

# Regressão Linear Com Múltiplas Variáveis

Beariz Camargo Câmara<sup>1</sup>

<sup>1</sup>Inteligência Artificial - Universidade Federal de Mato Grosso do Sul (UFMS-CPPP)

## 1. Regressão Linear com Múltiplas Variáveis

A análise de regressão linear com múltiplas variáveis é semelhante a regressão linear, no entanto teremos mais de uma variável independente.

## 2. Descrição do Trabalho

A tarefa consistiu na implementação da regressão linear para prever a qualidade de um determinado vinho tinto com base em suas características disponíveis que são: *fixed acidity*, *volatile acidity*, *citric acid*, *residual sugar*, *chlorides*, *free sulfur dioxide*, *total sulfur dioxide*, *density*, *pH*, *sulphates*, *alcohol*. O dataset utilizado foi o Wine Quality, disponível em: UCI Machine Learning Repository

Iniciou-se o  $\theta = 0$  e ao decorrer dos testes houve a atualização dos valores de  $\alpha$  e quantidade de iterações.

## 3. Código

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3 from reportlab.pdfgen import canvas
4 import pandas as pd
5
6
7 # reading the archives
8 wine = pd.read_csv('winequality-red.csv', delimiter=';')
9
10 # getting the x and y values
11 x = wine[wine.columns[:-1]].to_numpy()
12 y = wine[wine.columns[-1:]].to_numpy()
13
14 # normalizing x
15 X = np.divide(x-np.mean(x, axis=0), np.std(x, axis=0))
16 print(np.shape(X))
17 m = len(wine.columns[:-1])
18
19
20 # adding a column with 1
21 X = np.insert(X, 0, 1, axis=1)
22 print(np.shape(X))
23
24 theta = np.zeros(m+1).reshape(m+1, 1)
25
26 alpha = 0.0001
27 it = 1000
```

