

Regressão Geométrica

Mário Leite

...

Quando se tem uma vasta quantidade de dados obtidos experimentalmente, por exemplo, os valores de faturas de energia elétrica durante um ano com valores diferentes (alguns podem ser iguais) é possível obter, por exemplo, a média de gastos com energia elétrica nestes 12 meses passados. Também é possível obter dados mais complexos sobre essa amostra como: *desvio padrão*, *variância*, etc. Mas, como prever um possível valor num próximo mês para se ter uma expectativa de uma despesa adicional, baseando-se nesses valores obtidos nos 12 meses anteriores? É aí que entra o que se chama de “Regressão”, que pode nos dar uma ideia dessa expectativa. Vamos considerar um exemplo simples de fatura de energia elétrica de Agosto/2022 a Julho/2023 mostrado na **tabela 1**, envolvendo o consumo de energia (KwH) versus valor da fatura (R\$):

Ano/mês	Consumo (KwH)	Fatura (R\$)
ago/22	320	120,00
set/22	310	117,00
out/22	311	118,00
nov/22	316	119,00
dez/22	330	130,00
jan/23	326	132,00
fev/23	328	135,00
mar/23	330	132,00
abr/23	329	139,00
mai/23	345	158,00
jun/23	344	157,00
jul/23	342	150,00

Tabela 1

Esses valores tendem a ser agrupados de maneira a representa uma tendência linear e podem ser representados por uma equação de reta do tipo: $y = 1.1867x - 254.83$, com coeficiente de correlação R^2 valendo **0,9319**. Isto quer dizer que é possível prever, com 93.19% de certeza, que para um valor x (KwH) consumido o valor pago será de y (R\$). Mas, observe que os valores consumidos de KwH e as faturas em R\$ seguem um padrão linear nas variações; mas, e se fosse um caso tal como na **tabela 2** em que as relações entre x e y não variam de forma linear!?

