Regressão Logística: Função Custo Simplificada e Método do Gradiente

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ
Engenharia Elétrica

Onde estamos e para onde vamos

Na aula anterior, definimos a **função perda** e a **função custo** para a Regressão Logística.

Nesta aula, iremos simplificar a definição da função custo e aplicar o Método do Gradiente para encontrar seu mínimo global.

Na aula passada, vimos que a **Função de perda** para Regressão Logística é dada por

$$L\left(f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right), y^{(i)}\right) = \begin{cases} -\log\left(f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right)\right) & \text{, se } y^{(i)} = 1 \\ -\log\left(1 - f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right)\right) & \text{, se } y^{(i)} = 0 \end{cases}$$

OBS: $y^{(i)}$ é sempre 1 ou 0 (tumor maligno ou não)

Note que é possível simplificar, escrevendo da seguinte forma equivalente:

$$L\left(f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right), y^{(i)}\right) = -y^{(i)} \log\left(f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right)\right) - (1 - y^{(i)}) \log\left(1 - f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right)\right)$$

For the simplified loss function:

$$L(f_{\mathbf{w},b}(\mathbf{x}^{(i)}), y^{(i)}) = -y^{(i)}log(f_{\mathbf{w},b}(\mathbf{x}^{(i)})) - (1 - y^{(i)})log(1 - f_{\mathbf{w},b}(\mathbf{x}^{(i)}))$$

if the target $y^{(i)} = 1$, then what does this expression simplify to?

- $\boxed{\bigcirc}$ $-log(f_{\vec{w},b}(\mathbf{x}^{(i)})$
- $\bigcirc$ $-log(1 - f_{\vec{\mathbf{w}},b}(\mathbf{x}^{(i)}))$

Fonte: **Machine Learning Specialization**, *deeplearning.ai*, Stanford Online, Coursera.org.

Função custo simplificada

A função custo que queremos minimizar na **Regressão Logística** é a média das perdas:

$$J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^{m} L\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right), y^{(i)}\right)$$

onde

$$L\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right), y^{(i)}\right) = -y^{(i)} \log\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right) - (1 - y^{(i)}) \log\left(1 - f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right)$$

Sendo assim, note que podemos reescrever

$$J(\vec{w}, b) = -\frac{1}{m} \sum_{i=1}^{m} \left[y^{(i)} \log\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right) + (1 - y^{(i)}) \log\left(1 - f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right)\right]$$

Função custo simplificada

A função custo que queremos minimizar na Regressão Logística é a média das perdas:

$$J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^{m} L\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right), y^{(i)}\right)$$

onde

$$L\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right), y^{(i)}\right) = -y^{(i)} \log\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right) - (1 - y^{(i)}) \log\left(1 - f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right)$$

Sendo assim, note que podemos reescrever

$$J(\vec{w}, b) = -\frac{1}{m} \sum_{i=1}^{m} \left[y^{(i)} \log\left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right) + (1 - y^{(i)}) \log\left(1 - f_{\vec{w},b}\left(\vec{x}^{(i)}\right)\right)\right]$$

Observação:

- O Método do Gradiente agora pode ser aplicado para encontrar quais são os parâmetros $\vec{w}, b$ que minimizam $J(\vec{w}, b)$
- Lembrando que $J(\vec{w}, b)$ será convexa (único mínimo), ainda que o modelo seja:

$$f_{\vec{w},b}(\vec{x}) = \frac{1}{1 + e^{-(\vec{w} \cdot \vec{x} + b)}}$$

onde $f_{\vec{w},b}(\vec{x})$ é a probabilidade de y ser 1.

Precisamos encontrar os valores de $\overrightarrow{w}, b$ que minimizam

$$J(\overrightarrow{w}, b) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log\left(f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right)\right) + (1-y^{(i)})\log\left(1 - f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right)\right)\right]$$

Sabemos que o Método do Gradiente consiste em repetir:

$$w_j = w_j - \alpha\frac{\partial}{\partial w_j}J(\overrightarrow{w}, b) \qquad j = 1, \cdots, n$$

$$b = b - \alpha\frac{\partial}{\partial b}J(\overrightarrow{w}, b)$$

É possível provar que, para a função $J(\overrightarrow{w}, b)$ acima, tem-se:

$$\frac{\partial}{\partial w_j}J(\overrightarrow{w}, b) = \frac{1}{m}\sum_{i=1}^{m}\left(f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right) - y^{(i)}\right)x_j^{(i)}$$

$$\frac{\partial}{\partial b}J(\overrightarrow{w}, b) = \frac{1}{m}\sum_{i=1}^{m}\left(f_{\overrightarrow{w},b}\left(\overrightarrow{x}^{(i)}\right) - y^{(i)}\right)$$

Método do Gradiente aplicado à Regressão Logística

Método do Gradiente para Regressão Logística consiste em repetir:

$$w_j = w_j - \alpha \left[\frac{1}{m} \sum_{i=1}^{m} \left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right) - y^{(i)} \right) x_j^{(i)} \right]$$

$$b = b - \alpha \left[\frac{1}{m} \sum_{i=1}^{m} \left(f_{\vec{w},b}\left(\vec{x}^{(i)}\right) - y^{(i)} \right) \right]$$

Parece idêntico à Regressão Linear. Porém, devemos lembrar que:

Regressão linear:

$$f_{\vec{w},b}\left(\vec{x}^{(i)}\right) = \vec{w} \cdot \vec{x} + b$$

Regressão Logística:

$$f_{\vec{w},b}(\vec{x}) = \frac{1}{1 + e^{-(\vec{w} \cdot \vec{x} + b)}}$$

Conceitos importantes

Quando estudamos Regressão Linear, vimos alguns conceitos importantes que também podem ser aplicados no contexto de Regressão Logística, quais sejam:

- Monitorar o Método do Gradiente (curva de aprendizado)
- Realizar implementação Vetorizada
- Escalonamento de Características

Todos estes conceitos continuam valendo agora para o contexto de Regressão Logística

De olho no código!

Vamos agora implementar o algoritmo de Regressão Logística na prática.

Nome do arquivo que trabalharemos agora:

codigo - Regressão Logistica.ipynb

Atividade de aula

Parte 1

Rode todo o "codigo - Regressão Logistica.ipynb" sem fazer qualquer tipo de alteração. Certifique-se de que você o compreendeu.

Parte 2

1. Explique, com as suas próprias palavras, quais são os passos necessários para implementar a Regressão Logística na prática.
2. Faça modificações no conjunto de dados que está no código, e verifique como essas modificações alteram a fronteira de decisão do modelo estimado.