

# **REDES NEURONAIS ARTIFICIAIS**

Luís Morgado

ISEL-DEETC

# ALGORITMO DE RETROPROPAGAÇÃO

## Erro de classificação

$$E = \sum_z (d_{sz} - o_{sz})^2$$

## Actualização dos pesos

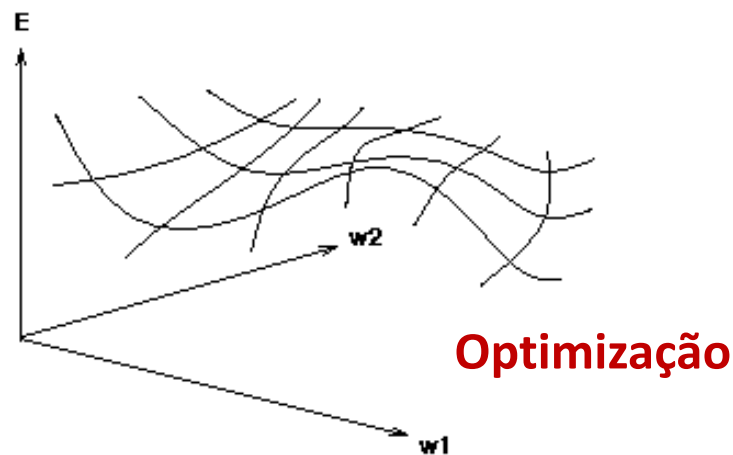
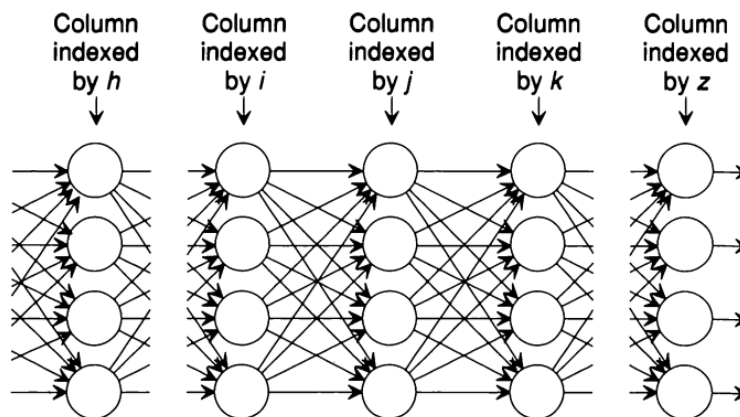
$$w_{ij}^{(n+1)} = w_{ij}^{(n)} + \Delta w_{ij}^{(n)}$$

$$\Delta w_{ij}^{(n)} = -\eta \left( \frac{\partial E}{\partial w_{ij}} \right)$$

$\eta$  - Taxa de aprendizagem

## Condição de paragem

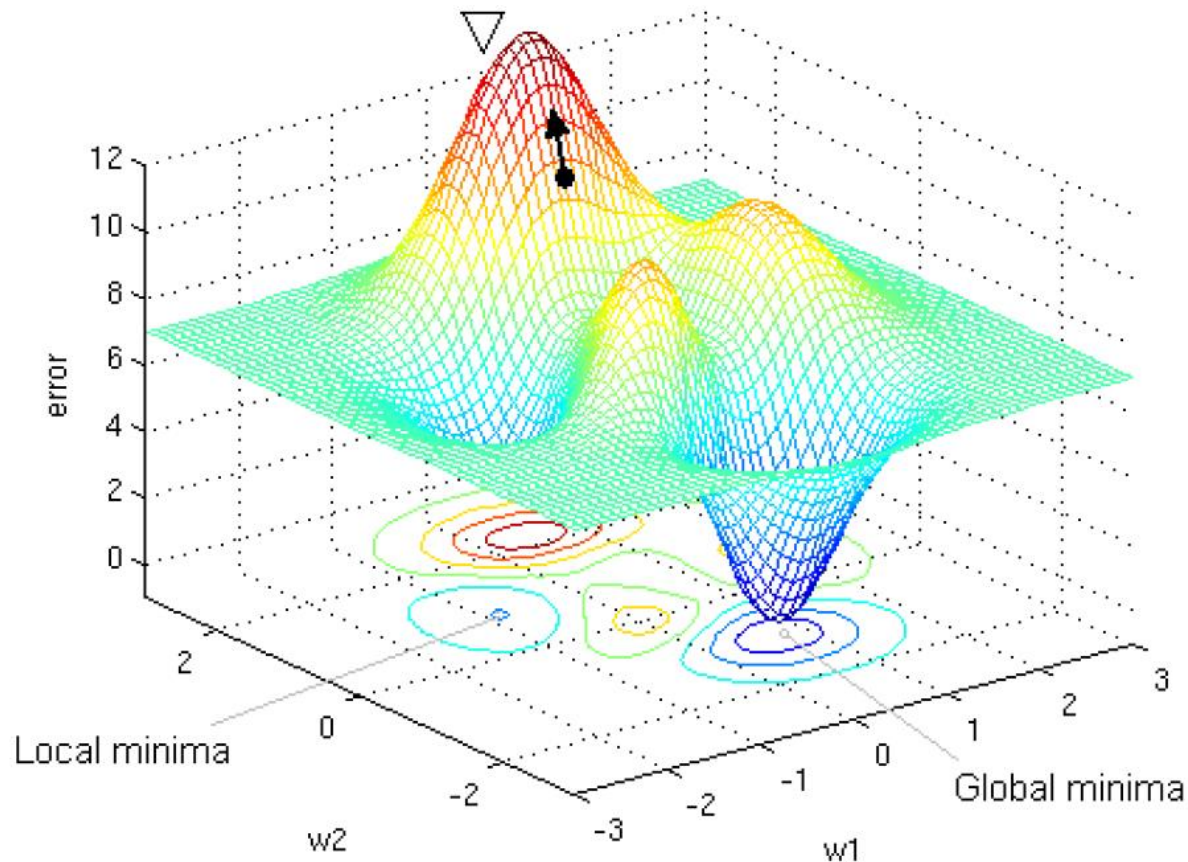
$$E = \sum_z (d_{sz} - o_{sz})^2 \leq \varepsilon$$



