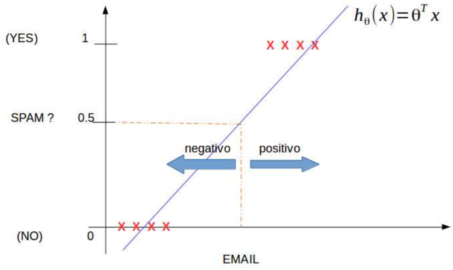
Inteligência Artificial
Aula 10 - vídeo 3 - Regressão Linear e Logística

João C. P. da Silva

Dept. Ciência da Computação - UFRJ

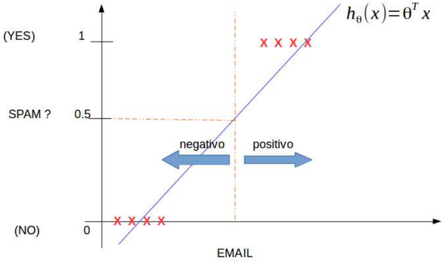
28 de outubro de 2020

Classificação - Regressão Logística



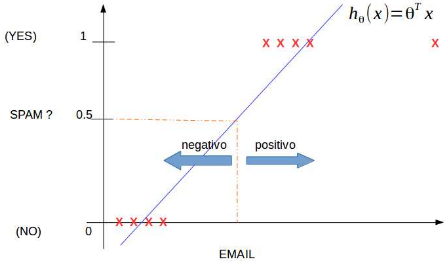
- Limiar de classificação da saída $h_\theta(x)$ em 0.5:
  - Se $h_\theta(x) \geq 0.5$, faça $y = 1$
  - Se $h_\theta(x) < 0.5$, faça $y = 0$

Classificação - Regressão Logística



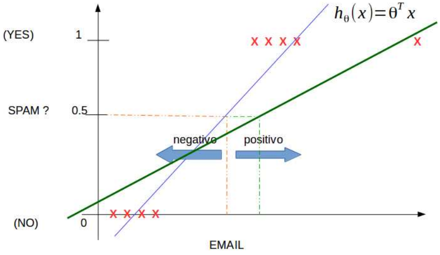
- Limiar de classificação da saída $h_\theta(x)$ em 0.5:
  - Se $h_\theta(x) \geq 0.5$, faça $y = 1$
  - Se $h_\theta(x) < 0.5$, faça $y = 0$

Classificação - Regressão Logística



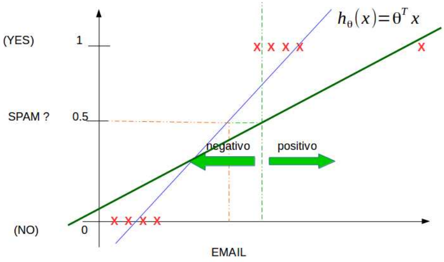
- Limiar de classificação da saída $h_\theta(x)$ em 0.5:
  - Se $h_\theta(x) \geq 0.5$, faça $y = 1$
  - Se $h_\theta(x) < 0.5$, faça $y = 0$

Classificação - Regressão Logística



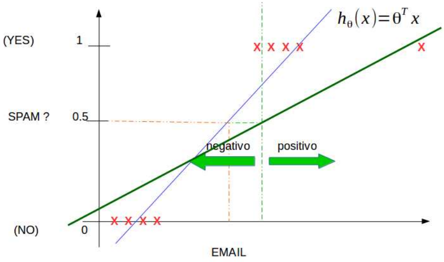
- Limiar de classificação da saída $h_\theta(x)$ em 0.5:
  - Se $h_\theta(x) \geq 0.5$, faça $y = 1$
  - Se $h_\theta(x) < 0.5$, faça $y = 0$

Classificação - Regressão Logística



- Limiar de classificação da saída $h_\theta(x)$ em 0.5:
  - Se $h_\theta(x) \geq 0.5$, faça $y = 1$
  - Se $h_\theta(x) < 0.5$, faça $y = 0$