X	Y
12	60
14	65
15	70
16	76
17	83
18	90
19	89
19	90
20	100
21	125
22	130
24	145

Tabela 2

Se fossemos aplicar uma regressão linear o valor de R^2 seria bem menor; **0.917**. Por isto, outro tipo de regressão deve ser aplicado: *logarítmica*, *exponencial* ou *geométrica*. O programa “FazRegressaoGeometrica”, codificado em Python, mostra esse tipo de regressão para a **tabela 2**.

```

'''
FazRegressaoGeometrica.py
-----

Faz uma "Regressão Geométrica" para uma equação do tipo  $y=a \cdot e^{bx}$  com
um exemplo de uma amostra de cinco dados.
-----
'''

import math
from decimal import Decimal, getcontext

def FazerRegressao(x, a, b):
    return a * math.exp(b * x)

#-----
def CalcularResiduos(yAtu, yPrev):
    residuos = [(yAtu[i] - yPrev[i])**2 for i in range(len(yAtu))]
    return sum(residuos)

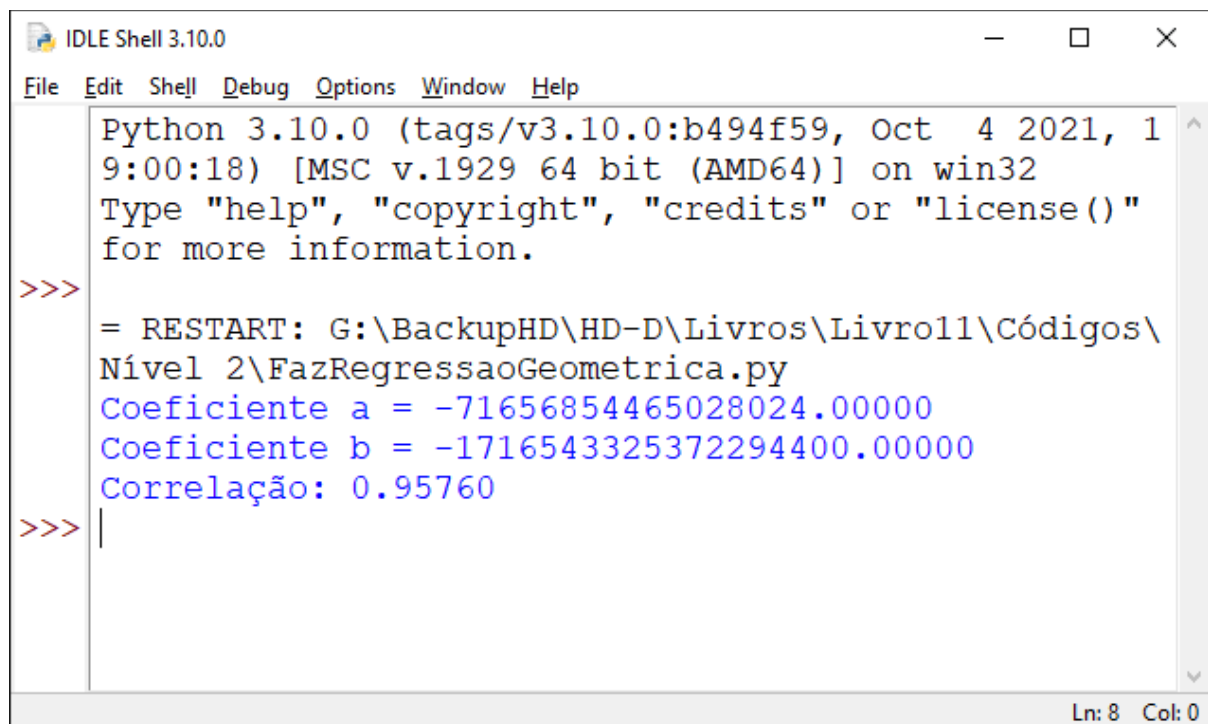
#-----
def FazerGradiente(xDado, yDado, taxa, itera):
    a = 1.0
    b = 1.0
    for _ in range(itera):
        yPrev = [FazerRegressao(x, a, b) for x in xDado]
        gradA = sum([(yPrev[i] - yDado[i]) * math.exp(b * xDado[i]) for
                     i in range(len(yDado))])
        gradB = sum([(yPrev[i] - yDado[i]) * a * xDado[i] * math.exp(b
                               * xDado[i]) for i in range(len(yDado))])

        a -= taxa * gradA
        b -= taxa * gradB
    return a, b

#-----
xDado = [12, 14, 15, 16, 17, 18, 19, 19, 20, 21, 22, 24]
yDado = [60, 65, 70, 76, 83, 90, 89, 90, 100, 125, 130, 145]
taxaCalculada = 0.0001
numIteracoes = 1000
a, b = FazerGradiente(xDado, yDado, taxaCalculada, numIteracoes)
yPrev = [FazerRegressao(x, a, b) for x in xDado]
yReal = sum(yDado) / len(yDado)
somaQuadTot = sum([(yDado[i] - yReal)**2 for i in range(len(yDado))])
somaQuadRes = CalcularResiduos(yDado, yPrev)
n = len(xDado)
sumXY = sum([xDado[i] * yDado[i] for i in range(n)])
sumX = sum(xDado)
sumY = sum(yDado)
sumX2 = sum([x**2 for x in xDado])
sumY2 = sum([y**2 for y in yDado])
numerador = n * sumXY - sumX * sumY
denominador = math.sqrt((n * sumX2 - sumX**2) * (n * sumY2 - sumY**2))
correlacao = numerador / denominador

#=====
print("Coeficiente a =", "{:.5f}".format(float(a)))
print("Coeficiente b =", "{:.5f}".format(float(b)))
print('Correlação:', "{:.5f}".format(correlacao))
#Fim do programa "FazRegressaoGeometrica" -----

```



```
Python 3.10.0 (tags/v3.10.0:b494f59, Oct 4 2021, 19:00:18) [MSC v.1929 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()"
for more information.
>>>
= RESTART: G:\BackupHD\HD-D\Livros\Livro11\Códigos\
Nível 2\FazRegressaoGeometrica.py
Coeficiente a = -71656854465028024.00000
Coeficiente b = -1716543325372294400.00000
Correlação: 0.95760
>>> |
```

Ln: 8 Col: 0

Figura 1 - Saída do programa “FazRegressaoGeometrica”