```

28 J = np.zeros(it)
29 e = np.zeros(it)
30 print(np.shape(theta))
31 for i in range(it):
32     H = np.dot(X, theta)
33     E = H - y
34     temp = (np.dot(E.T, E))
35     J = np.divide(temp, (2*m))
36     e[i] = J
37     theta = theta - (alpha * (np.dot(X.T, E)))
38
39 plt.figure()
40 plt.title("Loss X Updates")
41 plt.xlabel('J')
42 plt.ylabel('iteracoes')
43 plt.plot(e[-20:])
44 plt.show()

```

## 4. Comparação dos Resultados Obtidos

### 4.1. Primeira tentativa

Como primeiro teste foi utilizado o, alpha,  $\alpha = 1$  e o, número de iterações,  $it = 10$ . Nesse caso foi escolhido um alpha grande e um número de iterações pequeno para demonstrar o que ocorreria caso divergisse.

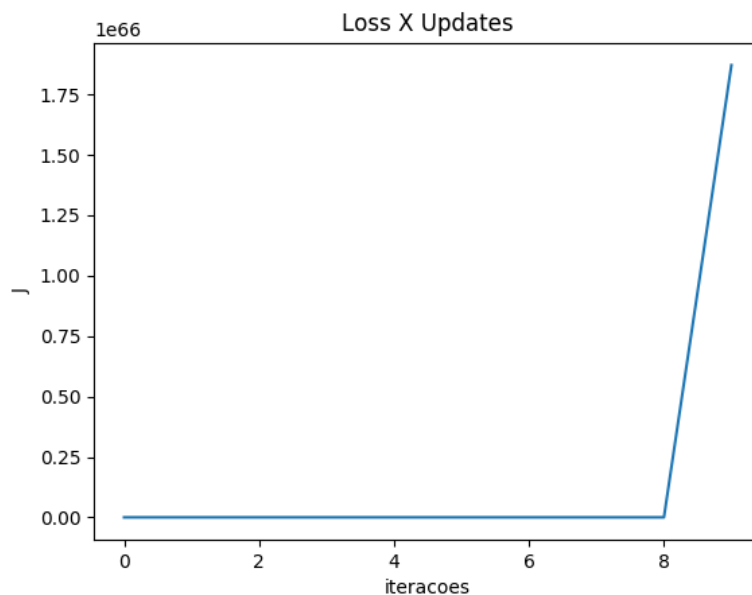
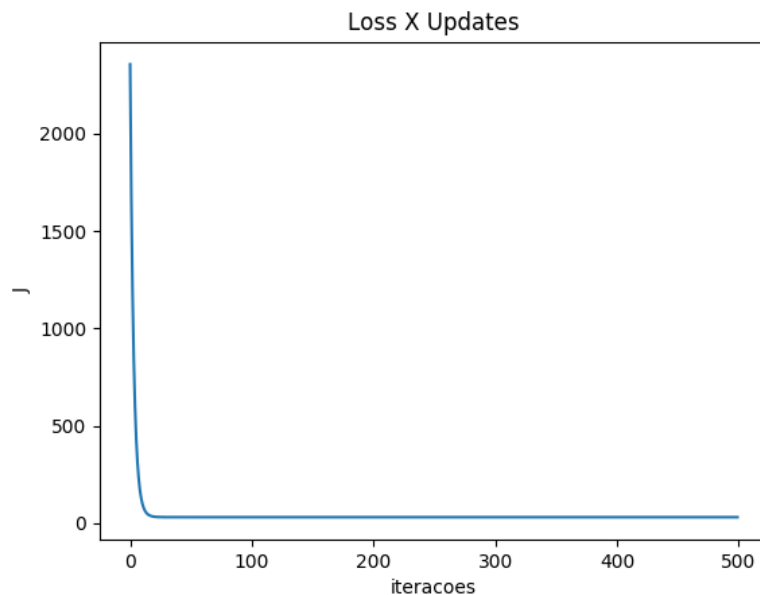


Figure 1. Gráfico 1

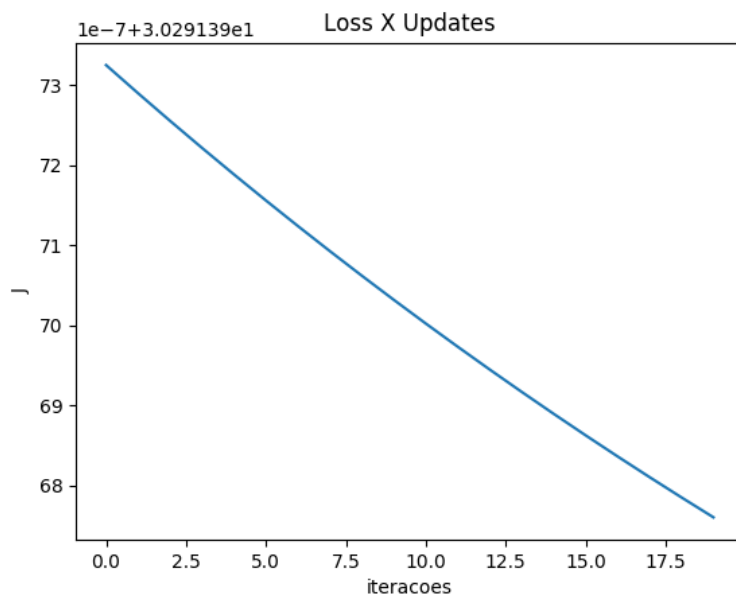
### 4.2. Segunda tentativa

No segundo teste diminuimos o tamanho do alpha para  $\alpha = 0.0001$  e aumentamos o número de iterações,  $it = 500$ . É interessante observar que se levamos em consideração apenas

o primeiro gráfico assumimos que nosso modelo convergiu. No entanto ao fazermos o recorte dos ultimos 20 valores, podemos observar que ainda da para melhorar um pouco .



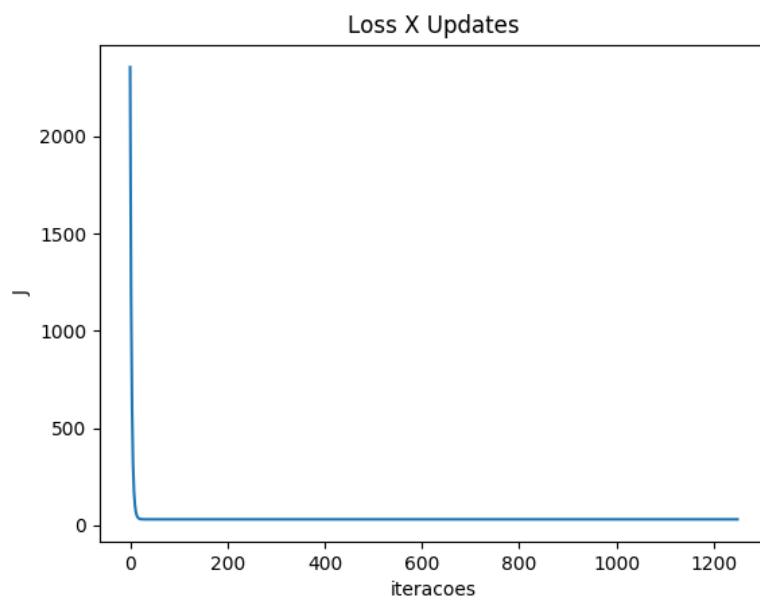
**Figure 2. Gráfico 2**



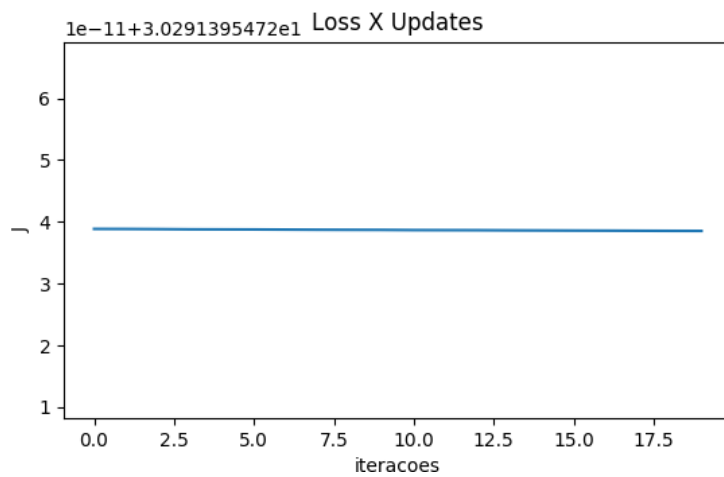
**Figure 3. Gráfico 2.1 - Recorte dos ultimo 20 valores**

### 4.3. Tentativa Final

Na tentativa final podemos observar que houve a convergência total do nosso custo. Para isso utilizamos o  $a = 0.0001$  e o  $it = 1250$ .



**Figure 4. Gráfico 3**



**Figure 5. Gráfico 3.1 - Recorte dos ultimo 20 valores**