Classificação:
Regressão Logística

Prof.: Eric A. Antonelo

Slides baseados no curso de *Machine Learning*
de Andrew Ng

DAS-UFSC

Classificação

E-mail: Spam / Não Spam?

Transações online: Fraudulentas (Sim/não)?

Tumor: Maligno (sim/não)?

$$y \in \{0, 1\}$$

0 (não)
1 (sim)

Regressão Logística

Queremos que $$0 \leq h_\theta(x) \leq 1$$

$h_\theta(x)$ = probabilidade estimada de que y = 1 dado x

Exemplo:

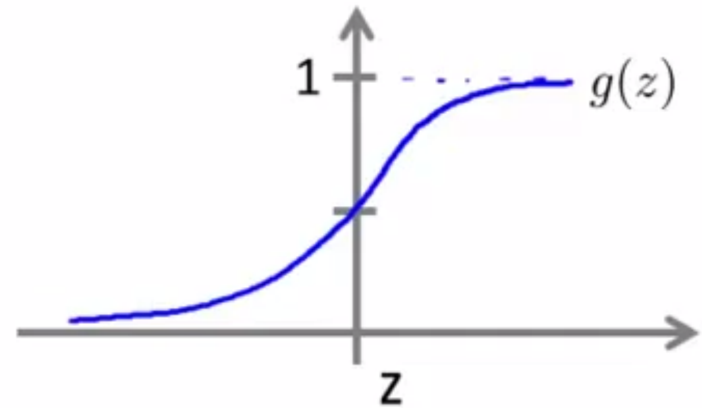
$$x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \text{tumorSize} \end{bmatrix}$$

$$h_\theta(x) = 0.7$$

70% de chance de tumor maligno

Fronteira de Decisão

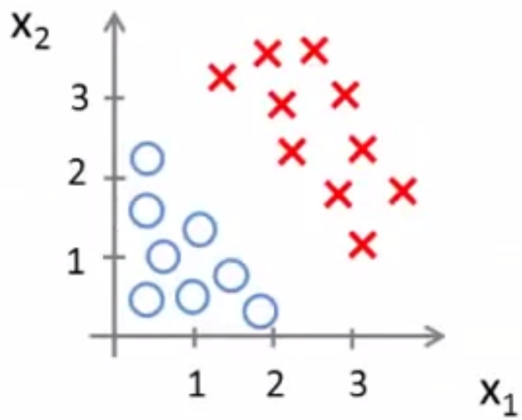
$$h_\theta(x) = g(\theta^T x)$$
$$g(z) = \frac{1}{1+e^{-z}}$$



