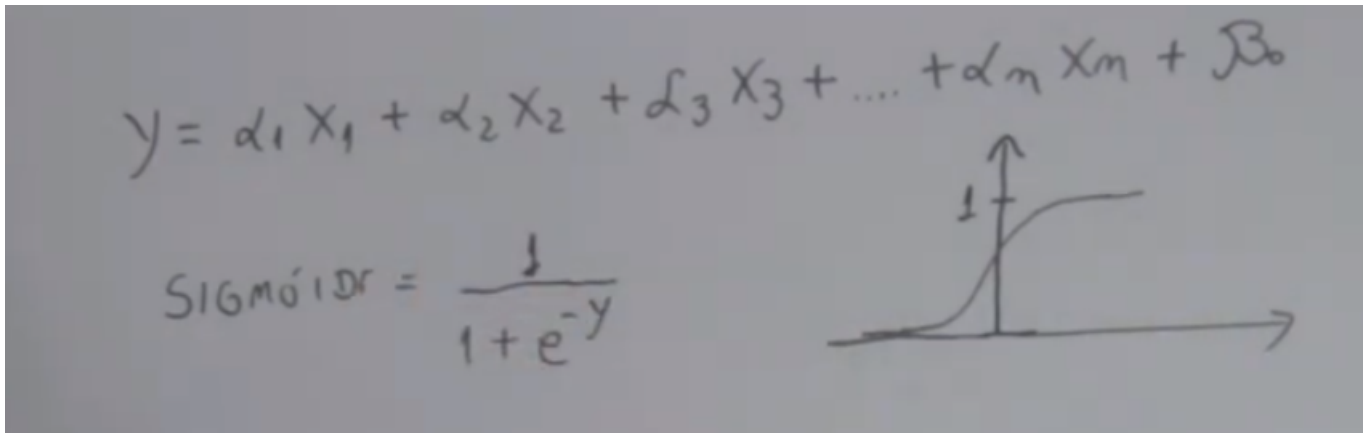


Regressão Logística

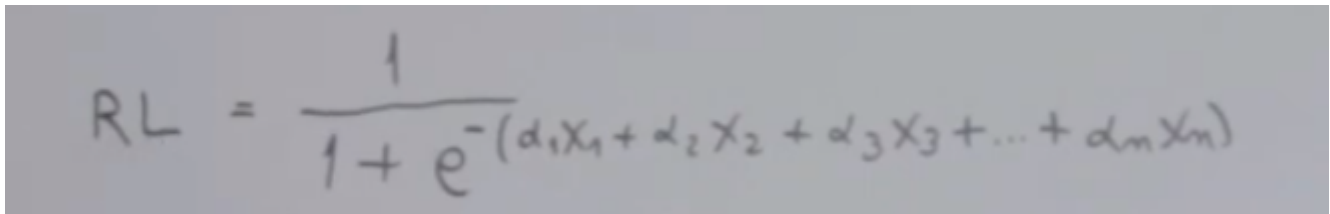
Diferente da regressão linear, esse algoritmo tem como objetivo classificar entre 0 ou 1 ou mais classes.

Ela utiliza a mesma função encontrada na regressão linear, onde há um peso multiplicado por X_i e somado com o intercepto B , porém é aplicada a função sigmoide em cima do resultado dessa função y conforme abaixo.



The image shows handwritten mathematical expressions and a graph. At the top, the linear combination is written as $y = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \dots + \alpha_n X_n + \beta_0$. Below this, the sigmoid function is defined as $\text{SIGMOIDE} = \frac{1}{1 + e^{-y}}$. To the right of the equation is a hand-drawn graph of the sigmoid function, which is an S-shaped curve starting near 0, passing through 0.5 at the origin, and approaching 1 as the input increases. The y-axis is labeled with '1' at the top.

