MÉTODOS QUANTITATIVOS

AULA 1

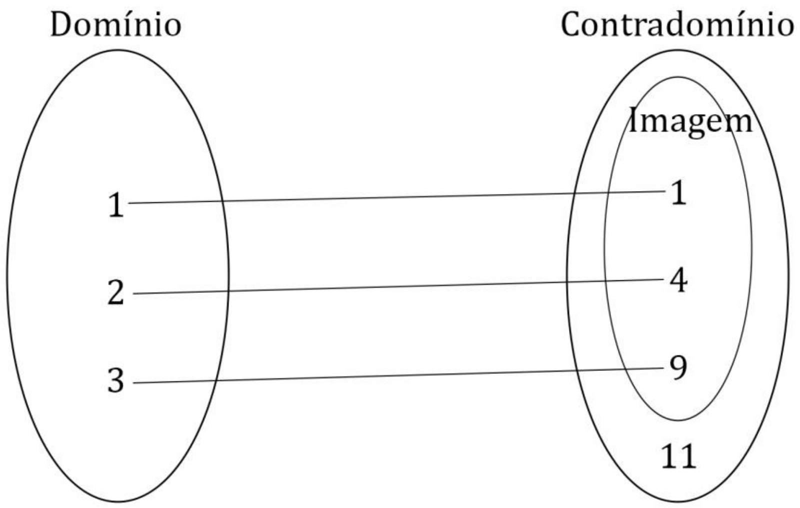
Prof. Ricardo Alexandre Deckmann Zanardini

CONVERSA INICIAL

Uma função, do ponto de vista matemático, é uma relação que há entre quantidades em que cada elemento de um conjunto está relacionado a apenas um elemento do outro conjunto. Na prática, há muitos exemplos importantes que podem ser representados por funções. Nesta etapa, abordaremos os principais tópicos associados às funções.

TEMA 1 – FUNÇÕES

Dizemos que uma função y=f(x) é uma relação que associa cada elemento x de um conjunto chamado de *domínio* um único elemento y de outro conjunto chamado de *contradomínio*. Chamamos de conjunto-imagem o conjunto formado por todos os elementos do contradomínio que estão associados aos elementos do domínio.



O usual é utilizar x e y para representar matematicamente funções de uma variável, mas pode-se utilizar outras letras também. Uma função está associada a uma expressão que relaciona a variável dependente y com a variável independente x.

Exemplo: considere uma indústria que produz sanduicheiras elétricas. Supondo que o custo unitário de produção corresponde a R$ 32,00 e que mensalmente x unidades são produzidas, expresse o custo mensal C em função da quantidade de sanduicheiras produzidas.

Resolução: como o custo unitário corresponde a R$ 32,00 e a quantidade mensal produzida é x, o custo mensal associado à produção das sanduicheiras elétricas é dado pela multiplicação do custo unitário (32 reais) pela quantidade produzida (x unidades). Sendo assim, a função que relaciona o custo mensal com a quantidade de sanduicheiras é:

$$C(x)=32x.$$

Podemos representar graficamente essa função.

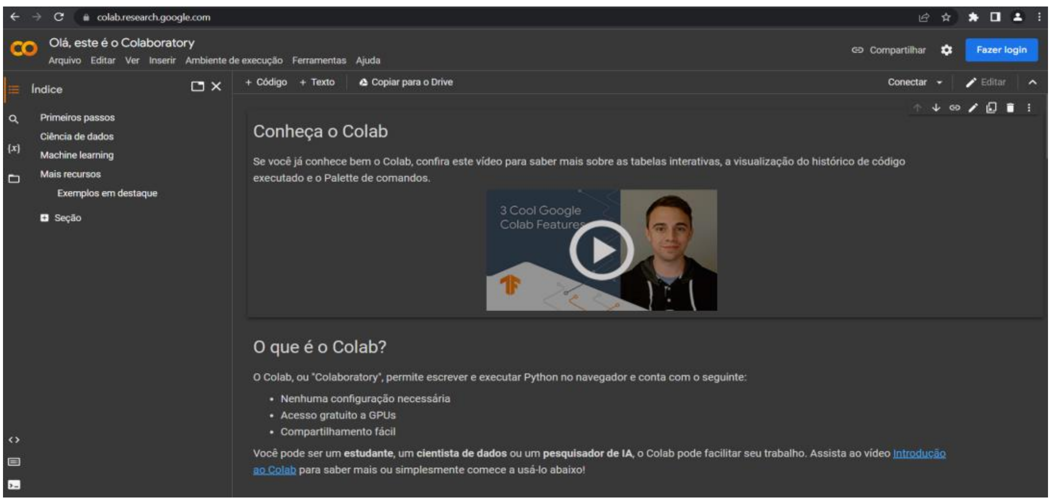
A vantagem de utilizar gráficos para representar funções é que a compreensão do comportamento destas é maior quando podemos visualizar a função além de sua expressão matemática.

Há diversos aplicativos que possibilitam a construção de gráficos.

Nesta nossa caminhada sobre métodos quantitativos, utilizaremos a biblioteca *matplotlib* do Python para a construção de gráficos de funções de uma variável. Por mais que o Python seja uma linguagem de programação, não precisamos de conhecimentos de programação, pois iremos utilizar de uma maneira bem simples funções já desenvolvidas.

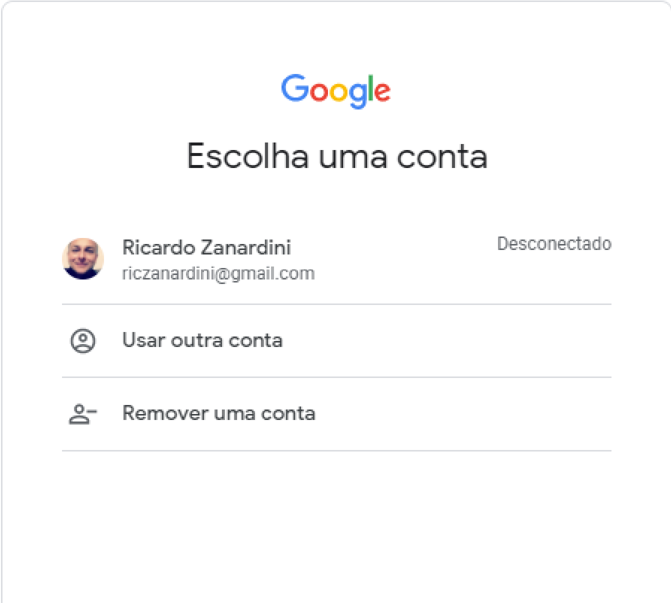
Também não será necessário instalar qualquer tipo de software, pois utilizaremos um ambiente colaborativo on-line denominado de Google Colaboratory e mais conhecido como Google Colab, pois possibilita o uso das funcionalidades do Python diretamente no navegador.

Para utilizar o Google Colab, é preciso ter uma conta Google e acessar o endereço <colab.research.google.com> (acesso em: 11 out. 2022).



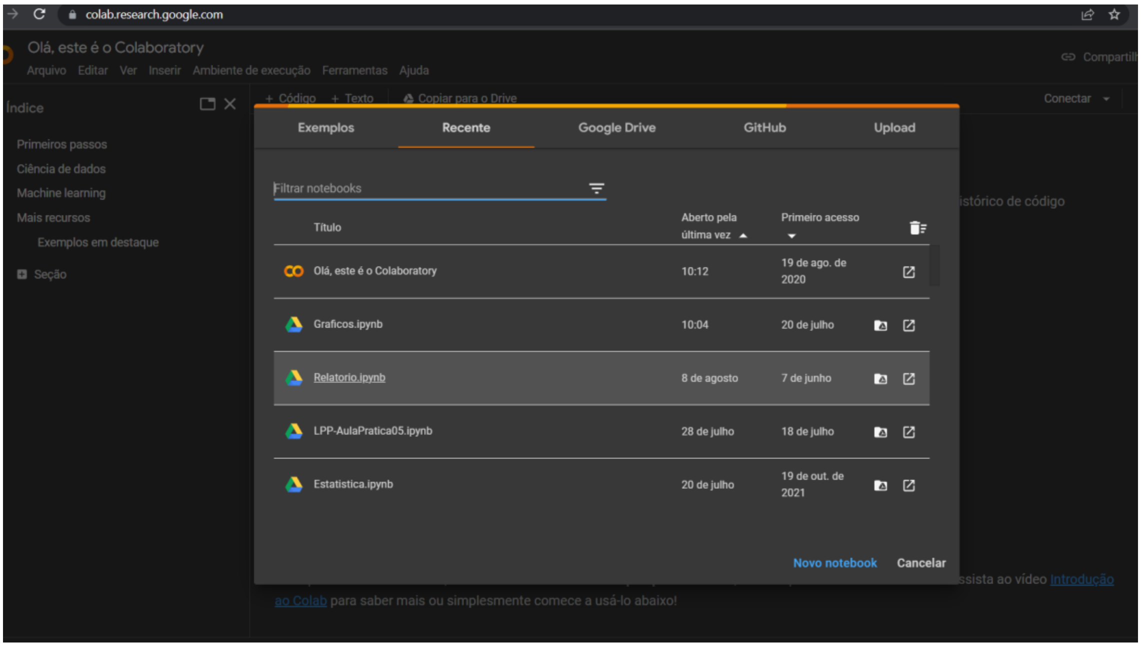
Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

O primeiro passo consiste em efetuar o login utilizando a conta Google. Caso ainda não tenha, é preciso criar uma nova conta.



