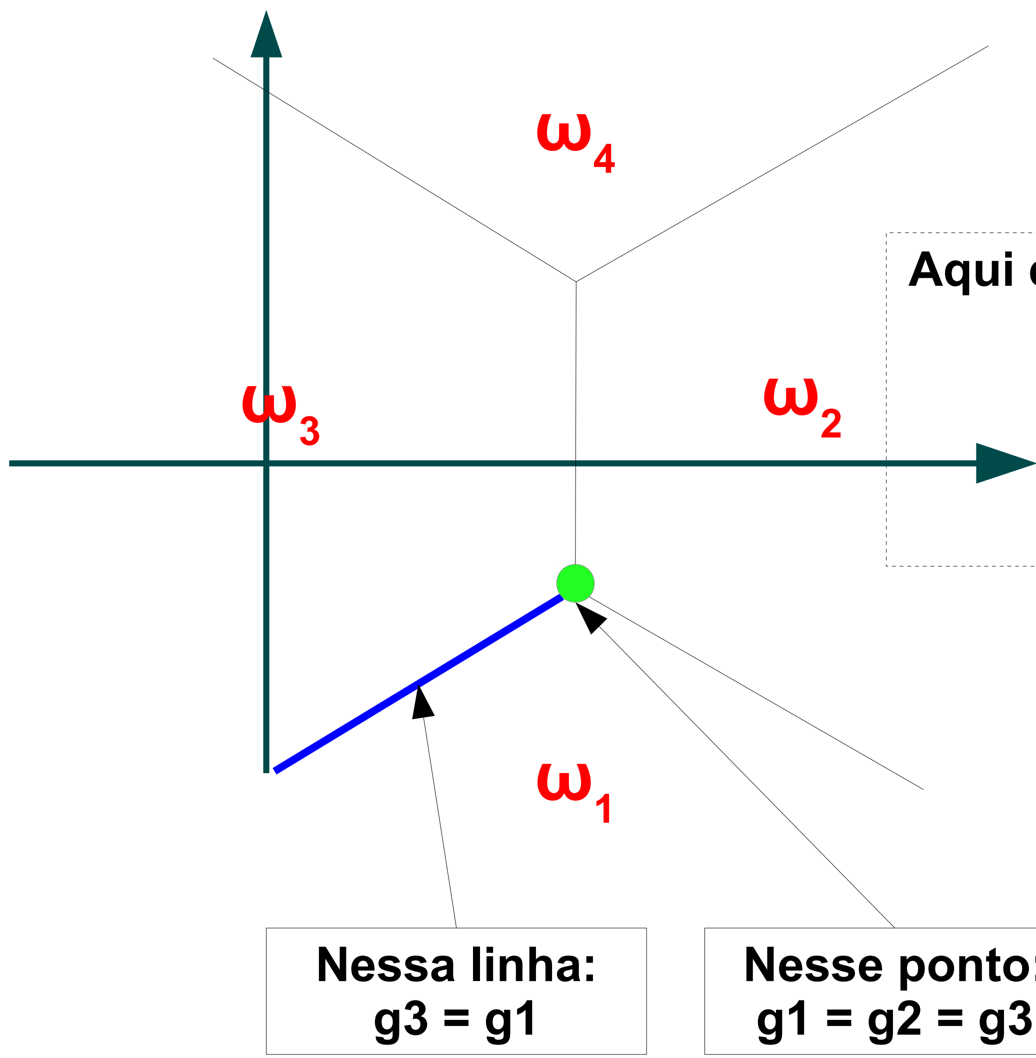
- A solução para isso é construir uma Máquina Linear.
- Suas características:
  - Não possui função de ativação;
  - Possui saída contínua, não sendo a saída de sinal.
  - Não tem ambiguidades, pois a maior saída corresponde a classe selecionada.
  - Ou seja, o neurônio com maior resultado representa a classe.



Máquina Linear

g é a função discriminativa

$g_i(\underline{x}; \underline{w}_i, b_i) = \underline{w}_i \cdot \underline{x} + b_i, \; i=1, \dots, c \in \mathbb{R}$



Aqui é a região de decisão de ω_1, onde:

$g_2 > g_1$

$g_2 > g_3$

$g_2 > g_4$

Máquina Linear

- Se $g_i > g_j$, $i \in \{1, \ldots, c\}$, $j = \{1, \ldots, c\} \backslash i \Rightarrow \underline{x} \in \omega_i$.
- Cada classe ω_i tem vetor de peso próprio:

  $w_i = (w_{i0}, w_{i1}, \ldots, w_{id})^T$.
- Pode ser feito o agrupamento dos pesos $\underline{w}_i$ como colunas de uma matriz:

