

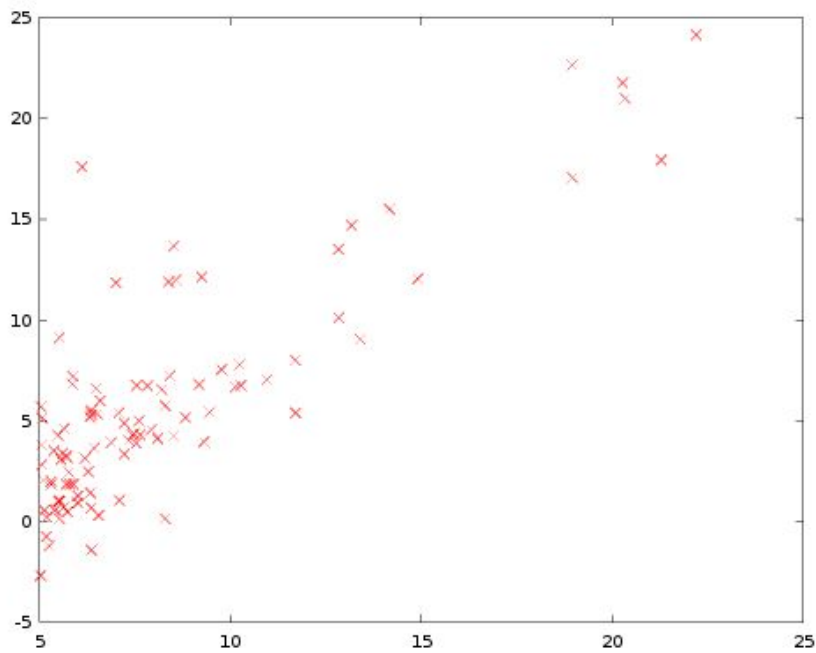
Lista 1

Nome: Carolina Herbster

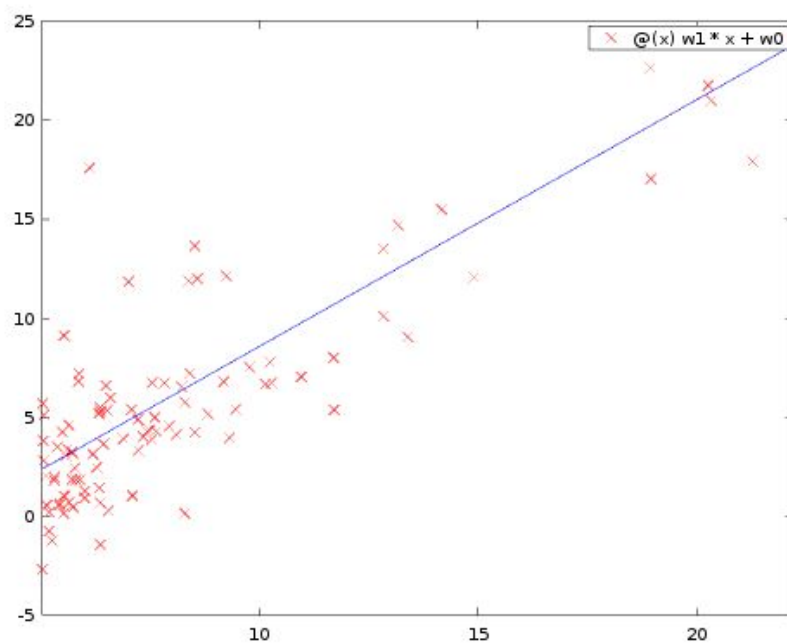
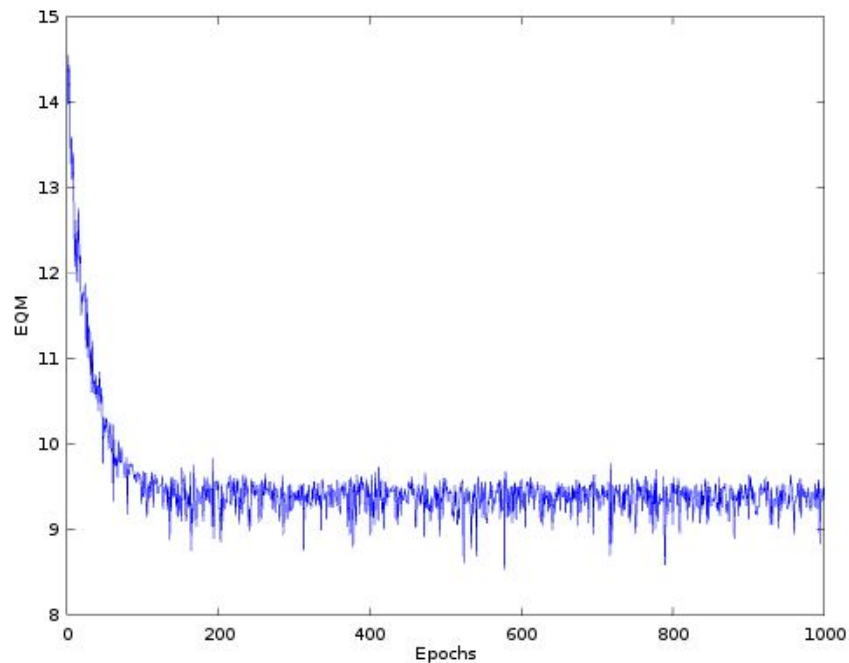
Matrícula: 354044

1 -

- Um modelo linear parece ser adequado, pois podemos passar uma linha por esse conjunto de dados de modo que os valores sejam aproximados com um erro não muito grande.