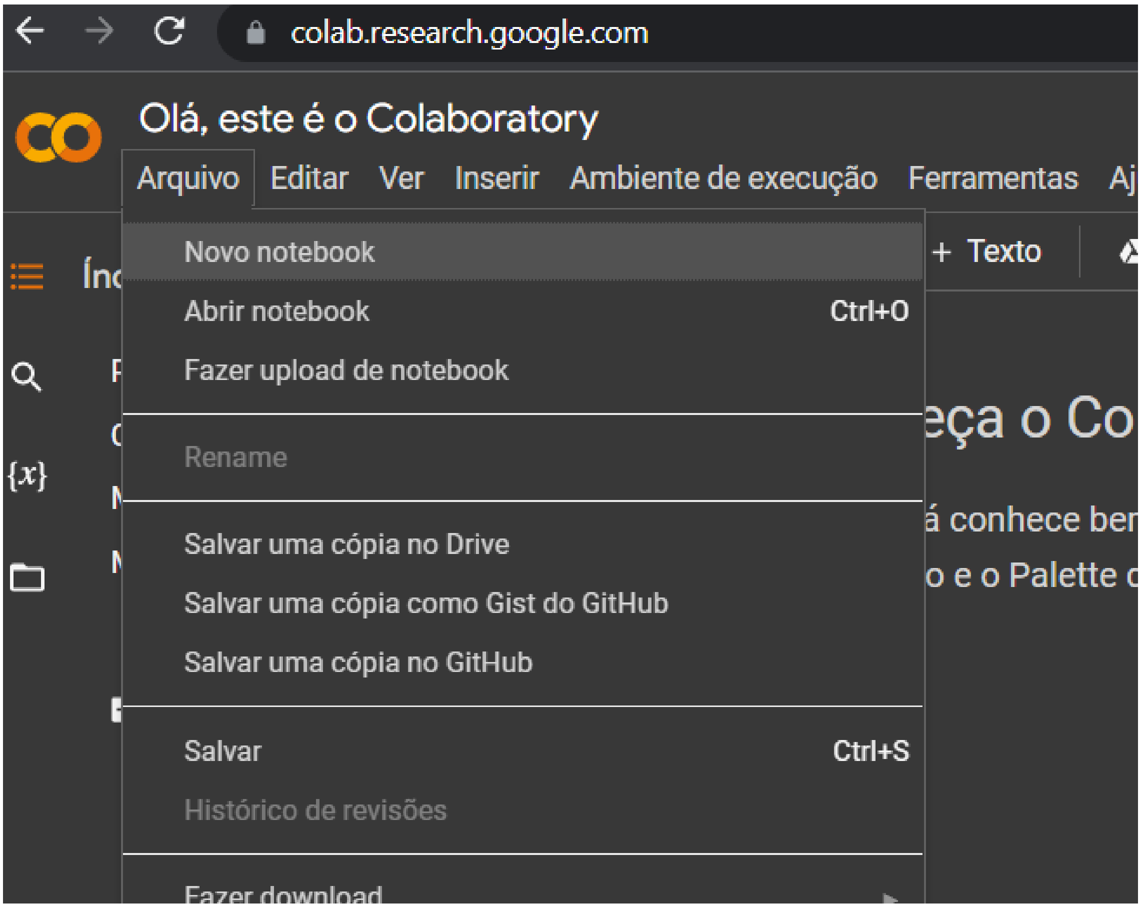
Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

Na página inicial, são apresentadas algumas informações relacionadas ao Google Colab.



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

Para começar a trabalhar com a biblioteca matplotlib, primeiro vamos clicar em Arquivo e, em seguida, em Novo notebook.



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

Os comandos necessários serão inseridos em uma célula de código. Para inserir uma nova célula, é só clicar em “+ Código”.

Para a criação de gráficos, a sequência de comandos segue sempre a mesma estrutura.

O primeiro passo é importar a biblioteca matplotlib e chamá-la de plt para simplificar a escrita dos comandos. Para isso, basta digitar em uma célula de código:

```
import matplotlib.pyplot as plt
```

Também é necessário importar a biblioteca numpy para gerar valores de x necessários para a construção dos gráficos. Chamaremos essa biblioteca de np e, em uma célula de código, vamos digitar:

```
import numpy as np
```

Agora que importamos as duas bibliotecas, podemos pensar nos passos necessários para a construção do gráfico.

Precisamos gerar valores para x e, em seguida, obter os respectivos valores de y.

O comando linspace gera um determinado número de valores igualmente espaçados para um intervalo de x. Por exemplo, se queremos 100 valores para x, variando de 0 a 10, basta digitar:

```
x=np.linspace(0, 10, 100)
```

Agora que temos os valores de x, vamos obter os respectivos valores de y tais que y=32x. Assim, basta digitar:

```
y=32*x
```

em que * indica a multiplicação.

O comando plot gera o gráfico:

```
plt.plot(x,y)
```

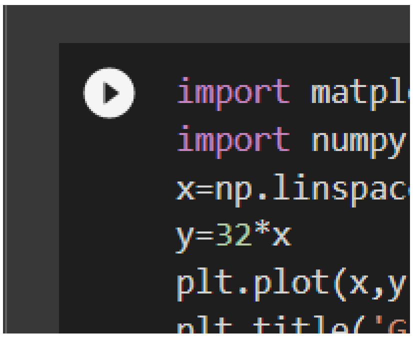
Para adicionar um título ao gráfico, utiliza-se o comando title:

```
plt.title('Gráfico da função C(x)=32x')
```

Finalmente, basta utilizar o comando show para que o gráfico possa ser apresentado corretamente.

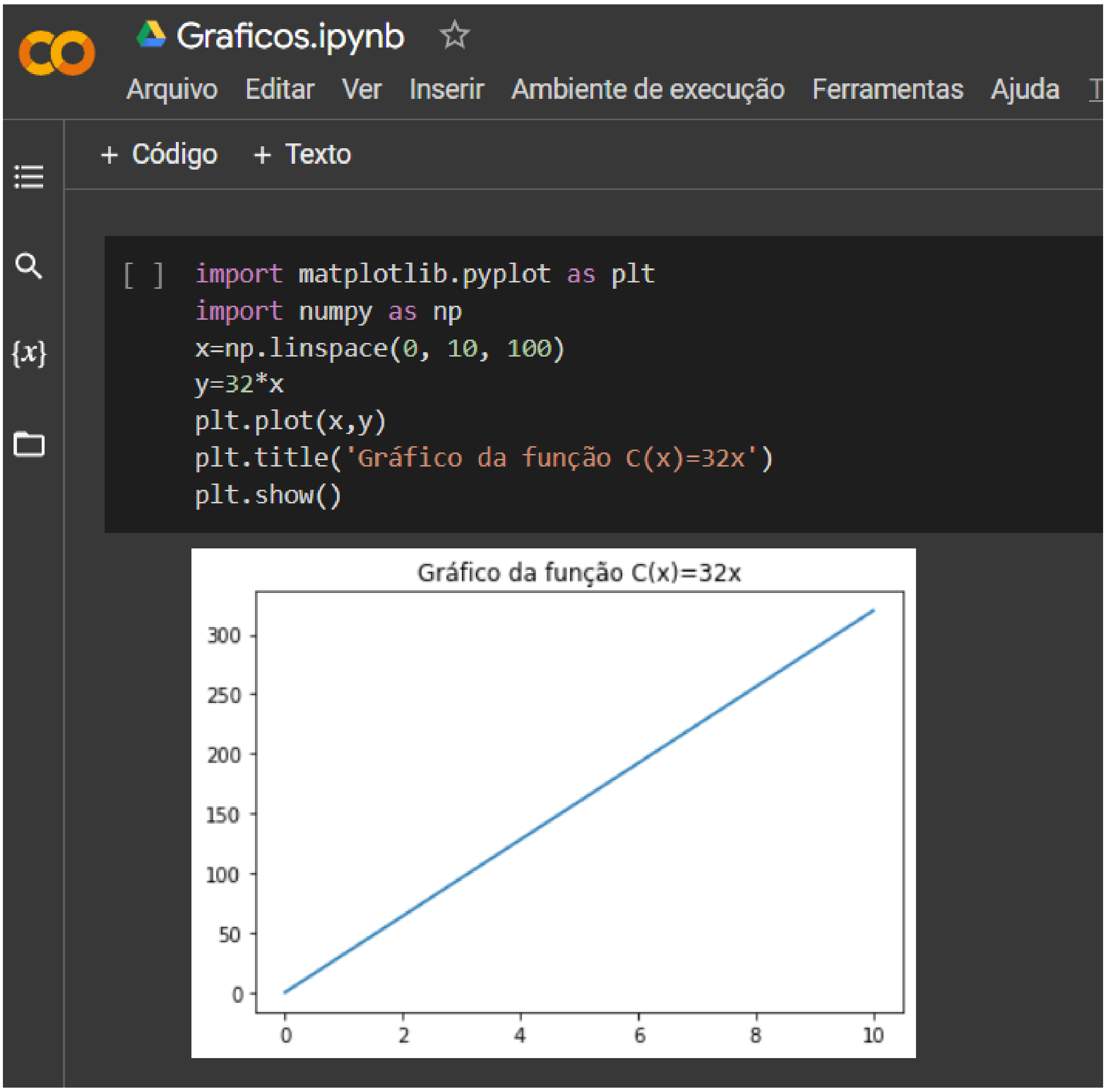
```
plt.show()
```

Para executar os comandos e obter o gráfico desejado, precisa-se apertar as teclas CTRL e ENTER ou então clicar no ícone com uma seta, ao lado da célula de comando:



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

A sequência total dos comandos com o respectivo gráfico é:



