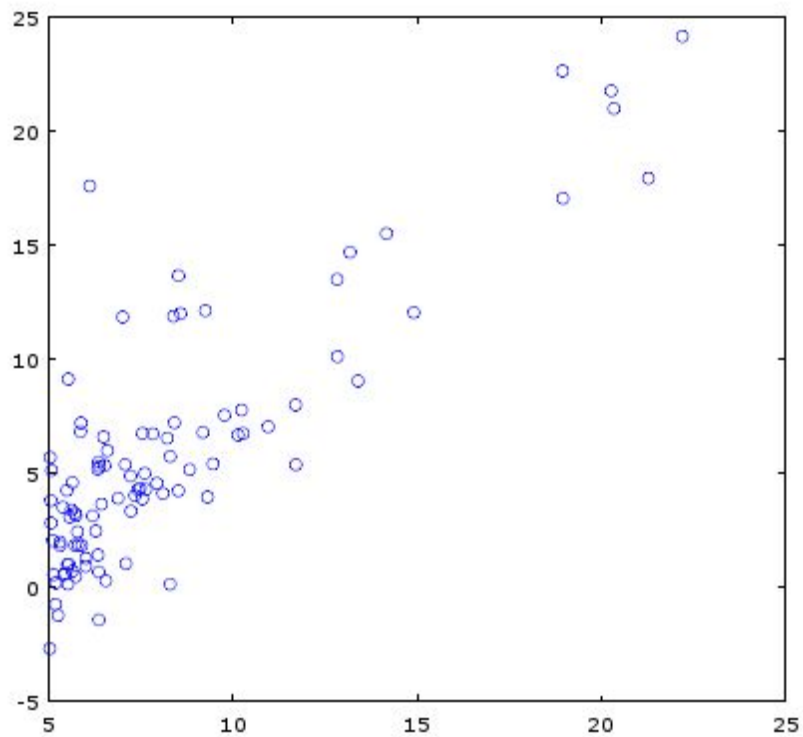


Lista 1

Ricardo Bustamante de Queiroz

Regressão Linear Univariada

Apresentar: Figura com os dados

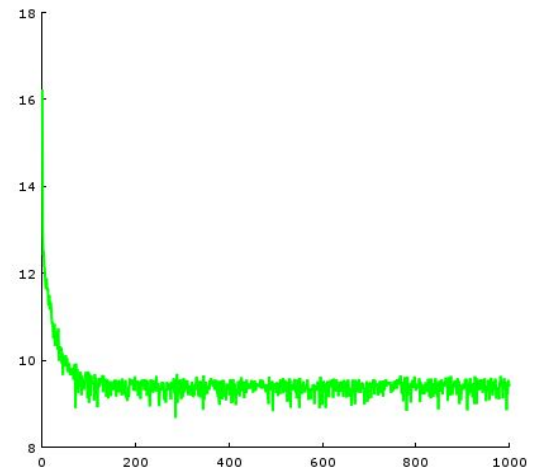
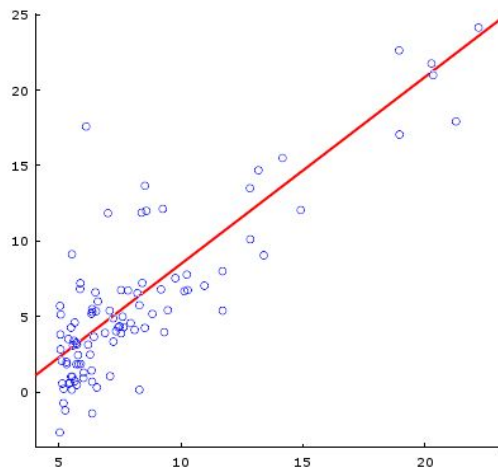


Comentários: Um modelo de regressão linear parece ser adequado para os dados em questão? Comente.

Parece. Os dados estão alongados na diagonal, o que indica que provavelmente o melhor modelo para aproximá-los é uma reta.

Apresentar: Valor final dos coeficientes e gráfico épocas x EQM.

w double 2x1 [-3.8897; 1.2379]



Comentários: Através do gráfico “épocas x EQM” é possível verificar que o algoritmo está “aprendendo” ? Comente.

