

Ajuste de Curvas - python

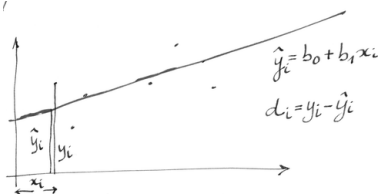
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico xOy conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: no ajuste de curvas busca-se uma expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(sx2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar faça: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coefficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5 \times [116.33 - 22.9^2/5]}$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625

para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

43 44 50 54 59 64 66 68

associados aos seguintes valores de y=

9 10 3 9 4 9 3 9

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

69 70

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 228 318 308 295 534 558 645 350 792 751

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 5.5 5.3 5.7 8.2 6.9 7.2 8.3 7.8 8.4 8.3

Para o ano próximo, o governo pretende efetuar uma emissão de

800

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-68993 - 29/05

Ajuste de Curvas - python

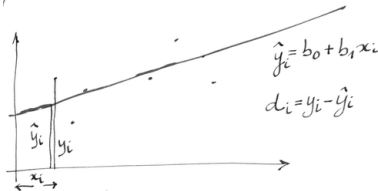
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico xOy conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: no ajuste de curvas busca-se uma expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 \cdot x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 \cdot x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 \cdot x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 \cdot x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i \cdot b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(sx2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar faça: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coeficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5} \times [116.33 - 22.9^2/5]$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625
 para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

45 51 52 54 56 58 59 60

associados aos seguintes valores de y=

143 162 166 176 182 182 192 195

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

61 65

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 223 273 290 223 381 563 471 750 596 662

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 5.7 5.8 5.9 5.1 6.9 6.1 10.3 7.2 9.5 6.1

Para o ano próximo, o governo pretende efetuar uma emissão de

560

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69789 - 29/05

Ajuste de Curvas - python

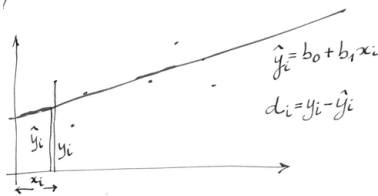
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico $x0y$ conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o *cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro*.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: o ajuste de curvas busca-se a expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 \cdot x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

. Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 \cdot x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 \cdot x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i \cdot b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(s2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar faca: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coeficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5} \times [116.33 - 22.9^2/5]$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625

para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

12 14 16 19 25 29 31 34

associados aos seguintes valores de y=

11 19 19 21 31 28 32 40

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

39 40

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2
-------	-------

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 212 330 211 474 364 327 498 432 817 459

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 5.2 6.9 5.7 7.8 6.2 7.4 10.2 8.2 8.7 6.9

Para o ano próximo, o governo pretende efetuar uma emissão de

423

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69008 - 29/05

Ajuste de Curvas - python

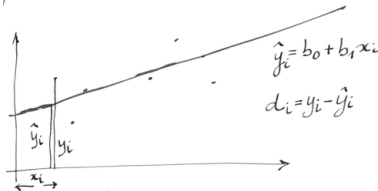
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico $x0y$ conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: no ajuste de curvas busca-se a expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

. Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(sx2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar faça: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coefficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5 \times [116.33 - 22.9^2/5]}$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625
 para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

15 18 19 24 27 29 31 34

associados aos seguintes valores de y=

14 11 13 15 17 10 8 8

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

36 39

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 238 213 340 385 498 503 502 455 803 286

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 5.1 7.1 6.9 7.2 8.5 7.4 9.2 6.2 10.6 5.2

Para o ano próximo, o governo pretende efetuar uma emissão de

716

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69015 - 29/05

Ajuste de Curvas - python

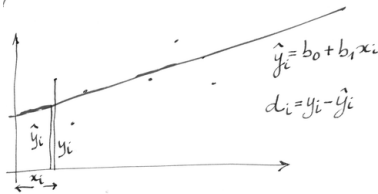
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico x_0y conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: o ajuste de curvas busca-se a expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 \cdot x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

. Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 \cdot x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 \cdot x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i \cdot b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(s2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar fica: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coefficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5} \times [116.33 - 22.9^2/5]$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625
 para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

32 36 40 41 42 45 52 56

associados aos seguintes valores de y=

78 85 89 93 95 107 118 123

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

58 60

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2
-------	-------

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSION ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 259 316 368 286 363 369 658 417 733 828

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 5.2 7.4 5.3 6.4 8.4 5.5 9.6 6.2 9.6 6.5