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

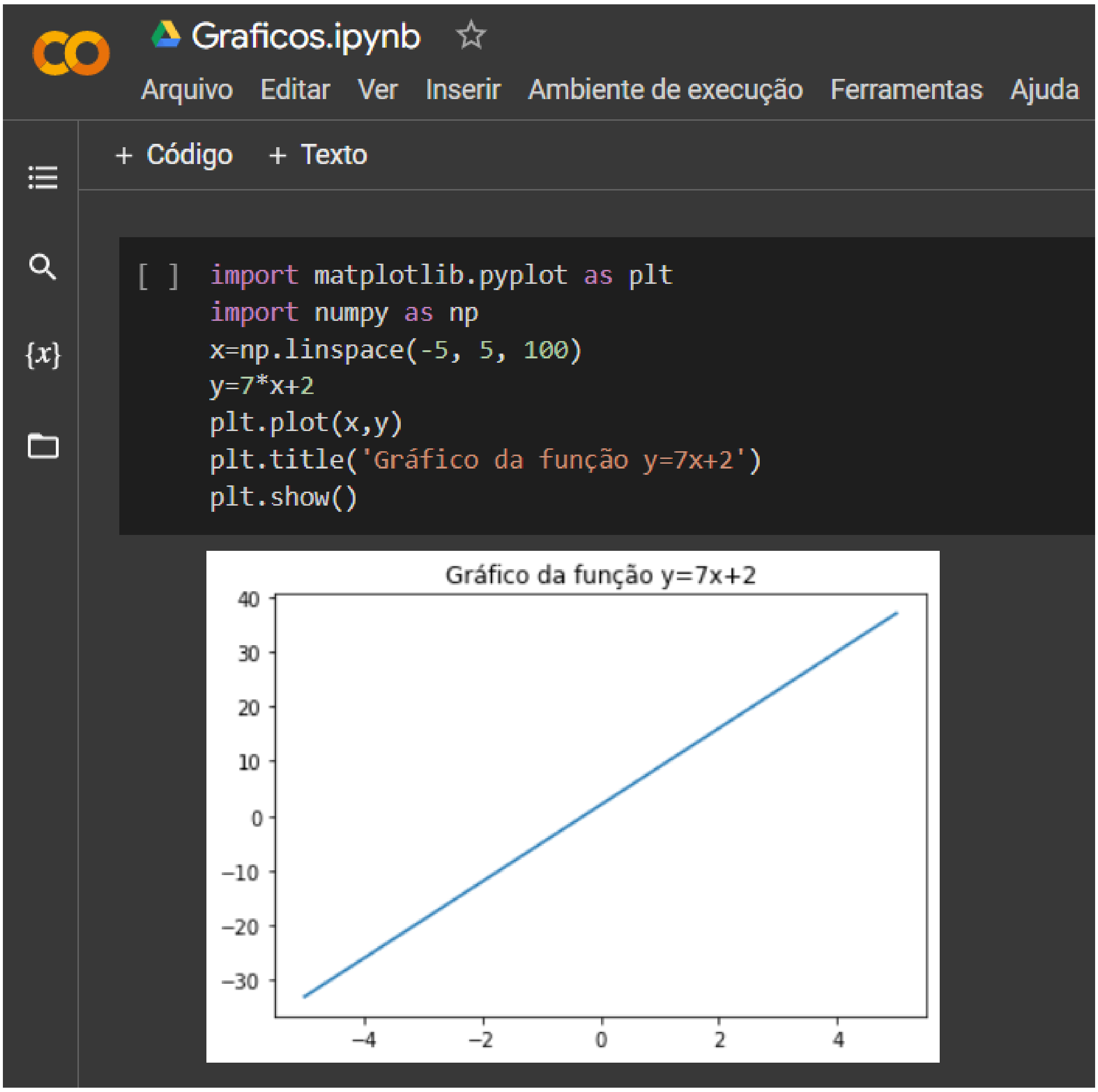
De forma geral, a sequência de comando para a construção de um gráfico por meio da biblioteca matplotlib do Python é:

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(XXXX, XXXX, XXXX)
y=XXXX
plt.plot(x,y)
plt.title('XXXX')
plt.show()
```

em que precisamos apenas substituir os termos XXXX pelos que forem necessários.

Por exemplo, se queremos o gráfico da função y=7x+2 em um intervalo no qual x varia de -5 a 5, tem-se:

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(-5, 5, 100)
y=7*x+2
plt.plot(x,y)
plt.title('Gráfico da função y=7x+2')
plt.show()
```



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

O gráfico da função y=7x+2 é uma linha reta e o da função C(x)=32x também. Por esse motivo, essas funções são conhecidas como lineares.

TEMA 2 – FUNÇÕES LINEARES

Uma função é chamada de linear quando tem a forma

$$y=ax+b$$

em que "a" corresponde ao coeficiente angular e indica a taxa de crescimento ou de decrescimento da função e o termo "b" é o coeficiente linear e indica o ponto no qual o gráfico da função intercepta o eixo y.

Também podemos utilizar a notação

$$f(x)=ax+b$$

em que f indica a função, e o termo entre parênteses indica a variável. Assim, f(x) indica a função f em que a respectiva variável independente é representada por x.

Há muitos exemplos reais que podem ser associados a funções lineares.

Quando é colocado combustível em um tanque de um veículo, tem-se uma função linear que associa o total a ser pago com a quantidade de combustível colocada. Supondo que o preço do litro da gasolina custa R$ 5,70 e que x indica a quantidade de combustível, a função que relaciona o custo total C em função da quantidade de combustível é

$$C(x)=5{,}7x.$$

Sendo assim, se um veículo for abastecido com 10 litros de gasolina, basta substituir x por 10 e multiplicar 5,7 por 10 para saber qual será o respectivo custo:

$$C(x)=5{,}7x$$
$$C(10)=5{,}7(10)$$
$$C(10)=57$$

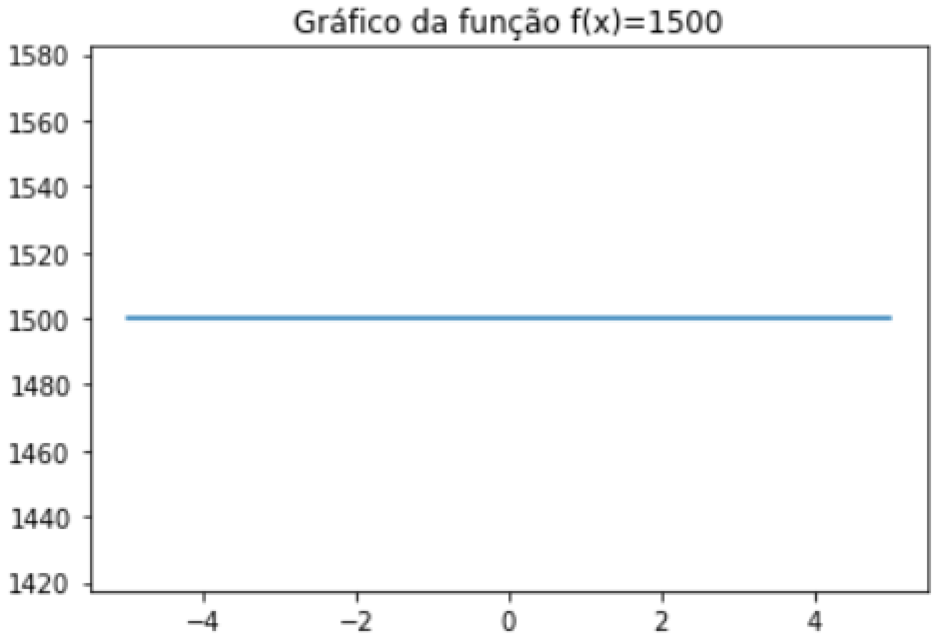
Portanto, o custo corresponde a R$ 57,00.

Quando tratamos de funções lineares, podemos ter funções constantes, ou seja, funções cujo valor não se altera com a variação de x.

Por exemplo, se considerar uma função que relaciona o aluguel de um imóvel com o tempo e não há alteração no valor do aluguel, tem-se uma função constante. Supondo que o aluguel corresponde a R$ 1.500,00, a respectiva função é dada por

$$f(x)=1500.$$

Note que, nesse caso, a=0 e b=1500 e o respectivo gráfico corresponde a



Fonte: Ricardo Alexandre Deckmann Zanardini.