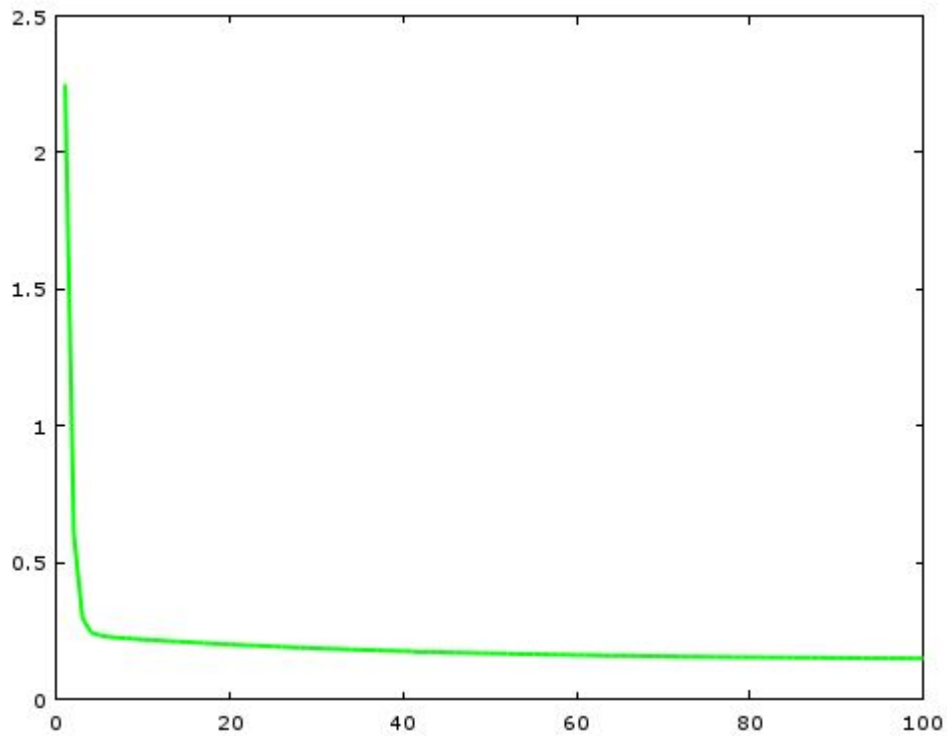
Sim! O EQM está caindo a cada época e por volta da época 200 já parece ter convergido pra algo entre 9 e 10.

Regressão Linear Múltipla

Apresentar: Valor final dos coeficientes e gráfico épocas x EQM.

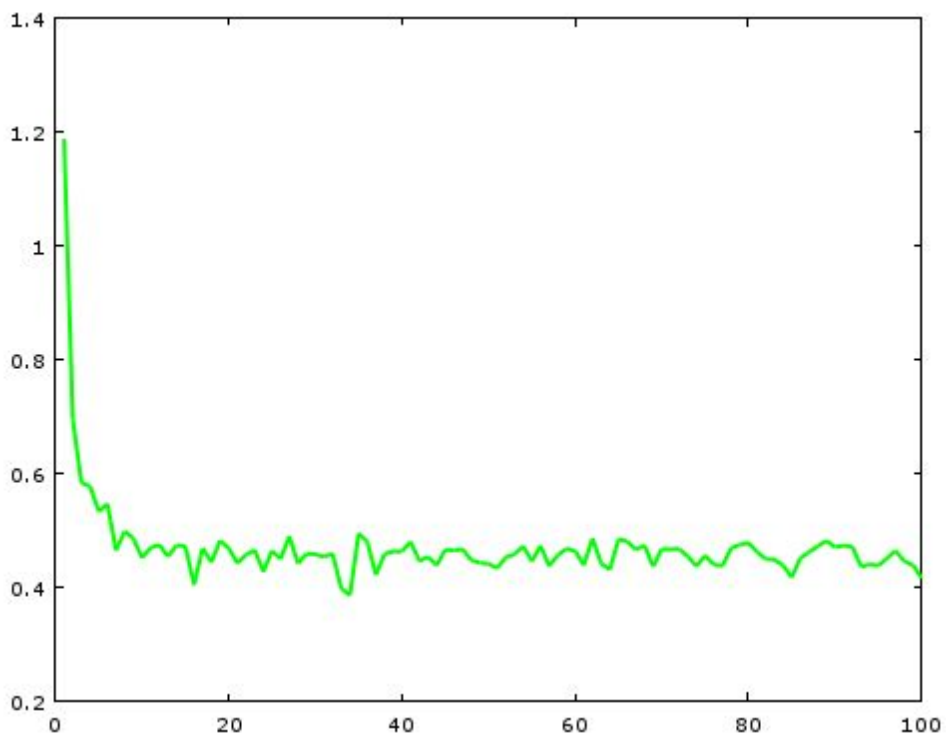
fiz por engano com $\alpha = 0.001$:

w	double	1x3	[0.24791, 1.2324, 0.20586]
---	--------	-----	----------------------------



depois de corrigir, com $\alpha = 0.01$:

`w` `double` `1x3` `[0.86271, 1.3745, -0.10712]`



Comentários: Através do gráfico “épocas x EQM” é possível verificar que o algoritmo está “aprendendo” ? Comente.

Sim. Ele parece estar convergindo por volta da época 60.