Para o ano próximo, o governo pretende efetuar uma emissão de

552

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69022 - 29/05

Ajuste de Curvas - python

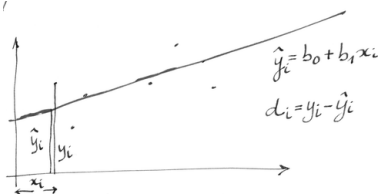
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico $x0y$ conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: no ajuste de curvas busca-se a expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

. Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 - b_1 x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(sx2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar faça: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coefficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5} \times [116.33 - 22.9^2/5]$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625
 para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

42 45 47 52 56 58 59 62

associados aos seguintes valores de y=

177 190 196 223 234 248 250 259

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

65 66

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 245 328 260 360 313 534 333 591 459 506

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 5.5 5.7 6.4 6.1 6.7 8.8 9.1 9.9 7.2 6.8

Para o ano próximo, o governo pretende efetuar uma emissão de

344

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69208 - 29/05

Ajuste de Curvas - python

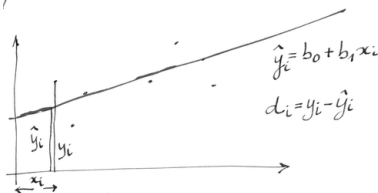
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico x_0y conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o *cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro*.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: o ajuste de curvas busca-se a expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1.x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1.x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1.x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

. Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1.x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1.x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1.x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1.x_i).x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1.x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1.x_i).x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1.x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i . b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i).b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(s2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar faça: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coeficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5} \times [116.33 - 22.9^2/5]$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625

para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

14 17 19 21 23 27 34 35

associados aos seguintes valores de y=

-7 -14 -17 -16 -14 -24 -29 -29

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

36 37

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2
-------	-------

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 229 271 407 235 229 580 615 300 734 531

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 5.1 6.0 7.2 8.5 6.0 5.8 6.7 10.5 10.8 9.0

Para o ano próximo, o governo pretende efetuar uma emissão de

537

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69039 - 29/05

Ajuste de Curvas - python

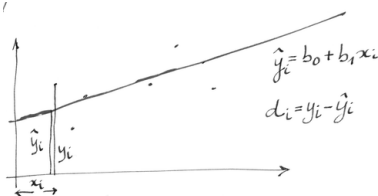
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico $x0y$ conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o *cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro*.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: o ajuste de curvas busca-se a expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 \cdot x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

. Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 \cdot x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 \cdot x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i \cdot b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(s2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
    #para chamar faça: x=[1.3,3.4,5.1,6.8,8.0] e
    # y=[2.0, 5.2, 3.8, 6.1, 5.8]
    #depois: minquad(x,y) e informe quantos x_i
    # quiser, seguido de -1
```

Coeficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5} \times [116.33 - 22.9^2/5]$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625

para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

41 42 44 45 51 53 61 64

associados aos seguintes valores de y=

258 267 277 285 317 324 380 396

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

67 68

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2
-------	-------

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhoes de dolares)
 213 266 212 309 475 405 467 677 346 409

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 6.3 5.7 6.7 7.0 6.2 7.5 8.1 7.9 9.5 9.9

Para o ano próximo, o governo pretende efetuar uma emissão de

962

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69046 - 29/05

Ajuste de Curvas - python

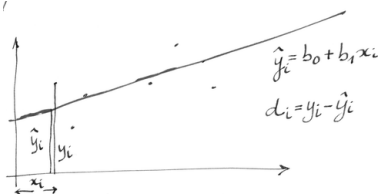
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico $x0y$ conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: no ajuste de curvas busca-se uma expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(sx2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar fica: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coefficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5 \times [116.33 - 22.9^2/5]}$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625
 para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

32 33 36 37 38 41 43 45

associados aos seguintes valores de y=

66 67 74 76 84 85 95 95

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

47 48

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 246 324 222 422 465 367 587 254 777 572

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 6.3 7.3 5.1 5.4 5.7 8.3 9.0 10.2 9.8 7.1

Para o ano próximo, o governo pretende efetuar uma emissão de

374

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69060 - 29/05

Ajuste de Curvas - python

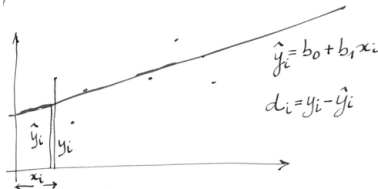
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico $x0y$ conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: no ajuste de curvas busca-se a expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 \cdot x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

. Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 \cdot x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 \cdot x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 \cdot x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i \cdot b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(sx2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar faça: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coefficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5} \times [116.33 - 22.9^2/5]$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625

para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

12 14 17 18 19 21 27 29

associados aos seguintes valores de y=

25 24 23 27 33 34 33 42

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

35 37

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2
-------	-------

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 264 263 243 413 259 394 207 240 545 340

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 6.1 5.2 6.0 6.7 7.0 8.7 9.6 9.1 10.2 5.5

Para o ano próximo, o governo pretende efetuar uma emissão de

548

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69077 - 29/05

Ajuste de Curvas - python

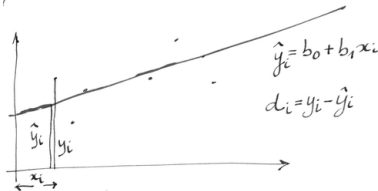
Ao estudar um fenômeno, pode-se propor um modelo matemático. Nele, definem-se

- variável resposta (ou variável dependente) – é aquela que é respondida pelo modelo, ou seja, é a variável de saída do modelo.
- variável explicativa (ou independente) – que é aquela exigida pelo modelo ou seja é a variável de entrada.

Para estudar melhor o fenômeno, deve-se construir um gráfico xOy conhecido como gráfico de dispersão, para verificar visualmente como os dados se distribuem. É bom lembrar que a decisão entre quem é entrada e quem é saída é uma decisão puramente arbitrária. A matemática “lava as mãos” nessa hora. Para ela tanto faz se o cachorro abana o rabo ou ao contrário é o rabo que abana o cachorro.

Este procedimento é parecido com a interpolação, mas lá havia a preocupação de analisar os dados desconhecidos apenas entre os limites obtidos na experimentação (senão seria extrapolação). Aqui não: no ajuste de curvas busca-se uma expressão de uma curva que se estende ao infinito no sentido matemático.

Veja-se um exemplo de um gráfico com alguns pontos:



Cada ponto no gráfico tem as suas coordenadas x e y e buscou-se uma reta ($y = b_0 + b_1 \cdot x$) que busque reproduzir o comportamento da distribuição dos pontos examinados. Cada ponto tem sua ordenada y real (y_i) e a ordenada correspondente $\hat{y}_i = b_0 + b_1 \cdot x_i$.

Para cada ponto i , pode-se calcular a distância entre y_i e $\hat{y}_i$ através da expressão $d_i = y_i - \hat{y}_i$. A soma das distâncias de todos os pontos levantados é

$$D = \sum_{i=1}^n d_i$$

D é a somatória das diferenças entre o valor real de y e aquele dado pela reta $y = b_0 + b_1 \cdot x$. Como D pode ser para mais ou para menos, o somatório simples pode mascarar os erros e não serve. Uma possível solução seria tomar o módulo ($|d_i|$) que resolveria o problema do sinal, mas criaria outro: devido a suas descontinuidades o módulo não tem uma boa derivada.

Resta tomar o quadrado das diferenças (d_i^2), que resolve os 2 problemas. Daqui, vem o nome do método que o dos “mínimos quadrados”.

Então,

$$D = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

. Substituindo d_i^2 por $(y_i - \hat{y}_i)^2$, chamado D de $D(b_0, b_1)$ que é a reta determinada por b_0 e b_1 e substituindo $\hat{y}_i$ por $b_0 + b_1 \cdot x_i$ fica

$$D(b_0, b_1) = \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)^2$$

Derivando esta equação em relação a b_0 fica:

$$\frac{\partial D}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i)$$

Derivando a mesma equação em relação a b_1 fica:

$$\frac{\partial D}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i$$

Os valores de b_0 e b_1 em que a função $D(b_0, b_1)$ apresenta um ponto de mínimo são obtidos – as usual – igualando-se as derivadas a zero. Fica: $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) = 0$ e $-2 \sum_{i=1}^n (y_i - b_0 - b_1 \cdot x_i) \cdot x_i = 0$. Daí, $\sum y_i - \sum b_0 - \sum b_1 \cdot x_i = 0$ e $\sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$. Chamando $\sum$ de S fica: $S y_i - S b_0 - S b_1 x_i = 0$ e $S y_i x_i - S b_0 x_i - S b_1 x_i^2 = 0$. Reescrevendo $S y_i = n b_0 + S x_i b_1$ e $S y_i x_i = S x_i \cdot b_0 + S x_i^2 b_1$ que é um sistema linear. Na sua forma matricial

$$\begin{bmatrix} n & S x_i \\ S x_i & S x_i^2 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} S y_i \\ S y_i x_i \end{bmatrix}$$

O objetivo final é obter b_0 e b_1 que, lembrando são os coeficientes da reta procurada ($\hat{y}_i = b_0 + b_1 x_i$). A solução deste sistema pelo método de escalonamento de Gauss é

$$b_1 = \frac{n \cdot \sum x_i y_i - \sum x_i \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

e

$$b_0 = \frac{\sum y_i - (\sum x_i) \cdot b_1}{n}$$

Obtidos estes parâmetros pode-se escrever a “menor” reta como

$$Y = b_0 + b_1 X$$

Seja um exemplo:

i	1	2	3	4	5
x_i	1.3	3.4	5.1	6.8	8.0
y_i	2.0	5.2	3.8	6.1	5.8

Calculando:

$$n = 5$$

$$\sum x_i = \dots = 24.6$$

$$\sum x_i^2 = (1.3)^2 + (3.4)^2 + \dots = 149.50$$

$$\sum y_i = 22.9$$

$$\sum x_i y_i = (1.3 \times 2.0) + (3.4 \times 5.2) + \dots = 127.54$$

$$b_1 = \frac{5 \times 127.54 - 24.6 \times 22.9}{5 \times 149.50 - 24.6^2} = 0.522$$

e

$$b_0 = \frac{22.9 - (24.6 \times 0.522)}{5} = 2.01$$

e daqui

$$Y = 2.01 + 0.522X$$

Python