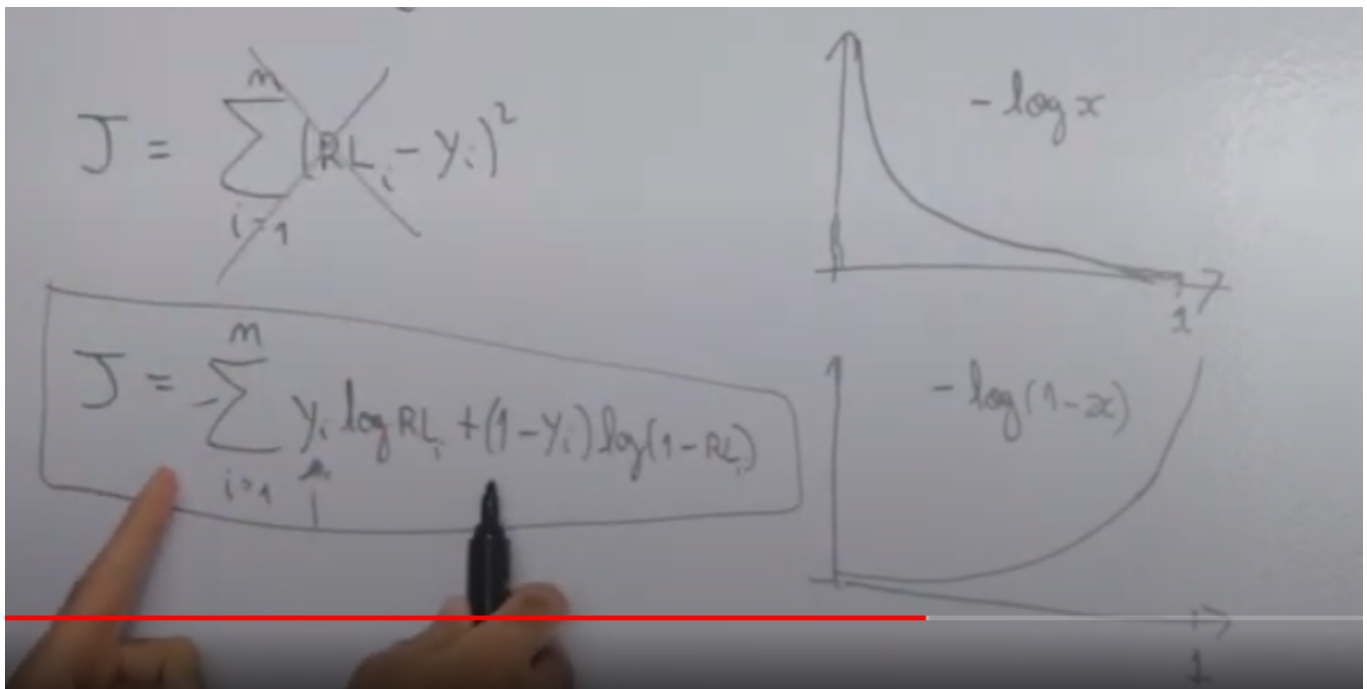
$$y = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \dots + \alpha_n X_n + \beta_0$$
$$\text{SIGMOIDE} = \frac{1}{1 + e^{-y}}$$



The image shows the logistic regression formula written in a handwritten style: $RL = \frac{1}{1 + e^{-(\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \dots + \alpha_n X_n)}}$.

$$RL = \frac{1}{1 + e^{-(\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \dots + \alpha_n X_n)}}$$

Função de custo:



A ideia da função de custo, na verdade é utilizar duas formas diferentes para quando o resultado real for 1 ou 0. Caso seja 1, somente o lado esquerdo da função de custo será utilizado, sobrando $-\log RL$, o que resulta no primeiro gráfico. Caso o resultado seja mesmo 1 o custo é próximo de 0, caso seja uma predição incorreta, o custo tende ao infinito. Da mesma forma acontece com o 0, ao cancelar o lado esquerdo, sobra somente $-\log(1-x)$, que resulta no segundo gráfico de cima pra baixo, onde se a predição estiver correta, tende a custo 0 e se for errado, no caso previu o número 1, tende ao infinito. Por fim soma-se todos os erros no final para saber como o modelo está se comportando.

SURTIU DA EQUAÇÃO DA VEROSSIMILHANÇA

Para múltiplas classes:

Basta refazer o procedimento, porém treinando uma função para cada classe. Por exemplo, dado 3 classes, A, B e C, devemos treinar o modelo para prever A, depois B, depois C, e utilizar essas 3 funções para classificar uma nova amostra. O resultado disso pode ser, 0.8 para classe A, 0.2 para classe B, e 0.5 para classe C, portanto a classe que pertence é a A.