Para a construção do gráfico, os comandos utilizados foram:

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(-5, 5, 100)
y=np.linspace(1500, 1500, 100)
plt.plot(x,y)
plt.title('Gráfico da função f(x)=1500')
plt.show()
```

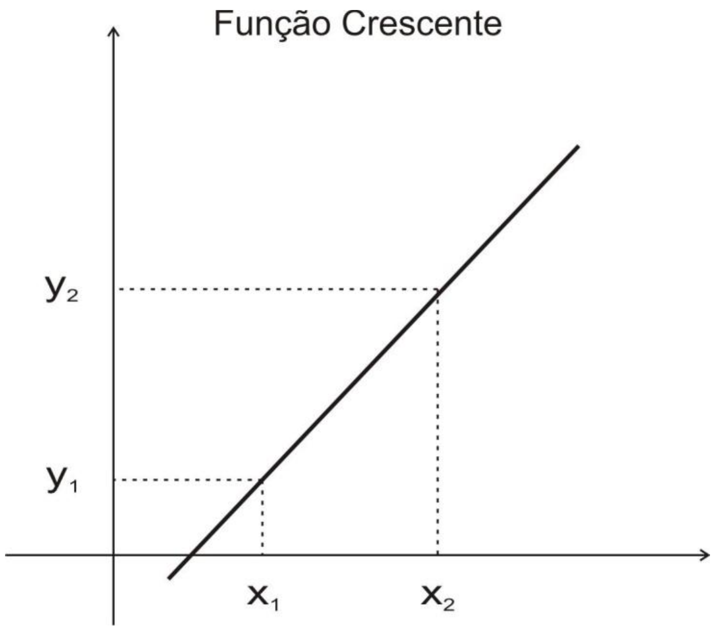
Uma função linear é chamada de crescente quando sempre que o valor de x aumenta, o de y também aumenta. De forma geral, uma função é dita crescente quando

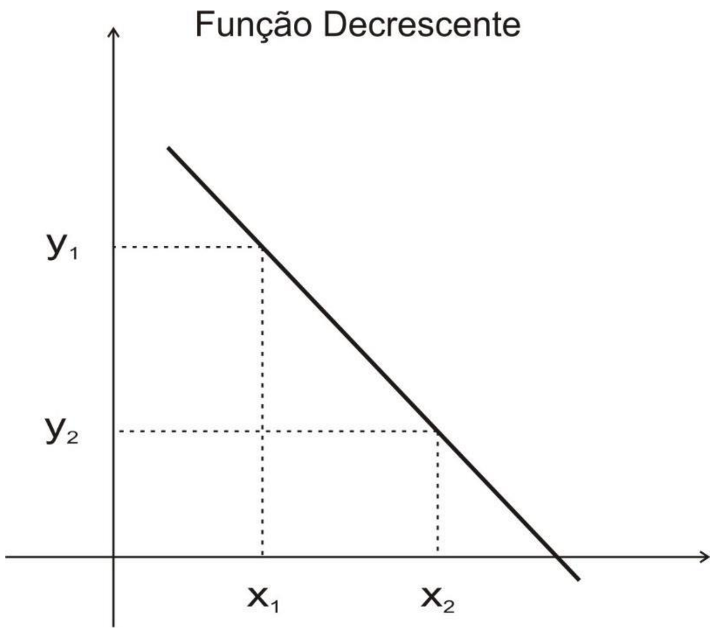
$$x_2>x_1 \Rightarrow y_2>y_1.$$

Quando aumentamos o valor de x, o valor de y diminui, a função é dita decrescente. Quando uma função é decrescente, tem-se que

$$x_2>x_1 \Rightarrow y_2<y_1.$$

Graficamente, tem-se:





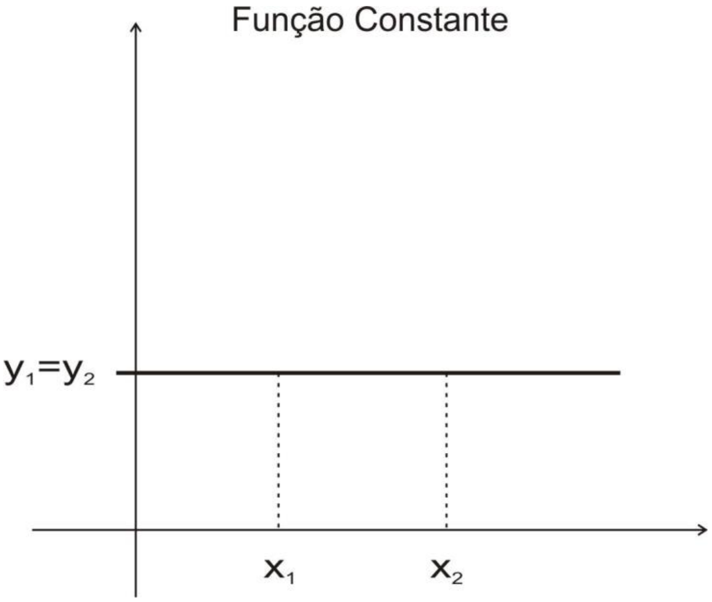
Fonte: Ricardo Alexandre Deckmann Zanardini.

O crescimento ou decrescimento de uma função linear está associado ao coeficiente angular “a”. Se o valor de “a” é positivo, a função é crescente e se “a” é negativo, ela é decrescente:

a>0: função crescente

a<0: função decrescente.

É importante ressaltar que quando a=0, tem-se uma função constante.



Fonte: Ricardo Alexandre Deckmann Zanardini.

Exemplo: considere uma indústria onde a produção de automóveis foi de 32 veículos por dia no primeiro mês e de 27 veículos por dia no 6° mês de um determinado ano. Sendo assim, obtenha a equação da reta que relaciona a produção diária com os respectivos meses. Em seguida, apresente a taxa de crescimento ou de decrescimento da produção nesse período.

Resolução: para determinar a equação da reta associada ao problema, é preciso de dois pontos. O primeiro ponto é obtido a partir da informação de que no primeiro mês a produção diária foi de 32 veículos. Assim, tem-se o par ordenado A(1, 32). Como o par ordenado é da forma (x, y), o primeiro valor corresponde ao mês (x=1) e o segundo valor corresponde à produção diária (y=32). Para obter o segundo ponto, tem-se a informação de que no sexto mês a produção diária foi de 27 veículos, ou seja, x=6 e y=27, o que resulta em B(6, 27).

Agora que se tem os pontos A(1, 32) e B(6, 27), precisa-se substituir as coordenadas de cada um deles na equação reduzida da reta y=ax+b.

Considerando o ponto A(1, 32), tem-se:

$y=ax+b$

$32=a(1)+b$

32=a+b

a+b=32

Para o ponto B(6, 27), tem-se:

y=ax+b

27=a(6)+b

27=6a+b

6a+b=27

A partir das equações obtidas, basta resolver o sistema

$$a+b=32$$
$$6a+b=27$$

Para resolver o sistema, iremos utilizar o método da adição.

Como os coeficientes de b são iguais a 1 em cada uma das equações, basta multiplicar uma das equações por -1 e, em seguida, somar as duas equações.

Multiplicando a primeira equação por -1, tem-se:

$$-a-b=-32$$
$$6a+b=27$$

Somando os termos das duas equações, tem-se:

$$-a+6a=5a,\ -b+b=0 \text{ e } -32+27=-5$$

Logo,

$$5a=-5$$
$$a=-5/5$$
$$a=-1$$

Para obter o valor de b, basta substituir a=-1 na primeira equação:

$$a+b=32$$
$$-1+b=32$$
$$b=32+1$$
$$b=33$$

Logo, a função linear que associa a produção diária com os meses é dada por:

$$f(x)=-1x+33.$$

Como o coeficiente de x é igual a -1, tem-se que a função é decrescente e a taxa corresponde a -1, ou seja, a cada mês a produção diária diminuiu em uma unidade.

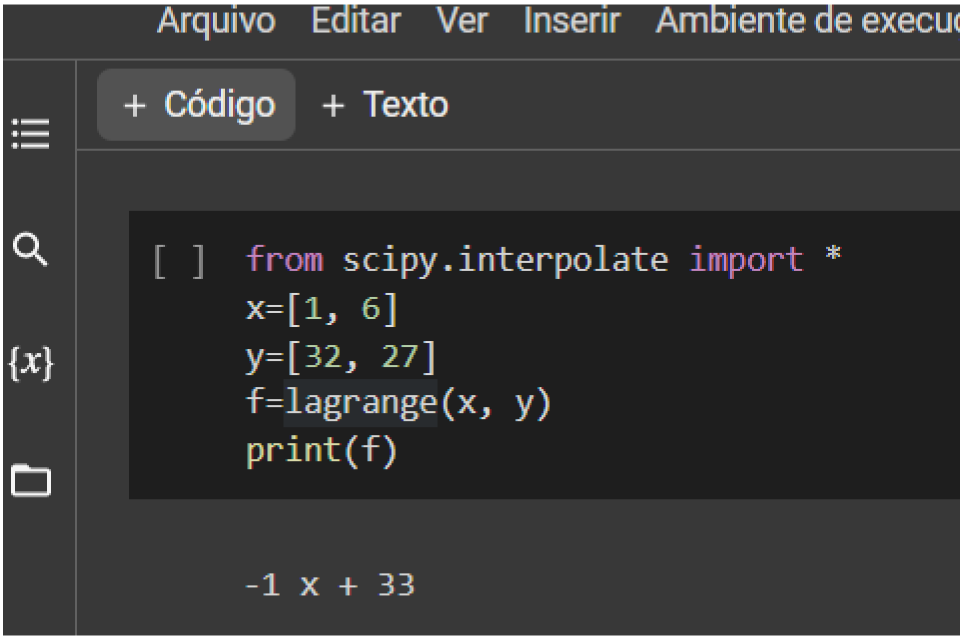
Utilizamos uma forma analítica para obter a função linear que passa por dois pontos, mas também é possível obtê-la por meio do Python.

O procedimento é bem simples, pois iremos utilizar o comando lagrange que gera a função que interpola um conjunto de pontos, ou seja, a função que passa pelos pontos dados.

Para esse nosso exemplo, basta importar a biblioteca scipy.interpolate, definir um vetor contendo os valores associados ao meses, um vetor contendo as respectivas produções diárias, utilizar o comando lagrange para gerar a função e, em seguida, utilizar o comando print para apresentar a função

