

Relatório 4 - Regressão linear

1 Introdução

Os resultados das funções desse relatório foram obtidos com base na seguinte tabela provida pelo professor:

x	y
0.00	2.26
0.50	3.80
1.00	4.43
1.50	5.91
2.00	6.18
2.50	7.26
3.00	8.15
3.50	9.14
4.00	10.87
4.50	11.58
5.00	12.55

A regressão linear busca o modelamento da relação entre variáveis, especificamente a relação linear entre a variável dependente e as variáveis independentes (explanatórias). A regressão linear simples contém apenas uma variável explanatória e é dada pela seguinte fórmula:

$$y = ax + b$$

Sendo o y a variável dependente, x a variável independente, a o coeficiente angular e b o intercepto.

Além da regressão linear simples, há também a regressão linear múltipla que é dada pela seguinte fórmula:

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \dots + b$$

Neste relatório além de trabalhar com a regressão linear simples ira se trabalhar com o coeficiente de correlação de Pearson e o coeficiente de determinação. A correlação é a medida do grau de relação entre duas variáveis e o coeficiente de correlação de Pearson determina o grau de correlação entre duas variáveis emparelhadas. Já o coeficiente de determinação mede o percentual da variação de y que é explicado pela variação de x .

2 Objetivo

- Implementar um Adaline que realize a regressão linear para os dados fornecidos pelo professor
- Comparar os resultados de regressão obtidos utilizando as equações de a e b.
- Encontrar coeficiente de correlação de Pearson e o coeficiente de determinação.

3 Desenvolvimento

Para realizar o treinamento do adaline foi utilizada a linguagem Common Lisp. Como o código para o treinamento do adaline já estava pronto, já que havia sido implementado no relatório da semana passada, muitas das funções foram reaproveitadas, como por exemplo, toda a parte de treinamento do adaline. Todavia, pela tabela fornecida pelo professor agora possuir apenas uma entrada foi necessário fazer uma alteração na função que gerava os pesos aleatórios para que essa retornasse uma lista com apenas um peso inicial e um bias inicial:

Além disso, para facilitar a execução do programa e das funções foram criadas duas listas, uma com um lista contando os valores de x e 1 (para que o bias possa ser calculado na função do ajuste de peso), e outra lista contendo os valores de y.

Usando essas tabelas na função de treinamento do adaline com coeficiente de aprendizado igual a 0.01 e com condição de parada como sendo 1000 ciclos foram obtidos os seguintes peso e bias:

$$\text{Peso} = 2.0170005 \quad \text{Bias} = 2.4392002$$

Com o peso e bias em mãos foi hora de calcular os valores de a e b. Para facilitar o calculo desses valores foram criadas as duas funções: sum, que soma os valores de uma lista, e average, que retorna a média de uma dada lista de números:

Com essas funções, foi possível implementar as seguintes funções que calculam a e b:

Que retornaram os seguintes valores:

$$a = 2.0058184 \quad b = 2.4518175$$

Com isso pronto era hora de comparar os resultados obtidos com o peso e bias do treinamento com os resultados obtidos com a e b. Para isso foi necessário criar uma função que calcula a regressão linear simples:

Com a função que calcula a regressão linear simples pronta só foi necessário alterar a função de comparação de saídas para que essa retornasse uma comparação entre os resultados obtidos com o peso e o bias e os obtidos com a e b:

Os resultados dessa comparação se mostraram satisfatórios e com um erro baixo.

Por fim, chegou a hora de calcular o coeficiente de correlação de Pearson (r) e o coeficiente de determinação (r^2). Para isso foram criadas as seguintes funções:

Que produzem como resultado:

$$r = 0.9961132 \quad r^2 = 0.99224144$$

4 Conclusão

Foi possível observar, por meio dessa aplicação do adaline, seu poder e utilidade na solução de problemas do mundo real. Os resultados obtidos foram muito satisfatórios, sendo muito próximos não só dos obtidos usando a e b , mas também dos esperados.

Ademais, foi possível observar usando o coeficiente de correlação de Pearson que a correlação entre as variáveis é positiva, já que $r > 0$ e também forte, já que o coeficiente de determinação é igual a 0.99224144.

5 Resultados

5.1 Comparação entre os resultados obtidos utilizando o pesos e o bias e os esperados

Expected output: 2.26	Obtained output: 2.44
Expected output: 3.80	Obtained output: 3.45
Expected output: 4.43	Obtained output: 4.46
Expected output: 5.91	Obtained output: 5.46
Expected output: 6.18	Obtained output: 6.47
Expected output: 7.26	Obtained output: 7.48
Expected output: 8.15	Obtained output: 8.49
Expected output: 9.14	Obtained output: 9.50
Expected output: 10.87	Obtained output: 10.51
Expected output: 11.58	Obtained output: 11.52
Expected output: 12.55	Obtained output: 12.52

Weight: 2.017005 b: 2.4392002

5.2 Comparação entre os resultados obtidos utilizando a e b e os esperados

Expected output: 2.26	Obtained output: 2.45
Expected output: 3.80	Obtained output: 3.45
Expected output: 4.43	Obtained output: 4.46
Expected output: 5.91	Obtained output: 5.46
Expected output: 6.18	Obtained output: 6.46
Expected output: 7.26	Obtained output: 7.47
Expected output: 8.15	Obtained output: 8.47
Expected output: 9.14	Obtained output: 9.47
Expected output: 10.87	Obtained output: 10.48
Expected output: 11.58	Obtained output: 11.48
Expected output: 12.55	Obtained output: 12.48

$$a = 2.0058184 \quad b = 2.4518175$$

5.3 Comparação entre os resultados obtidos utilizando o peso e o bias e utilizando a e b

Expected output: 2.45	Obtained output: 2.44
Expected output: 3.45	Obtained output: 3.45
Expected output: 4.46	Obtained output: 4.46
Expected output: 5.46	Obtained output: 5.46
Expected output: 6.46	Obtained output: 6.47
Expected output: 7.47	Obtained output: 7.48
Expected output: 8.47	Obtained output: 8.49
Expected output: 9.47	Obtained output: 9.50
Expected output: 10.48	Obtained output: 10.51
Expected output: 11.48	Obtained output: 11.52
Expected output: 12.48	Obtained output: 12.52

$$a = 2.0058184 \quad b = 2.4518175$$
$$\text{Weight: } 2.017005 \quad b: 2.4392002$$

6 Bibliografia

- https://en.wikipedia.org/wiki/Simple_linear_regression
- https://pt.wikipedia.org/wiki/Coeficiente_de_correla%C3%A7%C3%A3o_de_Pearson
- https://en.wikipedia.org/wiki/Coefficient_of_determination