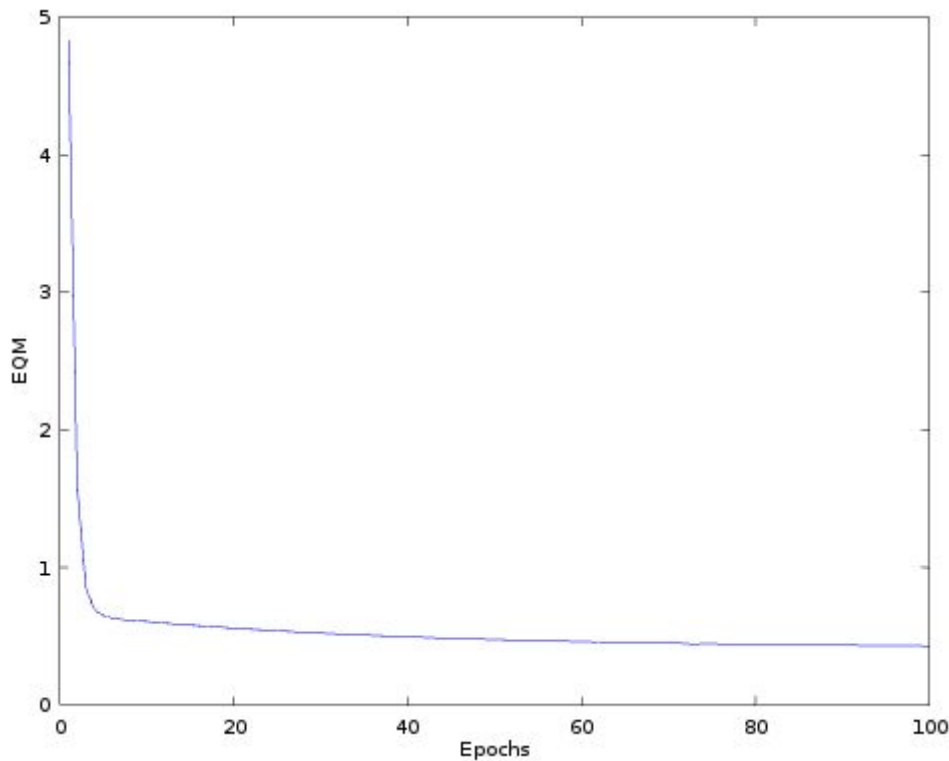


- Com valores iniciais de $w_0 = 1$ e $w_1 = 1$, os valores finais ficam $w_0 = -3.8668$ e $w_1 = 1.2452$. De acordo com o gráfico do erro médio em relação à época, podemos perceber uma grande diminuição do erro médio no início e uma relativa estabilização nas épocas mais posteriores, por isso podemos dizer que o algoritmo está aprendendo. Plotando o gráfico da função linear, podemos observar que ela se ajusta bem aos dados.



2 -

- Usando o gradiente descendente estocástico com chute inicial de [1; 1; 1]: Coeficientes: [0.762764; 1.216404; 0.067038]. Observando o gráfico de EQM X Épocas, podemos perceber que o erro diminui bastante à medida que se passam as épocas, até ficar mais ou menos estabilizado, o que indica o “aprendizado”.



- Usando os mínimos quadrados: Coeficientes: [0.895979; 1.392107; -0.087380]. Os valores são diferentes do gradiente descendente estocástico, pois o gradiente depende do chute inicial, do alfa e da quantidade de épocas de treinamento.

3 -

- Os valores dos coeficientes para cada lambda são:

- **(lambda = 0)**

0.792989

2.056501

0.121439

-1.161379

0.062729

-0.391102

- **(lambda = 1)**

0.218377
1.339161
0.225068
0.180160
0.048057
-0.247075

- **(lambda = 2)**

0.258168
1.162657
0.230547
0.249108
0.057040
-0.115261

- **(lambda = 3)**

0.318750
1.049327
0.233039
0.259881
0.068043
-0.031174

- **(lambda = 4)**

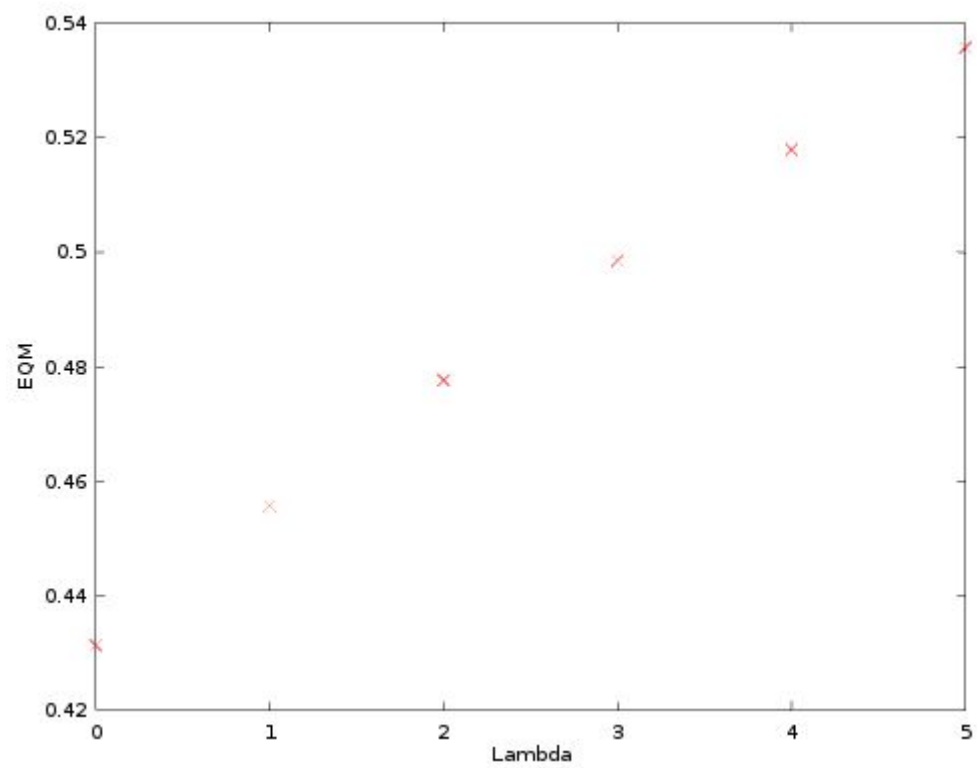
0.381665
0.965942
0.234412
0.257707
0.078538
0.025469

- **(lambda = 5)**

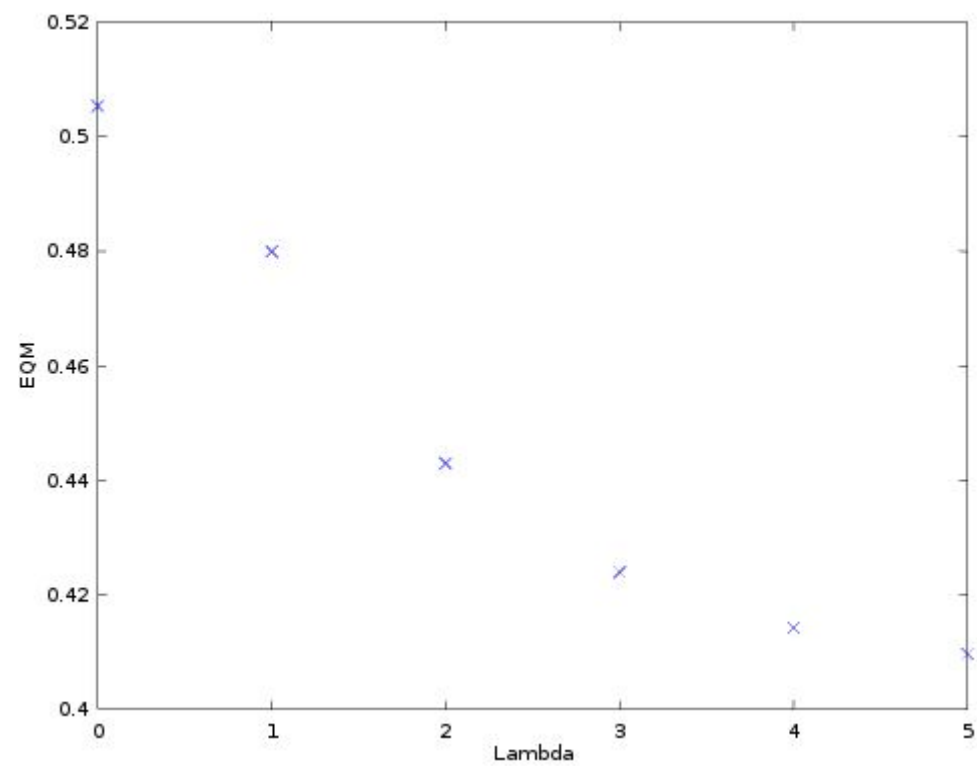
0.443831
0.900535
0.234909
0.251674
0.087866
0.065334

- Quanto mais aumentamos o valor do lambda, menos diminuimos os valores dos coeficientes w_4 e w_5 , logo elas parecem ser as menos relevantes.

- Erro no treinamento:



- Erro no teste:



- Os valores dos coeficientes diminuem com o λ , pois ele aumenta o custo dos coeficientes na função de custo. Logo, isso diminui o overfitting do modelo em relação ao conjunto de treinamento, tornando-o mais adequado ao conjunto de teste.