```
from scipy.interpolate import *
x=[1, 6]
y=[32, 27]
f=lagrange(x,y)
print(f)
```

O resultado é:



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

Assim, a função gerada pelo Python é f(x)=-1x+33, que equivale a f(x)=-x+33.

Note que, nesse caso, não escrevemos diretamente os pontos A(1, 32) e B(6, 27), mas escrevemos um vetor x contendo as primeiras coordenadas de cada ponto, ou seja, 1 e 6 e um vetor y contendo as últimas coordenadas de cada ponto, ou seja, 32 e 27.

Além das funções lineares, as funções quadráticas são muito importantes na resolução de diversos problemas reais.

TEMA 3 – FUNÇÕES QUADRÁTICAS

Muitas situações do nosso cotidiano podem ser descritas por meio de funções do segundo grau, também conhecidas como funções quadráticas.

Uma função da forma

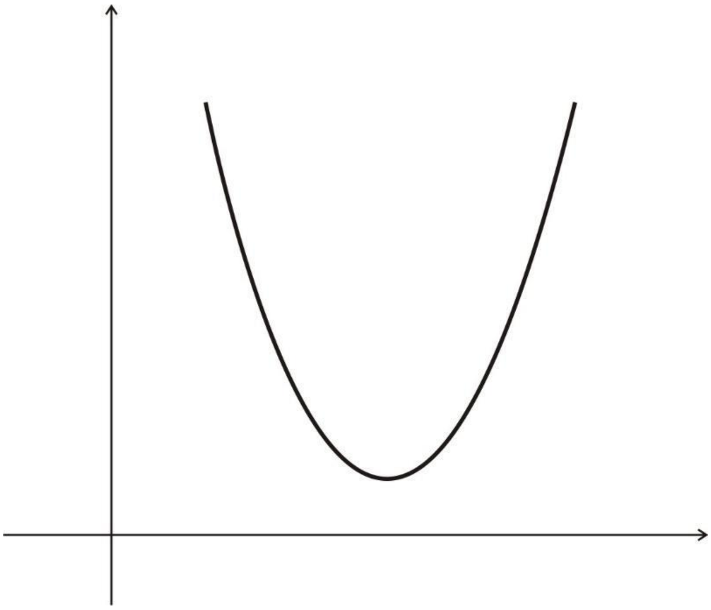
$$y=ax^2+bx+c$$

em que “a”, “b” e “c” são constantes e “a” é diferente de zero, é conhecida como função quadrática.

O gráfico de uma função quadrática é uma parábola.

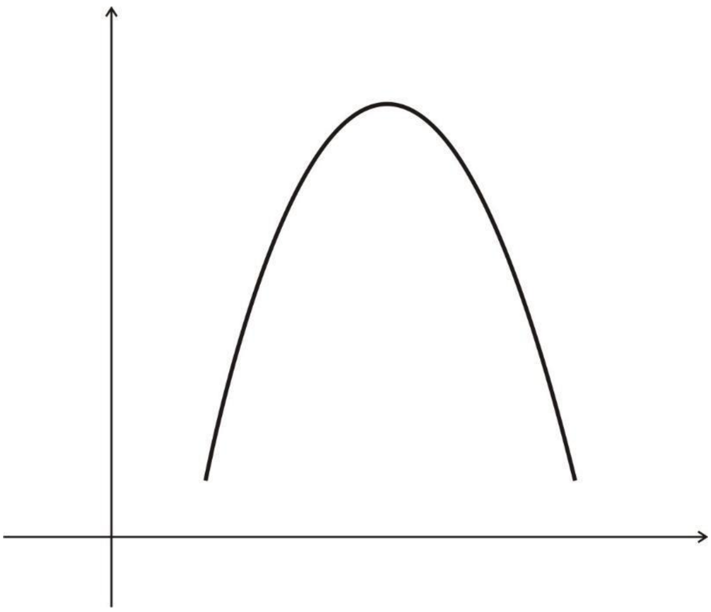
A concavidade desta pode estar para cima ou para baixo e isso depende do sinal do coeficiente de x^2, ou seja, depende do sinal de “a”.

Se a>0, a parábola tem concavidade voltada para cima.



Fonte: Ricardo Alexandre Deckmann Zanardini.

Se a<0, a parábola tem concavidade voltada para baixo.



Fonte: Ricardo Alexandre Deckmann Zanardini.

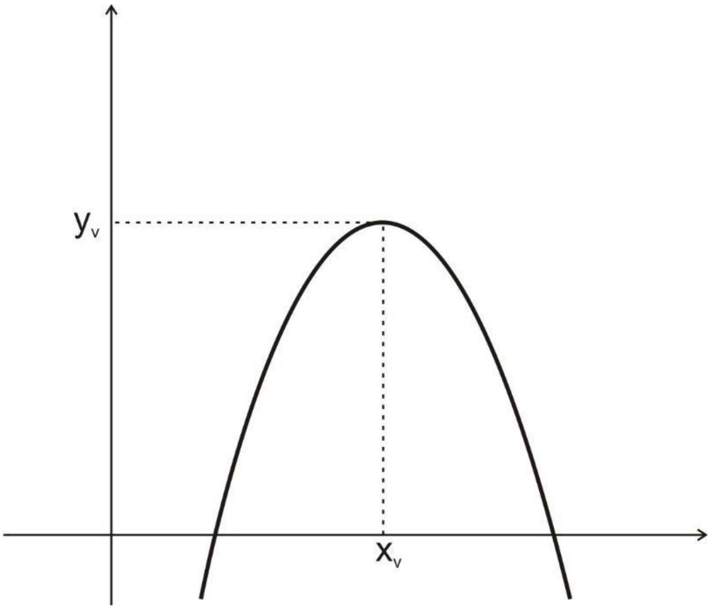
Há muitas situações reais nas quais as funções quadráticas estão presentes. Uma função quadrática pode relacionar a variação do lucro referente à venda de uma mercadoria com a variação do respectivo preço praticado.

Quando o gráfico de uma função quadrática tem concavidade para baixo, a função tem um ponto de máximo. Quando o gráfico de uma função quadrática tem a concavidade voltada para cima, a função tem um ponto de mínimo. Para obter o valor de x associado ao ponto de máximo ou ponto de mínimo da função, tem-se a fórmula do x referente ao vértice:

$$x_v = \frac{-b}{2a}$$

O respectivo valor de y referente ao vértice é dado por:

$$y_v = \frac{-\Delta}{4a} \text{ onde } \Delta = b^2 - 4ac$$



Fonte: Ricardo Alexandre Deckmann Zanardini.

Exemplo: considere um estacionamento no qual a função quadrática que relaciona o lucro mensal com o preço cobrado por hora é:

$$L(x)=-400x^2+6800x-12000$$

Nessa função, x é o preço cobrado por hora e L é o lucro mensal. A partir dessas informações, determine o preço que maximiza o lucro e o respectivo lucro máximo.

O preço que maximiza o lucro é o valor de x que corresponde ao vértice. Assim,

$$x_v = \frac{-b}{2a}$$

precisamos dos coeficientes a e b da função:

a = -400

b = 6800

O próximo passo é substituir os coeficientes na fórmula.

$$x_v = \frac{-6800}{2(-400)}$$

$$x_v = \frac{-6800}{-800}$$

$$x_v = 8{,}50$$

Podemos concluir que, para esse estacionamento, o preço que maximiza o lucro é R$ 8,50 por hora.

Para determinar o lucro máximo, é possível utilizar a fórmula referente ao y_v ou também substituir o valor do x encontrado no item anterior na função L(x)=-400x^2+6800x-12000.

Substituindo x por 8,5 (8,5 equivale a 8,50) na função quadrática L(x)=-400x^2+6800x-12000, tem-se:

L(x)=-400x^2+6800x-12000

L(8,5)=-400(8,5)^2+6800(8,5)-12000

L(8,5)=-400(72,25)+6800(8,5)-12000

L(8,5)=-28900+57800-12000

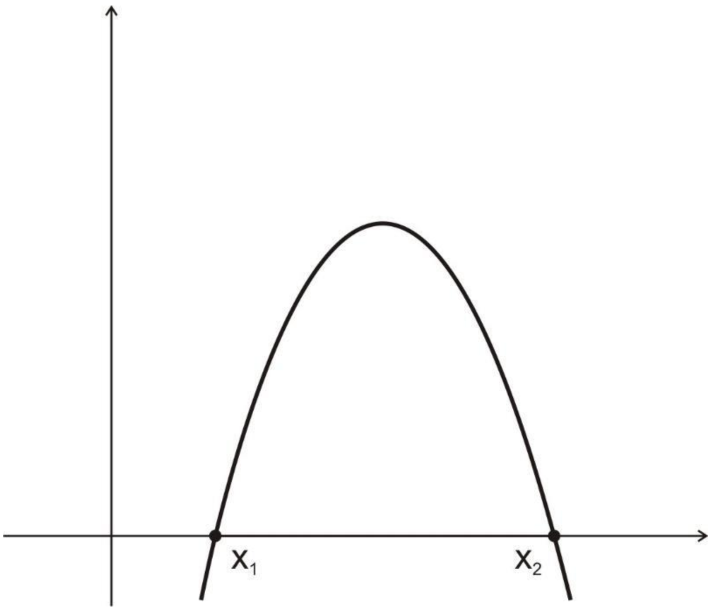
L(8,5)= 16.900,00

Portanto, o lucro mensal corresponde a R$ 16.900,00.

No caso das funções quadráticas, podemos determinar as raízes utilizando a fórmula quadrática:

$$x = \frac{-b \pm \sqrt{\Delta}}{2.a} \text{ onde } \Delta = b^2 - 4ac.$$

As raízes de uma função quadrática correspondem aos valores de x tais que y seja igual a zero.



Fonte: Ricardo Alexandre Deckmann Zanardini.

Podemos utilizar o Python para representar graficamente as funções quadráticas.

A sequência de passos é a mesma vista em funções lineares:

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(XXXX, XXXX, XXXX)
y=XXXX
plt.plot(x,y)
plt.title('XXXX')
plt.show()
```