# ALGORITMO DE RETROPROPAGAÇÃO

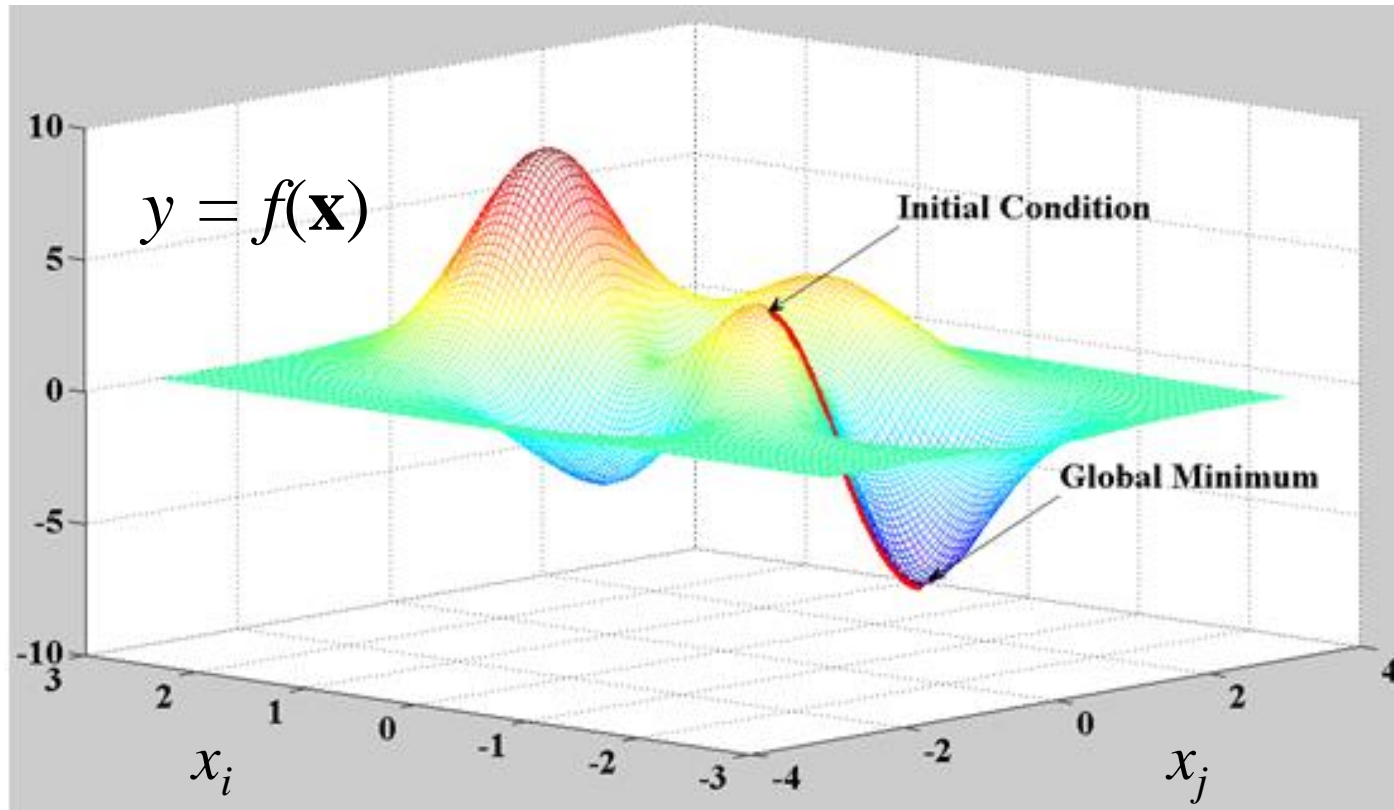
Optimização num espaço de pesos  $w_i$

Minimização do erro de classificação



# ALGORITMO DE DESCIDA DE GRADIENTE

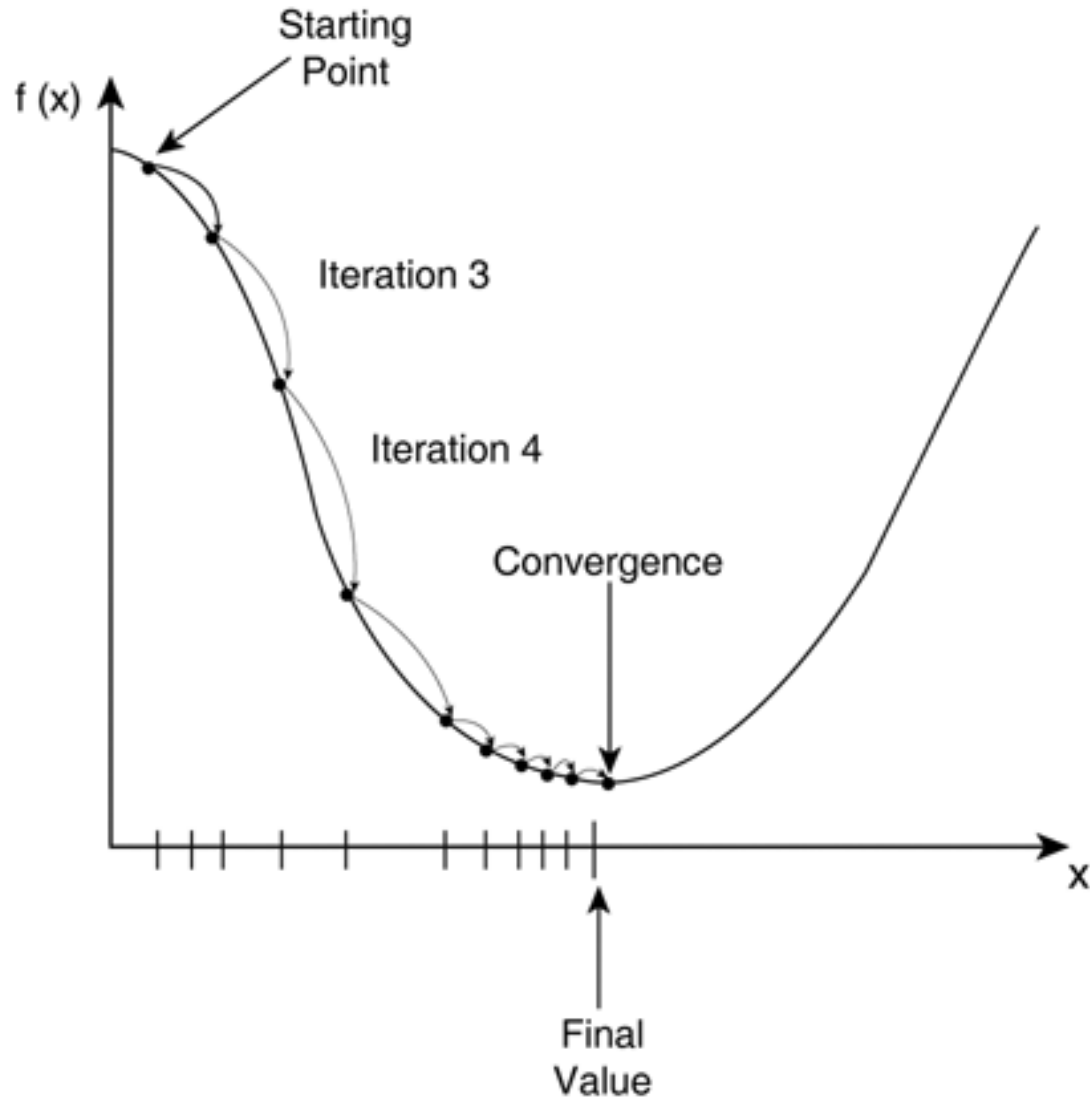
Objectivo: Minimizar erro



$$\mathbf{x}^{n+1} = \mathbf{x}^n + \Delta \mathbf{x}^n, n \geq 0$$

$$\Delta \mathbf{x}^n = -\eta \nabla f(\mathbf{x}) \quad \eta - \text{Dimensão do passo}$$