- Encontre os coeficientes da regressão utilizando o método dos mínimos quadrados. Apresentar: Valor final dos coeficientes

$w = 0.895979, 1.392107, -0.087380$

Comentários: Os valores obtidos pelos dois métodos são iguais? Comente.

Os valores ficaram bem próximos. Em algumas execuções do método estocástico fica levemente acima, e em outras levemente abaixo.

- Divida o conjunto de dados entre treino e teste. Para este problema, os primeiros 30 dados serão utilizados para treino e o restante será usado para teste.

- Encontre os coeficientes da regressão utilizando o método dos mínimos quadrados regularizado para os seguintes valores de $\lambda = [0 \ 1 \ 2 \ 3 \ 4 \ 5]$. Utilize o conjunto de treinamento.

Apresentar: Valores finais dos coeficientes

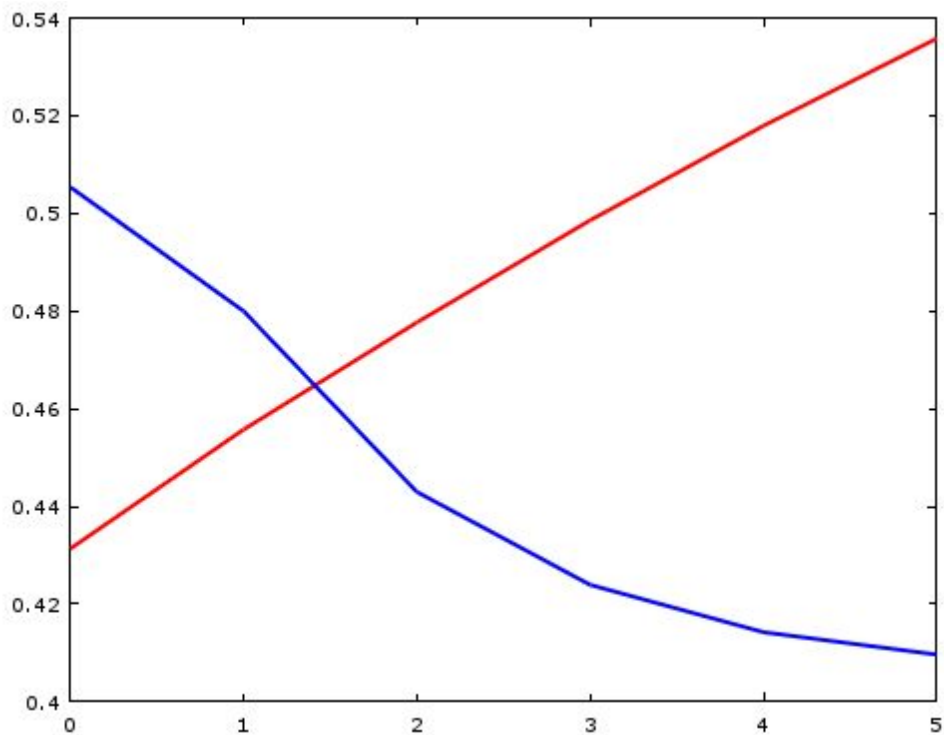
$\lambda = 0$; $w = 0.792989, 2.056501, 0.121439, -1.161379, 0.062729, -0.391102$
 $\lambda = 1$; $w = 0.218377 \ 1.339161 \ 0.225068 \ 0.180160 \ 0.048057 \ -0.247075$
 $\lambda = 2$; $w = 0.258168 \ 1.162657 \ 0.230547 \ 0.249108 \ 0.057040 \ -0.115261$
 $\lambda = 3$; $w = 0.318750 \ 1.049327 \ 0.233039 \ 0.259881 \ 0.068043 \ -0.031174$
 $\lambda = 4$; $w = 0.381665 \ 0.965942 \ 0.234412 \ 0.257707 \ 0.078538 \ 0.025469$
 $\lambda = 5$; $w = 0.443831 \ 0.900535 \ 0.234909 \ 0.251674 \ 0.087866 \ 0.065334$

Comentários: Quais variáveis parecem ser menos relevantes para a regressão?

As que possuem menor valor tem menor influência no valor final. Aparentemente os dois últimos valores de w possuem menor relevância, dependendo do valor de λ .

- Encontrar o valor do EQM para os dados de treinamento e de teste para cada um dos valores de λ .

Apresentar: Dois gráficos. EQM x λ no conjunto de treinamento e EQM x λ no conjunto de teste



**Comentários: Como os valores dos coeficientes variam com λ ? Explique o motivo.
Comente o crescimento/decrescimento dos erros presente nas figuras EQM x λ**

O conjunto de treinamento (linha vermelha) treina melhor com todos os W presentes, mas isso pode causar overfitting. Isso explicaria a diminuição do erro para o caso mais genérico, adicionando os elementos de teste.