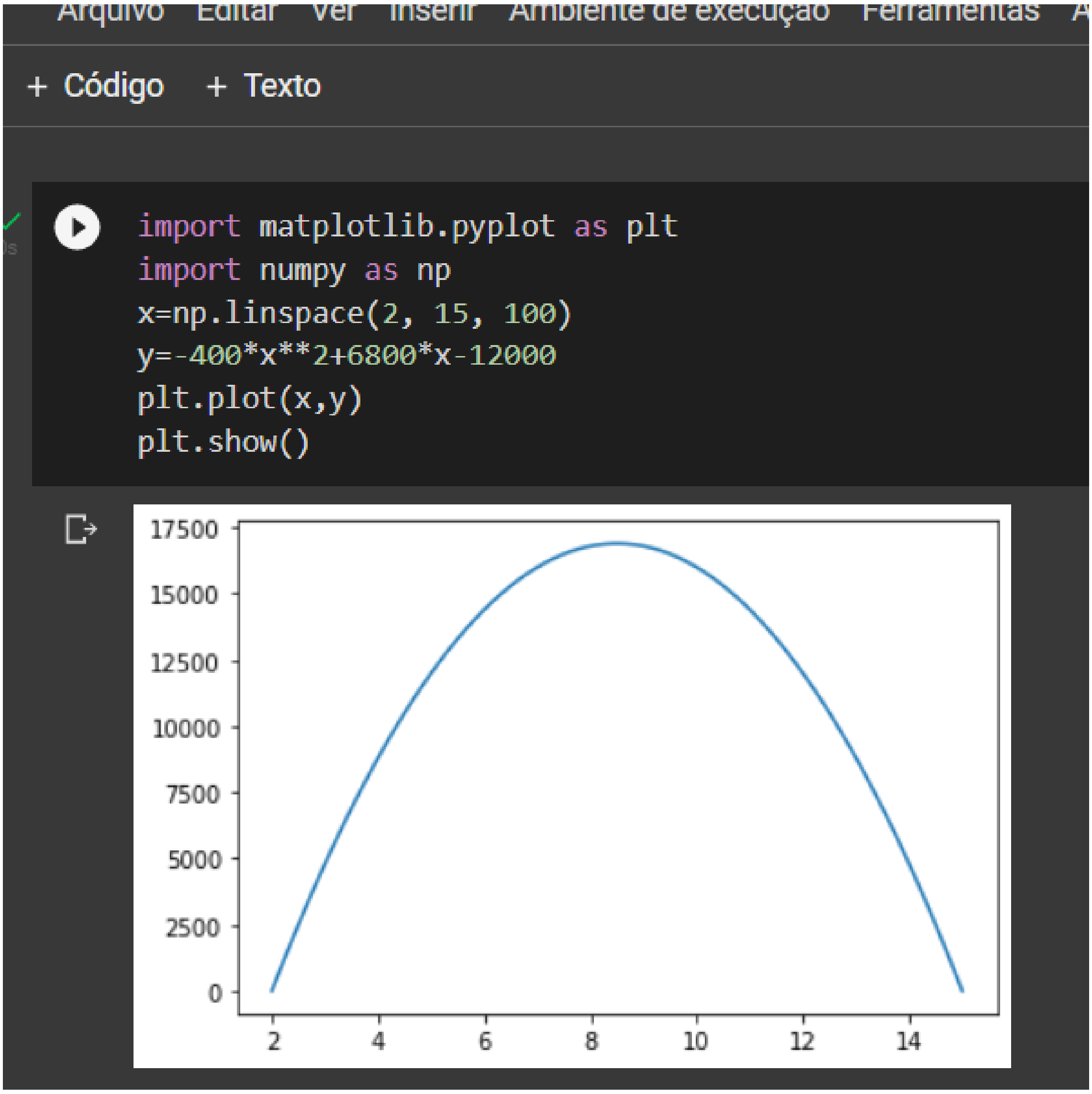
Considerando a função

$$L(x)=-400x^2+6800x-12000$$

utilizaremos a seguinte sequência de comandos:

```
import matplotlib.pyplot as plt
```

```
import numpy as np
x=np.linspace(2, 15, 100)
y=-400*x**2+6800*x-12000
plt.plot(x,y)
plt.show()
```



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

É importante ressaltar que o símbolo "**" indica a potenciação. Escolhemos o intervalo de x entre 2 e 15 e 100 valores nesse intervalo.

Em outro momento, aprenderemos a forma de obter uma função quadrática a partir de um problema real.

Vamos agora tratar de um outro tipo importante de funções, as exponenciais.

TEMA 4 – FUNÇÕES EXPONENCIAIS

Uma função que tem a forma

$$f(x)=a.b^x$$

em que “a” é diferente de zero, “b” é positivo e “b” é diferente de 1 é chamada de função exponencial. O termo “a” é o valor da função quando x é igual a zero. O termo “b” é chamado de base. Quando “b” é maior do que 1, a função é crescente; e se “b” está entre 0 e 1, a função é decrescente.

Utilizamos funções exponenciais em muitas aplicações reais, como crescimento populacional, problemas envolvendo oferta e demanda, depreciação, meia-vida de uma substância, juros compostos, entre muitos outros.

Para que a compreensão seja mais efetiva, vamos considerar o seguinte exemplo: uma determinada indústria tem um crescimento previsto de 8% ao ano. Atualmente, o valor dela corresponde a 120 milhões de reais. Construa uma tabela contendo a previsão do valor da empresa para cada um dos próximos cinco anos.

A cada ano, o valor da empresa é 8% maior do que o valor anterior. Assim, devemos multiplicar o valor atualizado da empresa por 1,08, o que corresponde a 100% mais o acréscimo de 8%, pois 100%+8%=108%=1,08. Esse fator de crescimento é a base “b” da função exponencial.

Os valores anuais previstos da empresa são:

Ano	Valor (em milhões de reais)
0	120
1	120 x 1,08 = 129,6
2	120 x $1,08^2$ = 139,968
3	120 x $1,08^3$ = 151,16544
4	120 x $1,08^4$ = 163,2586752
5	120 x $1,08^5$ = 176,319369216

Fonte: Ricardo Alexandre Deckmann Zanardini.

De forma geral, podemos escrever essa função para um ano qualquer da seguinte forma:

$$f(x)=120.1{,}08^x$$

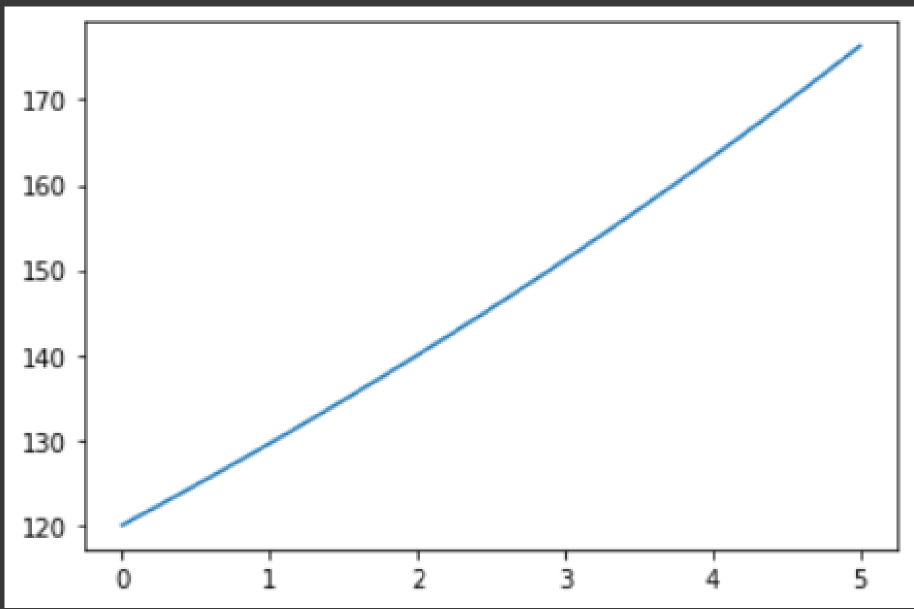
em que a=120 e b=1,08. A base “b” corresponde ao fator de aumento 1,08 e o valor de “a” corresponde ao valor da empresa quando x=0, ou seja, a=120.

Podemos utilizar o Python para fazer o gráfico da função $f(x)=120.1{,}08^x$. Um detalhe importante: como o separador de casas decimais é o ponto (.), no lugar de 1,08 precisamos digitar 1.08.