Suponha que a predição seja “y = 1” se $h_\Phi(x) >= 0.5$

a predição seja “y = 0” se $h_\Phi(x) < 0.5$

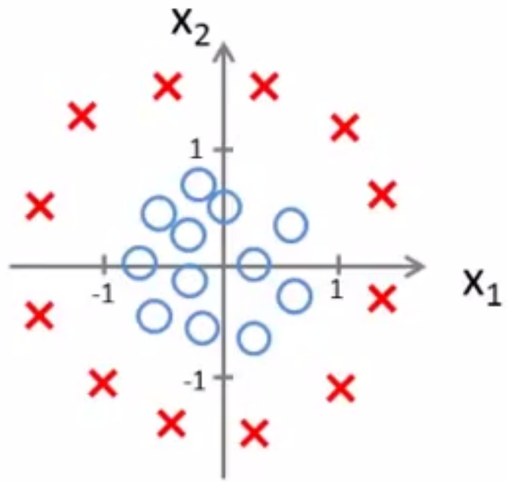
Fronteira de Decisão



$$h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$$

$$\text{“}y = 1\text{“ if } \; -3 + x_1 + x_2 \geq 0$$

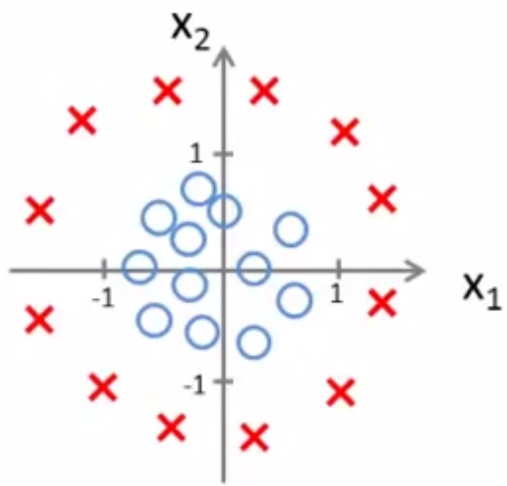
Fronteira de Decisão não-linear



$$h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2)$$

"$y = 1$" if $-1 + x_1^2 + x_2^2 \geq 0$

Fronteira de Decisão não-linear



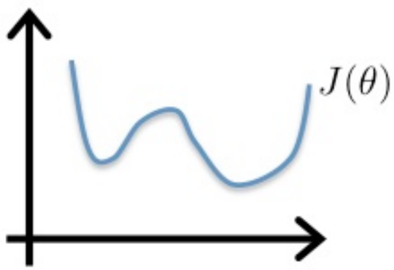
$$h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2)$$

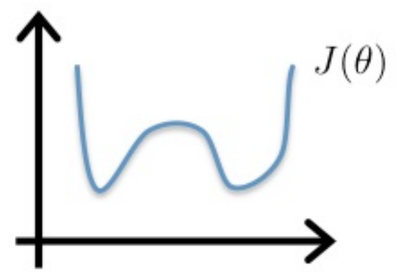
$$\text{“}y = 1\text{“ if } \; -1 + x_1^2 + x_2^2 \geq 0$$

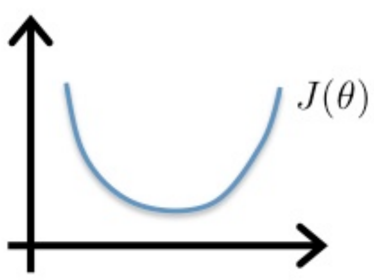
$$h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_1^2 x_2 + \theta_5 x_1^2 x_2^2 + \theta_6 x_1^3 x_2 + \dots)$$

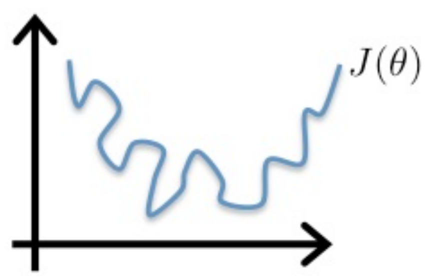
Função de custo

Qual das funções abaixo é convexa?









Função de custo para regressão logística

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

Função de custo para regressão logística

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{Cost}(h_\theta(x^{(i)}), y^{(i)})$$

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

Função de custo para regressão logística

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\mathrm{Cost}(h_\theta(x^{(i)}), y^{(i)})$$

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

$$\Downarrow$$

$$\begin{aligned} J(\theta) &= \frac{1}{m}\sum_{i=1}^{m}\mathrm{Cost}(h_\theta(x^{(i)}), y^{(i)}) \\ &= -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)}\log h_\theta(x^{(i)}) + (1 - y^{(i)})\log(1 - h_\theta(x^{(i)}))\right] \end{aligned}$$

Método do descenso do gradiente

$$J(\theta) = -\frac{1}{m}[\sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_\theta(x^{(i)}))]$$

Want $\min_\theta J(\theta)$:

Repeat $\{$

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

(simultaneously update all θ_j)

$\}$

Método do descenso do gradiente

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log\left(1 - h_\theta(x^{(i)})\right)\right]$$

Want $\min_\theta J(\theta)$:

Repeat $\{$

$$\theta_j := \theta_j - \alpha \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)}$$

(simultaneously update all θ_j)