# ALGORITMO DE DESCIDA DE GRADIENTE



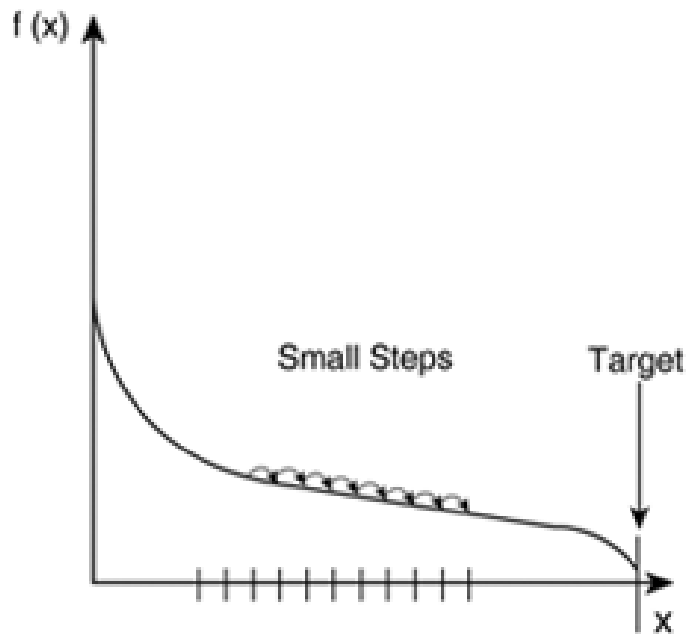
# ALGORITMO DE DESCIDA DE GRADIENTE

## Problemas

$$\mathbf{x}^{n+1} = \mathbf{x}^n + \Delta \mathbf{x}^n, n \geq 0$$

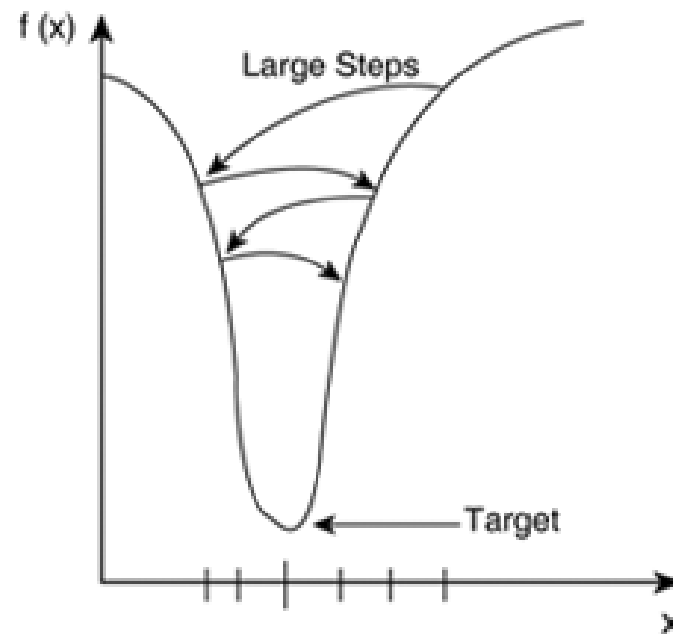
$$\Delta \mathbf{x}^n = -\eta \nabla f(\mathbf{x}) \quad \eta - \text{Dimensão do passo}$$

Passo  $\eta$  pequeno



Convergência lenta

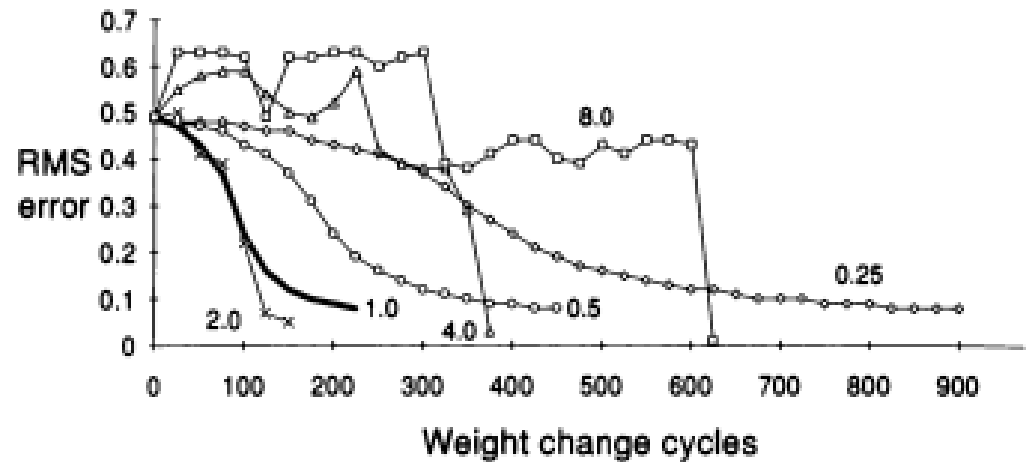
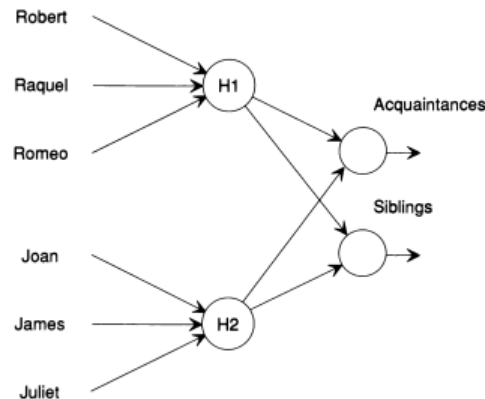
Passo  $\eta$  grande