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(0, 5, 100)
y=120*1.08**x
plt.plot(x,y)
plt.show()
```

+ Código + Texto

```
[ ] import matplotlib.pyplot as plt
    import numpy as np
    x=np.linspace(0, 5, 100)
    y=120*1.08**x
    plt.plot(x,y)
    plt.show()
```



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

A seguir, estudaremos diversas aplicações relacionadas às funções.

TEMA 5 – APLICAÇÕES DE FUNÇÕES

Vimos o que são as funções lineares, quadráticas e exponenciais. No entanto, além dessas, há outros tipos.

A seguir, apresentaremos exemplos relacionados a outros tipos delas.

Exemplo 1: um estudo afirmou que o custo para a remoção de poluentes despejados por uma determinada indústria é dado por:

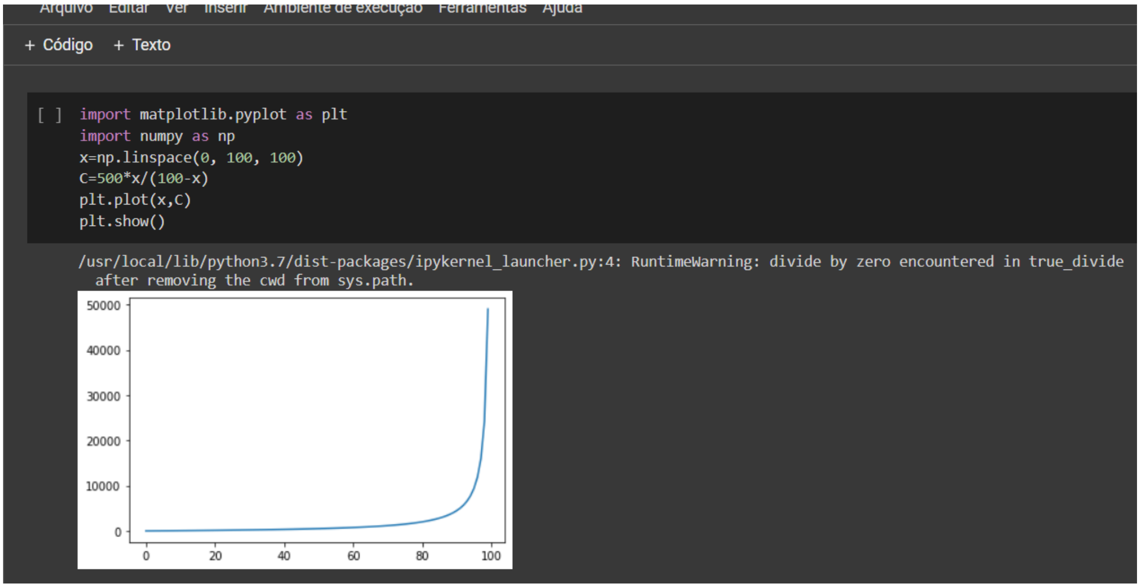
$$C(x)=500x/(100-x)$$

em que C é o custo e x é a porcentagem de poluentes a ser removida. Por meio do Python, construa o gráfico da função C com x pertencente ao intervalo [0, 100].

Como a função a ser representada graficamente é C(x)=500x/(100-x) com x variando de 0 a 100, tem-se:

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(0, 100, 100)
C=500*x/(100-x)
plt.plot(x,C)
plt.show()
```

que resulta em:



Fonte: Ricardo Alexandre Deckmann Zanardini/colab.research.google.com.

A mensagem que aparece diz que ocorreu uma divisão por zero. Observe que quando x é igual a 100, no denominador da fração, tem-se 100-100=0 e não há divisão por zero.

Analisando o gráfico da função, é possível observar que o custo de remoção de poluentes aumenta em função do aumento da porcentagem a ser removida, e quando x está próximo de 100%, o custo tem um crescimento bastante elevado.

Exemplo 2: estima-se que o custo de produção referente a x unidades produzidas de um determinado bem, por hora, é dado pela função

$$C(x)=0{,}05x^3-3x^2+58x+200$$

em que C é o custo, em reais. Determine o custo referente à produção de 30 unidades por hora.

$C(x)=0{,}05x^3-3x^2+58x+200$

$C(30)=0{,}05(30)^3-3(30)^2+58(30)+200$

$C(30)=0{,}05(27000)-3(900)+58(30)+200$

$C(30)=1350-2700+1740+200$

$C(30)=590$

O custo para a produção de 30 unidades por hora corresponde a R$ 590,00.

FINALIZANDO

Chegamos ao final desta etapa sobre métodos quantitativos. Estudamos, aqui, as funções lineares, quadráticas, exponenciais e algumas aplicações de outros tipos de funções. Além de algumas das diversas aplicações relacionadas às funções, vimos que é possível utilizar o Python para construir de maneira simples gráficos de funções.

REFERÊNCIAS

CASTANHERIA, N. P. **Matemática aplicada**. 3. ed. Curitiba: Ibpex, 2010.

DEMANA, F. D.; WAITS, B. K.; FOLEY, G. D.; KENNEDY, D. **Pré-cálculo**. 2. ed. São Paulo: Pearson, 2013.

FLEMING, D. M.; GONÇALVES, M. B. **Cálculo A**: função de uma variável. 2. ed. São Paulo: Pearson, 2007.

MÉTODOS QUANTITATIVOS

AULA 2

Prof. Ricardo Alexandre Deckmann Zanardini

CONVERSA INICIAL

Estudaremos limites e derivadas, que também são muito úteis em situações do cotidiano. Os conceitos relacionados a limites são úteis quando precisamos saber o comportamento de uma função quando a variável independente se aproxima de um certo valor. Quanto às derivadas, podemos utilizar no estudo de situações relacionadas a taxas de crescimento. Também podemos determinar máximos e mínimos de funções, tais como situações envolvendo lucro máximo ou custo mínimo, por meio das derivadas.

TEMA 1 – LIMITES

O conceito de limite está relacionado ao estudo do comportamento de uma função quando a variável independente *x* se aproxima de um determinado valor. Esse conceito pode ser utilizado, por exemplo, para fazermos o estudo do custo de remoção de poluentes de uma indústria quando a porcentagem de poluentes aumenta. Podemos determinar também o que acontece com o custo unitário de um determinado produto com o aumento da produção.

Uma forma de entendermos melhor é pensarmos em um exemplo:

Uma indústria produz camisetas estampadas com um custo de R$ 200,00 referente ao cliché (placa de metal em relevo destinada à impressão de imagens ou de textos) e um custo de R$ 15,00 referente a cada camiseta sem estampa. A função que relaciona o custo total C de produção de *x* camisetas é dada por

$$C(x) = 15x + 200.$$

Para calcularmos o custo unitário de produção, basta dividirmos o custo total C pelo número de camisetas produzidas *x*. Chamando de *U* o custo unitário, a função que fornece o custo unitário com base na produção corresponde a

$$U(x) = \frac{15x+200}{x}.$$

Dividindo cada termo do numerador por *x*, temos

$$U(x) = 15 + \frac{200}{x}.$$

Para termos uma ideia do comportamento do custo total e do custo unitário em relação ao aumento da produção, temos a seguinte tabela.

Tabela 1 – Custo unitário e custo total em relação ao aumento da produção

x	C(x)	U(x)
1	R$ 215,00	R$ 215,00
10	R$ 350,00	R$ 35,00
20	R$ 500,00	R$ 25,00
30	R$ 650,00	R$ 21,67
40	R$ 800,00	R$ 20,00
50	R$ 950,00	R$ 19,00
100	R$ 1.700,00	R$ 17,00
200	R$ 3.200,00	R$ 16,00
300	R$ 4.700,00	R$ 15,67
1000	R$ 15.200,00	R$ 15,20
10000	R$ 150.200,00	R$ 15,02

Observando o comportamento de *C* e de *U*, temos que, com o aumento da produção, o custo total aumenta e o custo unitário diminui.

Neste caso, qual é o custo unitário mínimo?

Para obtermos uma resposta, podemos pensar no que ocorre com o custo unitário quando a produção aumenta infinitamente. Assim, saberemos qual é o custo unitário mínimo. Em outras palavras, basta calcularmos o limite de U(x) quando *x* tende a infinito. Escrevemos a expressão da seguinte forma:

$$\lim_{x \to \infty} \left(15 + \frac{200}{x}\right).$$

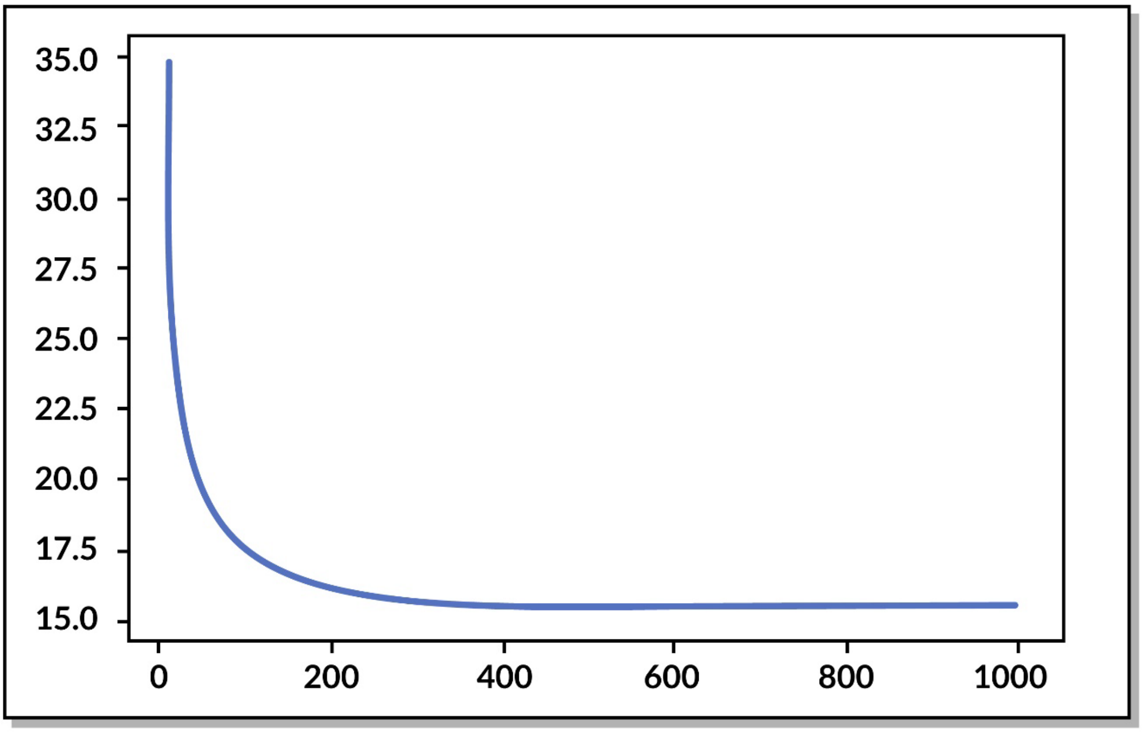
Como 15 é uma constante, a variação ocorre no termo 200/x. Quanto maior o valor de *x*, menor o resultado da divisão de 200 por *x*. Desta maneira, substituindo *x* por infinito, a expressão 200/x tende a zero.