Classificação - Regressão Logística

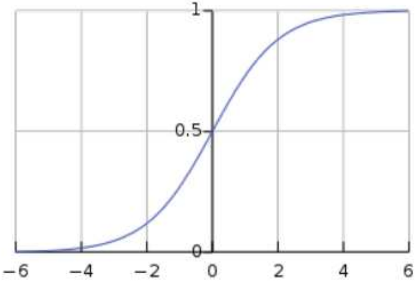


- Limiar de classificação da saída $h_\theta(x)$ em 0.5:
  - Se $h_\theta(x) \geq 0.5$, faça $y = 1$
  - Se $h_\theta(x) < 0.5$, faça $y = 0$
- Não parece ser uma boa ideia usar regressão linear para problemas de classificação.

Classificação - Regressão Logística

- Note que $y \in \{0, 1\}$, mas $h_\theta(x)$ pode ser > 1 ou < 0
- **Regressão Logística**: $0 \leq h_\theta(x) \leq 1$
- Para isso, nossa hipótese $h_\theta(x) = \theta^T x$ é transformado na hipótese $h_\theta(x) = g(\theta^T x)$, onde $g(z) = \frac{1}{1+e^{-z}}$ é a função *sigmoide/logística*. Logo,

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



Interpretação de $h_\theta(x)$: probabilidade estimada que $y = 1$ dado o input x.

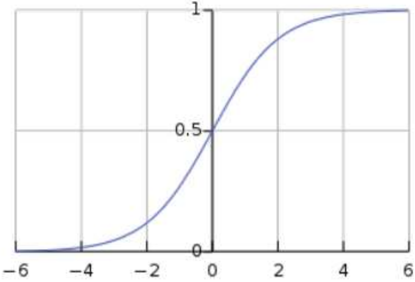
Ou seja, $h_\theta(x) = P(y = 1 \mid x; \theta)$.

Como consequência, $P(y = 0 \mid x; \theta) = 1 - P(y = 1 \mid x; \theta)$

Exemplo

Se $x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ tam.tumor \end{bmatrix}$

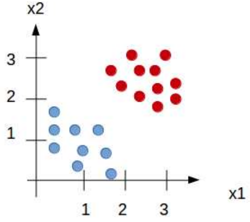
$h_\theta(x) = 0.7$ significa que existe uma chance de 70% do tumor ser malígno.



- Note que:
  - Se $h_\theta(x) = g(\theta^T x) \geq 0.5$, prevemos $y = 1$
  - Se $h_\theta(x) = g(\theta^T x) < 0.5$, prevemos $y = 0$
- Mas
  - $g(\theta^T x) \geq 0.5$ sempre que $\theta^T x \geq 0$ e
  - $g(\theta^T x) < 0.5$ sempre que $\theta^T x < 0$

Classificação - Regressão Logística

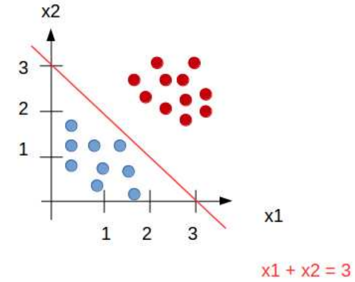
Exemplo



- Temos $h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$
- Considere os parâmetros $\theta_0 = -3, \theta_1 = 1, \theta_2 = 1$
- Logo, teremos $y = 1$ se $-3 + x_1 + x_2 \geq 0$, ou, $x_1 + x_2 \geq 3$