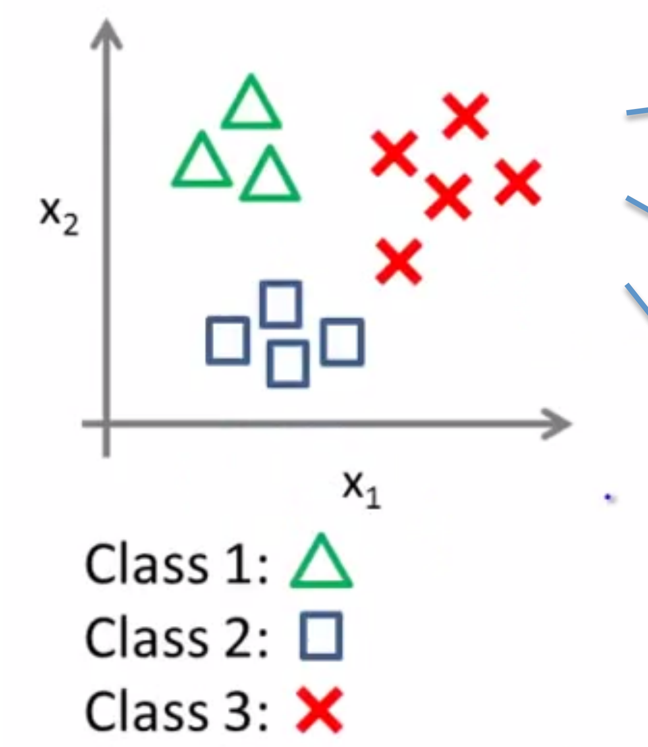
$\}$

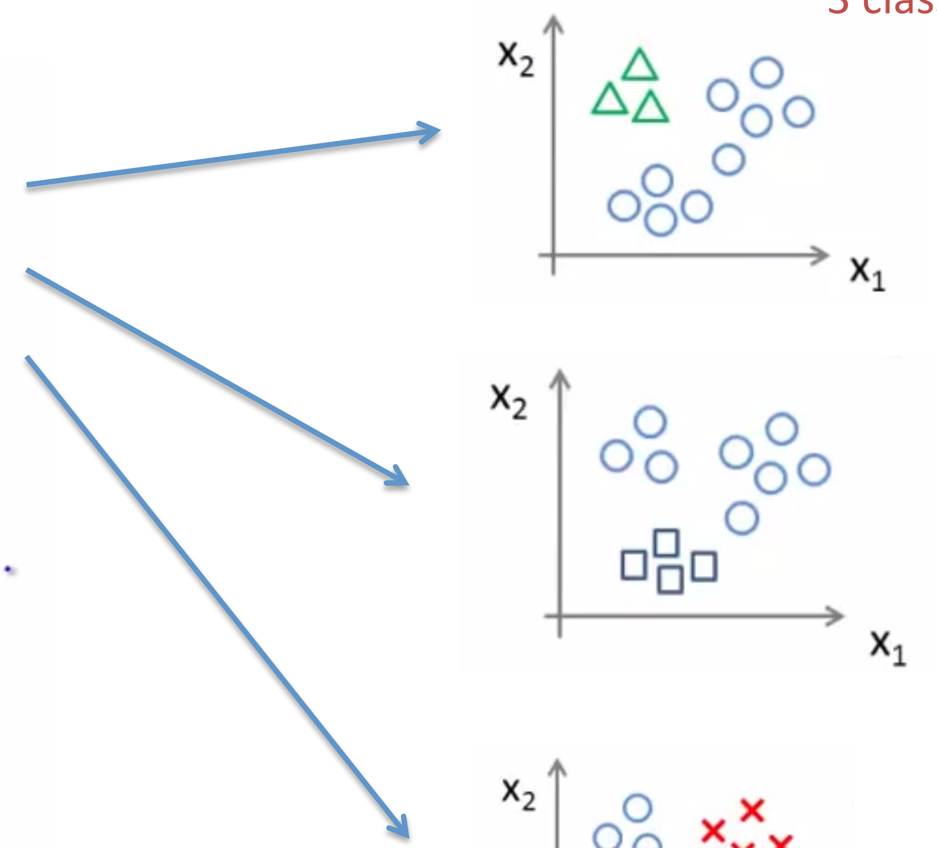
Algoritmo parece idêntico a regressão linear!

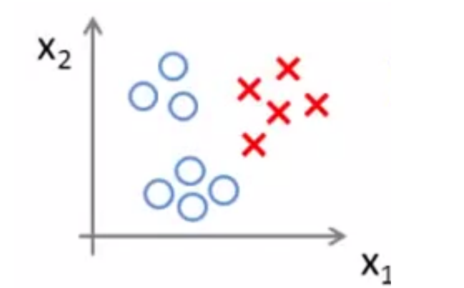
Mas a hipótese é diferente.

Classificação com múltiplas classes

3 classificadores







$$h_\theta^{(i)}(x) = P(y = i | x; \theta) \qquad (i = 1, 2, 3)$$

Classificação com múltiplas classes

- One-vs-all