Oscilação

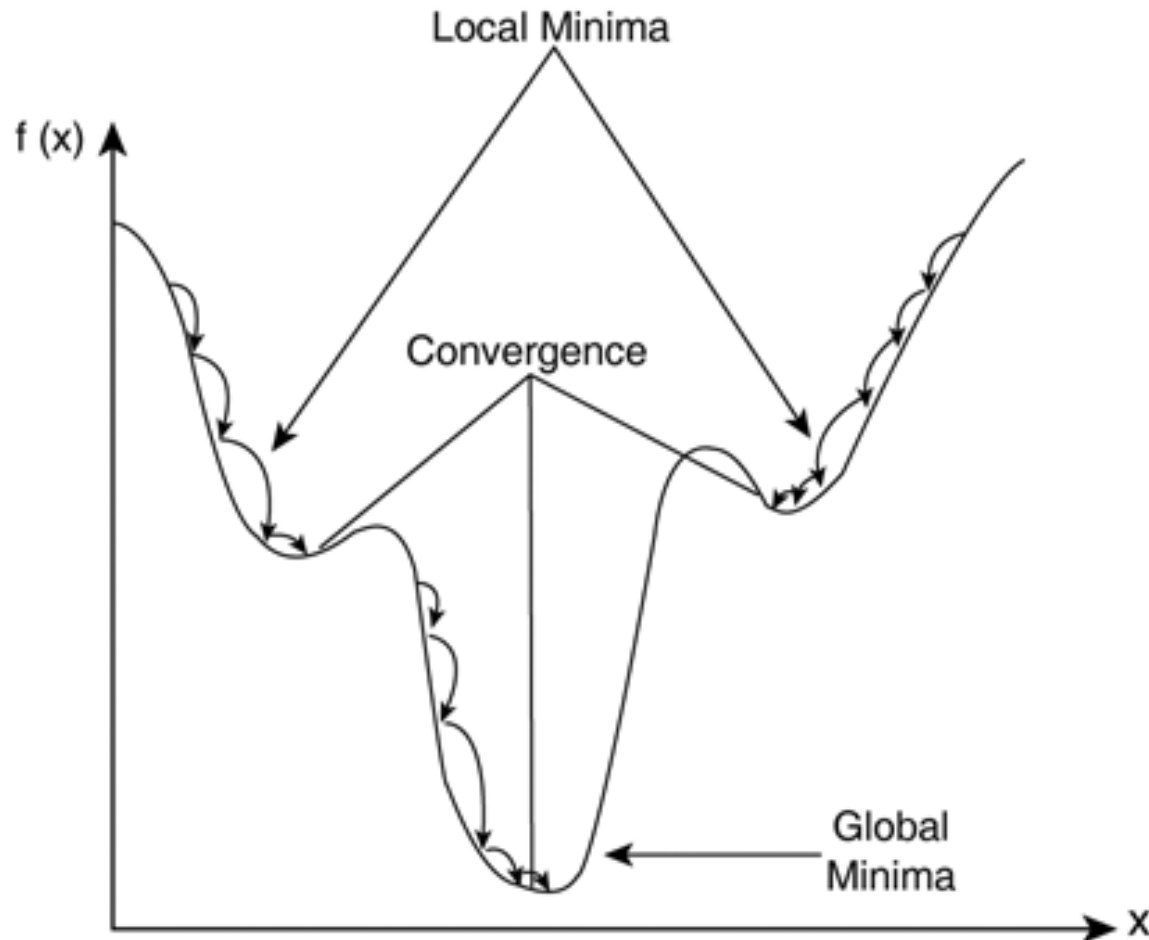
# ALGORITMO DE RETROPROPAGAÇÃO

## Efeito da variação da taxa de aprendizagem $\eta$



# ALGORITMO DE DESCIDA DE GRADIENTE

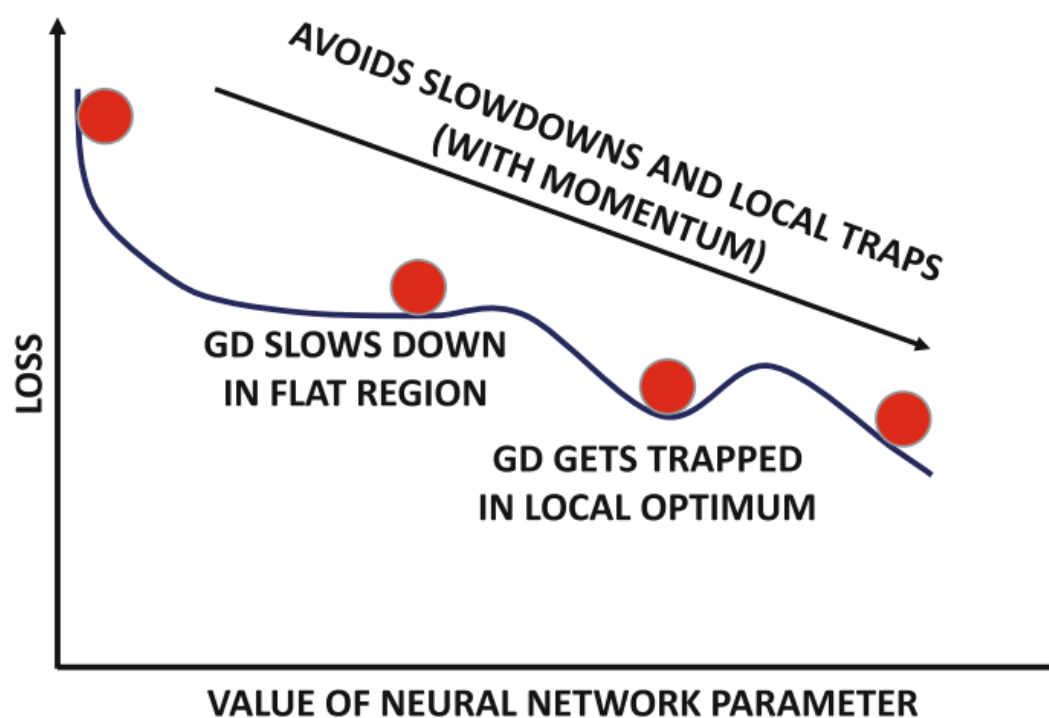
## Óptimos locais





# ALGORITMO DE DESCIDA DE GRADIENTE

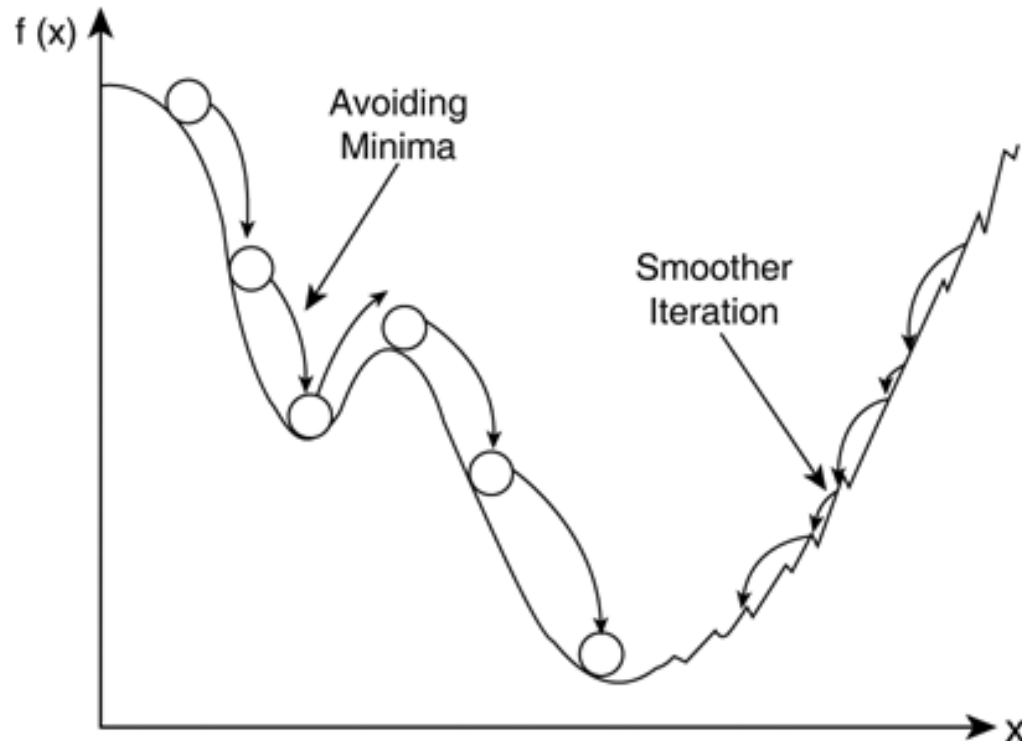
## Descida de gradiente com momento



# ALGORITMO DE DESCIDA DE GRADIENTE

## Descida de gradiente com momento

$$\Delta x_i \leftarrow \alpha \Delta x_{i-1} - \eta \nabla f(x_i)$$



# APRENDIZAGEM EM REDES NEURONAIS

## Métodos de Optimização para Treino de Redes Neuronais

### *Descida de Gradiente*

$$\Delta w_i = -\eta \frac{\partial L}{\partial w_i}$$

### *Descida de Gradiente com Momento*

$$\Delta w_i^t = \alpha \Delta w_i^{t-1} - \eta \frac{\partial L}{\partial w_i}$$