Classificação - Regressão Logística

Exemplo



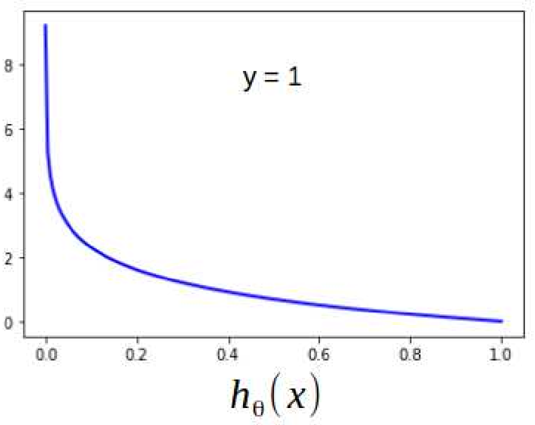
- Temos $h_\theta(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$
- Considere os parâmetros $\theta_0 = -3, \theta_1 = 1, \theta_2 = 1$
- Logo, teremos $y = 1$ se $-3 + x_1 + x_2 \geq 0$, ou, $x_1 + x_2 \geq 3$
- A equação $x_1 + x_2 = 3$ corresponde a $h_\theta(x) = 0.5$ (Decision Boundary)

Classificação - Função Custo

- Conjunto de treinamento: $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \cdots, (x^{(m)}, y^{(m)})\}$
- $x = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}$, $x_0 = 1, y \in \{0, 1\}$
- $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$
- Como escolher os parâmetros θ?
- **Regressão Linear**: $J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \frac{1}{2}(h_\theta(x^{(i)}) - y^{(i)})^2 = \frac{1}{m} \sum_{i=1}^{m} Cost(h_\theta(x^{(i)}), y^{(i)})$
- Usar esta função para regressão logística, podemos ter vários mínimos locais por causa de $h_\theta(x)$. ($J(\theta)$ é **não convexa**).
- **Regressão Logística**: $Cost(h_\theta(x), y) = \begin{cases} -log(h_\theta(x)) & \text{se } y = 1 \\ -log(1 - h_\theta(x)) & \text{se } y = 0 \end{cases}$

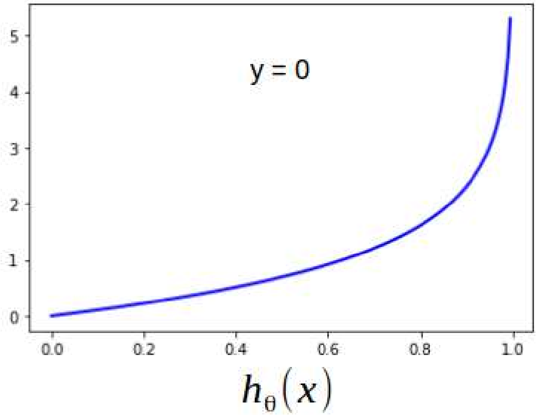
Classificação - Função Custo

Regressão Logística: $Cost(h_\theta(x), y) = \begin{cases} -log(h_\theta(x)) & \text{se } y = 1 \\ -log(1 - h_\theta(x)) & \text{se } y = 0 \end{cases}$



Classificação - Função Custo

Regressão Logística: $Cost(h_\theta(x), y) = \begin{cases} -log(h_\theta(x)) & \text{se } y = 1 \\ -log(1 - h_\theta(x)) & \text{se } y = 0 \end{cases}$



Classificação - Função Custo

Regressão Logística: $Cost(h_\theta(x), y) = \begin{cases} -log(h_\theta(x)) & \text{se } y = 1 \\ -log(1 - h_\theta(x)) & \text{se } y = 0 \end{cases}$

Uma forma equivalente de escrever o custo:

$$Cost(h_\theta(x), y) = -y * log(h_\theta(x)) - ((1 - y) * log(1 - h_\theta(x)))$$

Então:

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

Queremos determinar os parâmetros θ que minimizam $J(\theta)$ e fazer previsões para novos valores de x usando a fórmula $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$

Classificação - Função Custo

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

Para determinar os parâmetros θ que minimizam $J(\theta)$ usamos **gradiente descendente**:

- Repetir até convergir:
    - $\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$
- Atualização simultânea.

Inteligência Artificial
Aula 10 - vídeo 3 - Regressão Linear e Logística

João C. P. da Silva

Dept. Ciência da Computação - UFRJ

28 de outubro de 2020