```
# minquad
def minquad(x,y):
    import numpy as np
    n=len(x)
    sy=0
    sx=0
    sx2=0
    sxy=0
    sy2=0
    i=0
    while i<n:
        sx=sx+x[i]
        sy=sy+y[i]
        sx2=sx2+(x[i]**2)
        sxy=sxy+(x[i]*y[i])
        sy2=sy2+y[i]**2
        i=i+1
    print('sy ',sy)
    print('sx ',sx)
    print('sx2 ',sx2)
    print('sxy ',sxy)
    print('n ',n)
    print('sy2 ',sy2)
    b1=((sy*sx)-(n*sxy))/((sx**2)-(sx2*n))
    b0=(sy-(sx*b1))/n
    num=(sxy-((sx*sy)/n))*2
    den=((sx2-((sx**2)/n))*(sy2-((sy**2)/n)))
    print("coef = ",num/den)

    while (1==1):
        xx=float(input("Informe x (-1 para fim)"))
        if xx==-1:
            return
        yy=b0+(b1*xx)
        print("y = ",yy)
#para chamar faça: x=[1.3,3.4,5.1,6.8,8.0] e
# y=[2.0, 5.2, 3.8, 6.1, 5.8]
#depois: minquad(x,y) e informe quantos x_i
# quiser, seguido de -1
```

Coefficiente de determinação

mede a qualidade do ajuste linear, isto é quão próxima é a reta determinada em relação ao conjunto de pontos. É um valor (R^2) que varia entre zero e um ($0 \leq R^2 \leq 1$). Quanto mais perto de 1, melhor o ajuste. A fórmula é

$$R^2 = \frac{[\sum xy - ((\sum x \sum y)/n)]^2}{[\sum x^2 - ((\sum x)^2/n)] \cdot [\sum y^2 - ((\sum y)^2/n)]}$$

no exemplo acima, o número que faltou calcular é $\sum y^2 = 116.33$ e fica:

$$R^2 = \frac{[127.54 - 24.6 \times 22.9/5]^2}{149.5 - 24.6^2/5 \times [116.33 - 22.9^2/5]}$$

ou $R^2 = 0.679$.

Para você fazer

Para testar o algoritmo implementado, faça:

x = 26 31 33 36 38 41 42 44
 y = 40 41 45 49 51 55 56 58
 a = 9.249642005 e b = 1.103102625
 para x = 45, y = 58.88 e para x = 50, y = 64.40

Para você fazer

1. No exercício a seguir, considere valores de x=

11 14 18 24 25 27 32 35

associados aos seguintes valores de y=

-17 -20 -29 -40 -48 -54 -64 -63

Calcule os parâmetros b_0 e b_1 , e depois infira os valores de y para dois valores de x dados (x_1 e x_2)

36 37

Para calcular a resposta, arredonde para 2 casas decimais

y_1	y_2

2. Em um país distante chamado LISARB, há muitos problemas econômicos. Em particular, há uma inflação persistente que APARENTEMENTE depende entre outras coisas, da emissão de moeda pelo Banco Central daquele país. Economistas que apoiam esta visão levantaram a inflação e a emissão de moeda nos últimos 10 anos, obtendo as duas séries históricas, a seguir:

EMISSAO ANUAL DE MOEDA NOS ULTIMOS 10 ANOS (em bilhões de dólares)
 255 245 226 368 337 601 294 474 497 610

INFLAÇÃO ANUAL NOS ULTIMOS 10 ANOS
 5.2 5.7 7.2 8.0 7.7 6.0 7.5 5.2 6.4 7.1

Para o ano próximo, o governo pretende efetuar uma emissão de

316

bilhões de dólares. Segundo este modelo econométrico, baseado no método dos mínimos quadrados, qual é a inflação esperada ?

inflação



121-69796 - 29/05