### *Descida de Gradiente Estocástica*

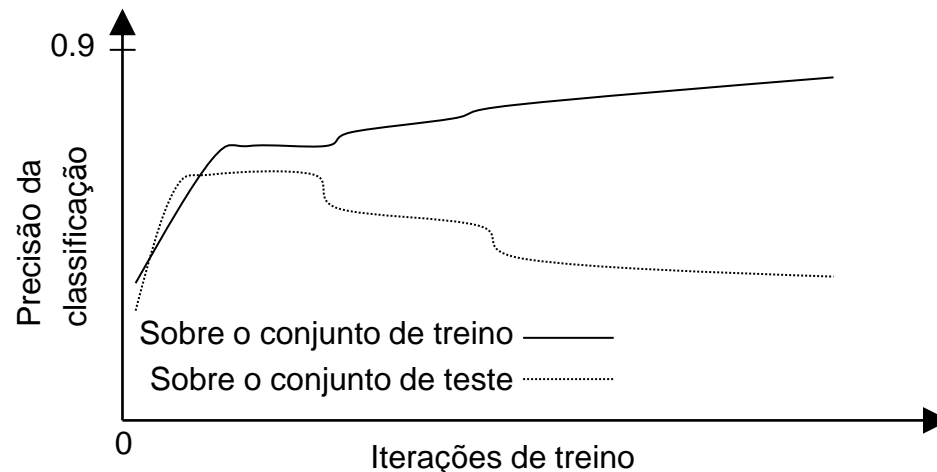
Em cada iteração de treino é utilizado um subconjunto amostra do conjunto de treino (*batch*)

### **Outros métodos**

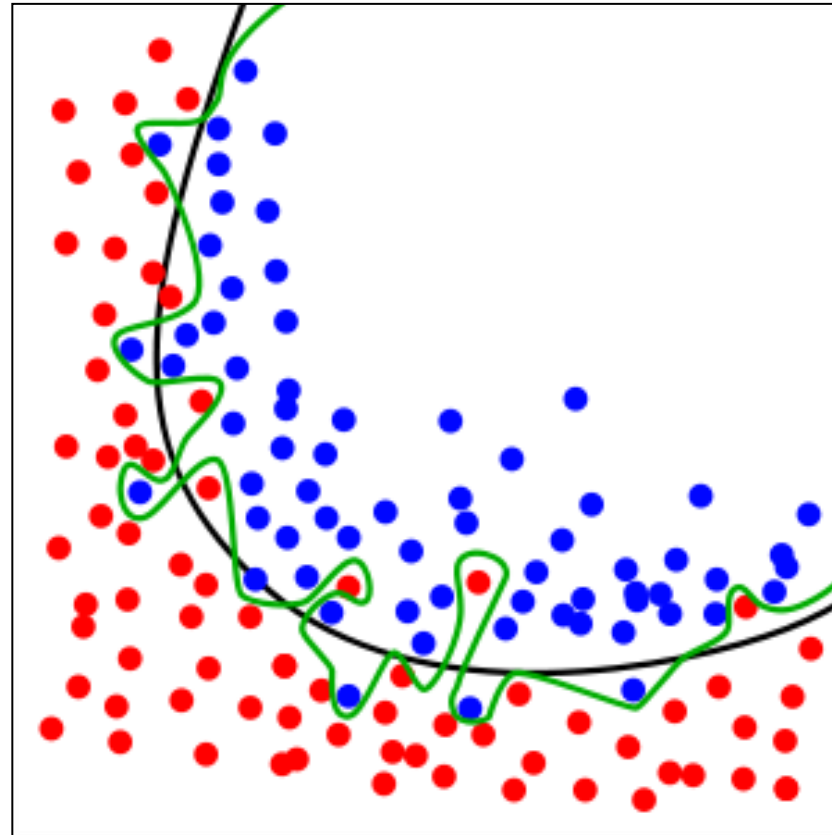
*Root Mean Square Propagation (RMSProp), Adaptive Moment Estimation (Adam), Adaptive Gradient Algorithm (AdaGrad)*

# SOBREPARAMETRIZAÇÃO

- A **sobreparametrização (*overfitting*)** é um dos problemas que pode surgir com frequência em redes neurais artificiais
- Se a aprendizagem se prolongar em demasia, entre outras causas, pode resultar numa **diminuição da capacidade de generalização da rede**
- Apesar da rede ter aprendido os detalhes de um conjunto de treino particular, pode não responder de forma adequada quando são apresentados exemplos novos