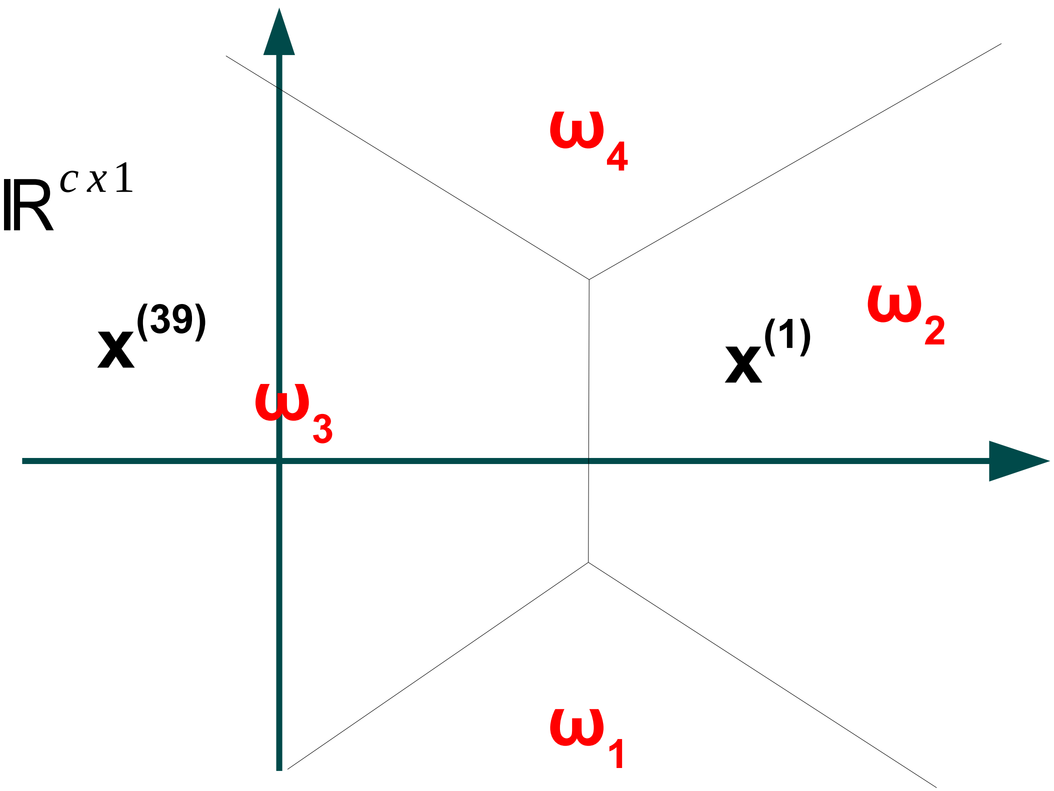
$$W_{(d+1)xc} = \begin{pmatrix} \underline{w}_1 & \underline{w}_2 & \underline{w}_3 & \ldots & \underline{w}_c \end{pmatrix} = \begin{pmatrix} w_{10} & & w_{c0} \\ w_{11} & \cdots & w_{c1} \\ \vdots & & \vdots \\ w_{1d} & & w_{cd} \end{pmatrix}$$

Cálculo da Máquina Linear

$$g_i(\underline{x};\underline{w}_i)=\underline{w}_i\cdot\underline{x}=\underline{w}_i^T\,\underline{x}\in\mathbb{R}$$

$$\underline{g}=\begin{pmatrix} g_1 \\ \vdots \\ g_c \end{pmatrix} \qquad \underline{g}(\underline{x},W)=W^T\underline{x}\in\mathbb{R}^{c\,x\,1}$$

$$D=\{(\underline{x}^{(k)},\underline{y}^{(k)})\}_{k=1}^{x}$$



$$\underline{x}^{(1)}=\begin{pmatrix} 1 \\ 4 \\ 0{,}5 \end{pmatrix} \quad \underline{y}^{(1)}=\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \underline{x}^{(39)}=\begin{pmatrix} 1 \\ -1 \\ 0{,}9 \end{pmatrix} \quad \underline{y}^{(39)}=\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Cálculo da Máquina Linear

- Então, o resultado desejado é:

$$w^T\underline{x}^{(1)} = y^{(1)} \qquad \dots \qquad w^T\underline{x}^{(\dots)} = y^{(\dots)}$$

- Como obter os valores de W ?
  - Para obter W, deve-se isolá-lo. Então:

$$X \,.\, W = Y$$

...

Cálculo da Máquina Linear

- Isolando W:

$$X.W=Y \mid X^{-1}.$$
$$\Rightarrow X^{-1}X.W=X^{-1}.Y \quad \Leftrightarrow \quad W=X^{-1}.Y$$

- Porém, X não é uma matriz quadrada e por isso não possui matriz inversa.
- Deve-se então utilizar sua pseudoinversa:

$$X.W=Y \quad \mid X^{T}.$$
$$X^{T}X.W=X^{T}.Y \quad \mid (X^{T}X)^{-1}.$$
$$(X^{T}X)^{-1}.(X^{T}X).W=(X^{T}X)^{-1}X^{T}.Y$$
$$W=X^{+}.Y$$
$$X^{+}=(X^{T}X)^{-1}X^{T}$$

Exercício

- Com as entradas:

  X1 = (-1.0 1.0);

  X2 = (1.0 3.0);

- E as saídas:

  Y1 = -1;

  Y2 = 1;

- Calcule os coeficientes de regressão W através da pseudoinversa da matriz de entradas.
- Antes disso, adicione a coluna do bias nas entradas.