# SOBREPARAMETRIZAÇÃO (*Overfitting*)

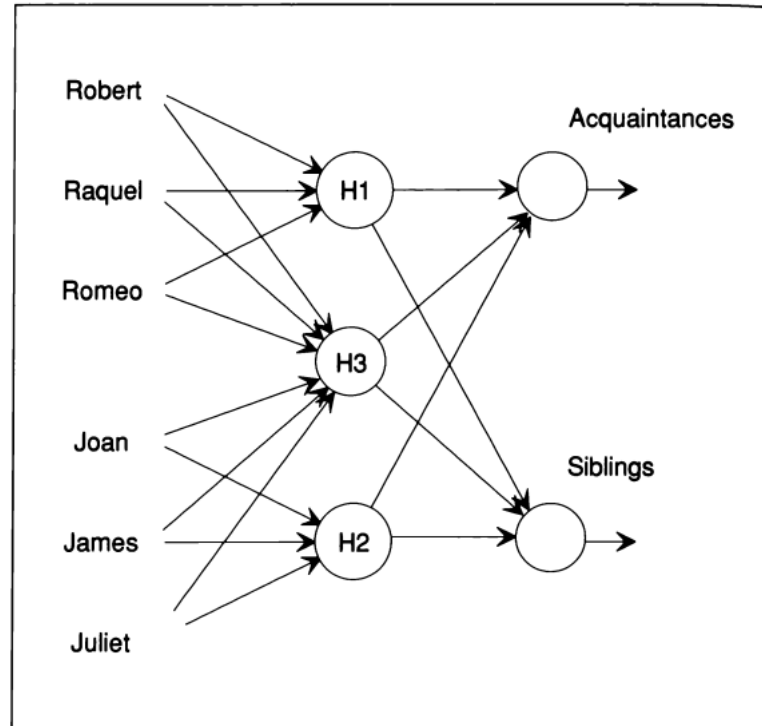


— Rede com  
sobreparametrização

Uma rede pode aprender os detalhes de um conjunto de treino particular, não respondendo de forma adequada quando são apresentados exemplos novos

# SOBREPARAMETRIZAÇÃO (*Overfitting*)

## Exemplo



Robert	Raquel	Romeo	Joan	James	Juliet	$A_d$	$A_o$	$S_d$	$S_o$
1	0	0	0	0	1	1	0.99	0	0.00
0	0	1	1	0	0	1	0.06 (?)	0	0.94 (?)
0	0	0	0	1	1	0	0.97 (?)	1	0.01 (?)

# **SOBREPARAMETRIZAÇÃO (*Overfitting*)**

- **Causas**

- Aprendizagem prolongada em demasia
- Poucos exemplos de treino
- Dados com ruído
- Rede tem demasiados parâmetros (neurónios, ligações)
  - Mais do que necessário para os dados em causa

# SOBREPAREMETRIZAÇÃO (*Overfitting*)

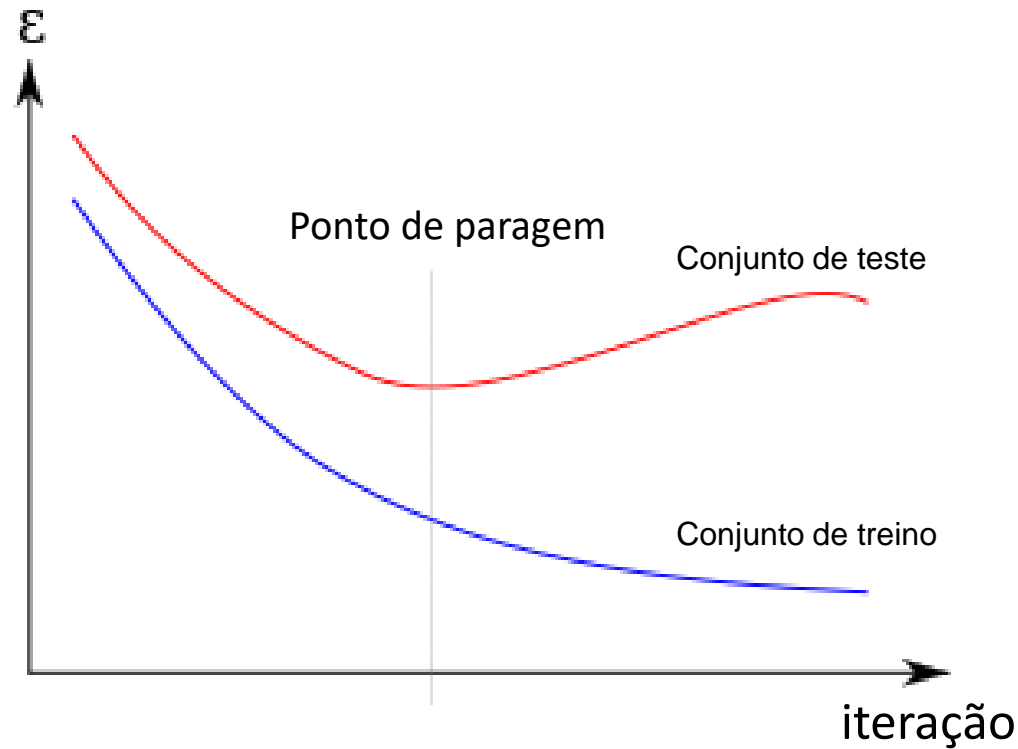
- Soluções

- Simplificação da rede
- Paragem prévia (*early stopping*)
- Aumento dos dados (*data augmentation*)
- Regularização
- Redução selectiva de parâmetros (*dropout*)



# SOBREPARAMETRIZAÇÃO (*Overfitting*)

Paragem prévia (*early stopping*)



# BIBLIOGRAFIA

[Aggarwal, 2018]

C. Aggarwal, *Neural Networks and Deep Learning*  
Springer, 2018

[Munakata, 1998]

T. Munakata, *Fundamentals of the New Artificial Intelligence*, Springer, 1998

[Winston, 1992]

P. Winston, *Artificial Intelligence*, 3rd Edition, Addison-Wesley, 1992

[Raizer *et al.*, 2009]

K. Raizer, H. Idagawa, E. Nobrega, L. Ferreira, *Training and Applying a Feedforward Multilayer Neural Network in GPU*, CILAMCE, 2009