Podemos dizer então que

$$\lim_{x \to \infty} \left(15 + \frac{200}{x}\right) = 15 + 0 = 15.$$

Quanto maior for a produção, menor o custo unitário, e quando *x* tende a infinito, o custo unitário tende a R$ 15,00.

Graficamente, temos



Os comandos utilizados para a construção do gráfico no Python foram:

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(0, 1000, 100)
C=15+200/x
plt.plot(x,C)
plt.show()
```

TEMA 2 – APLICAÇÕES E O PYTHON NO ESTUDO DE LIMITES

Problemas práticos e teóricos relacionados a limites podem ser resolvidos de uma forma analítica ou por meio do Python.

No nosso caso, utilizaremos a função “limit” da biblioteca sympy.

A forma necessária para a resolução de problemas relacionados a limites é bastante simples. Inicialmente vamos importar a biblioteca simbólica sympy digitando

```
from sympy import *
```

em uma célula de código do Google Colab.

Precisamos criar uma variável simbólica *x*

```
x=symbols("x")
```

Agora, vamos informar qual é a função, por exemplo,

$$y = 15 + 200/x$$

Para calcularmos o limite de *y* quando *x* tende a infinito, basta fazermos

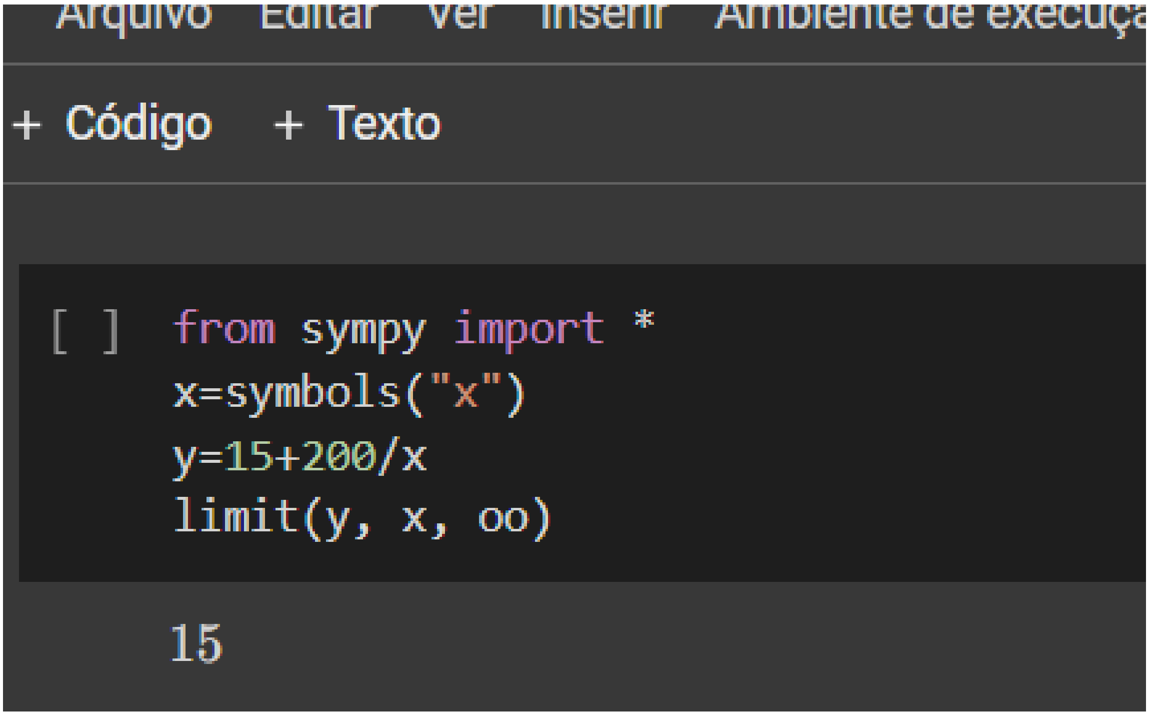
limit(y, x, oo)

Observe que o símbolo de infinito é representado por duas letras “o” seguidas.

A sequência total de comandos é:

```
from sympy import *
x=symbols("x")
y=15+200/x
limit(y, x, oo)
```

e o resultado é:



Esta função é a mesma vista anteriormente no exemplo relacionado ao custo unitário de produção de camisetas estampadas.

Ao utilizarmos o comando “limit”, precisamos informar qual é a função (y), qual é a variável (x) e para qual valor esta variável está tendendo (oo).

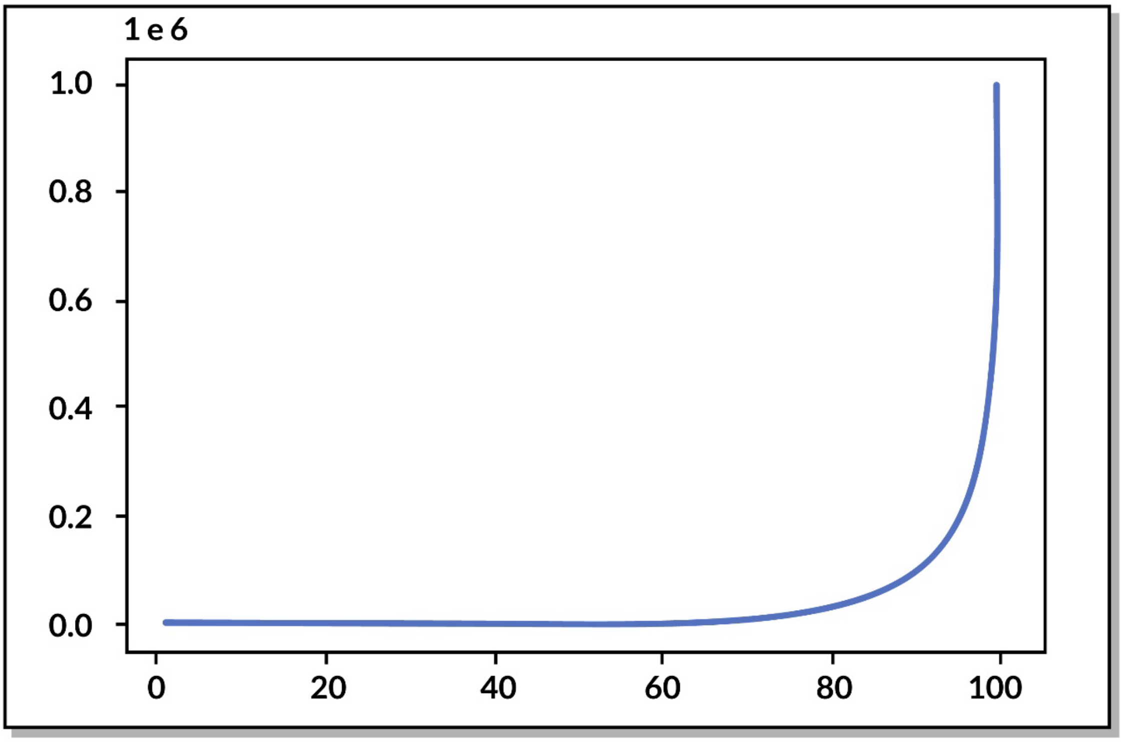
Exemplo:

O custo C de remoção dos poluentes de um rio é dado em função da porcentagem removida x. A relação entre o custo de remoção e a porcentagem removida é dada pela função

$$C(x) = 5000 + \frac{10000x}{100-x}.$$

Qual é o custo de remoção dos poluentes quando x tende a 100%.

Graficamente, temos:



Os comandos utilizados para o gráfico foram:

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(0, 100, 100)
C=5000+10000*x/(100-x)
plt.plot(x,C)
plt.show()
```

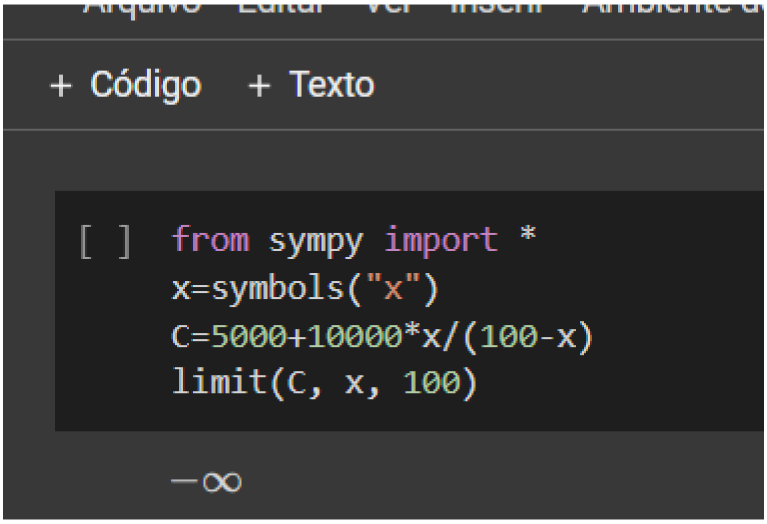
Observando o gráfico, é possível perceber que o custo tende a infinito quando x tende a 100%.

Podemos obter o resultado deste limite por meio do Python:

```
from sympy import *
x=symbols("x")
```

C=5000+10000*x/(100-x)

limit(C, x, 100)

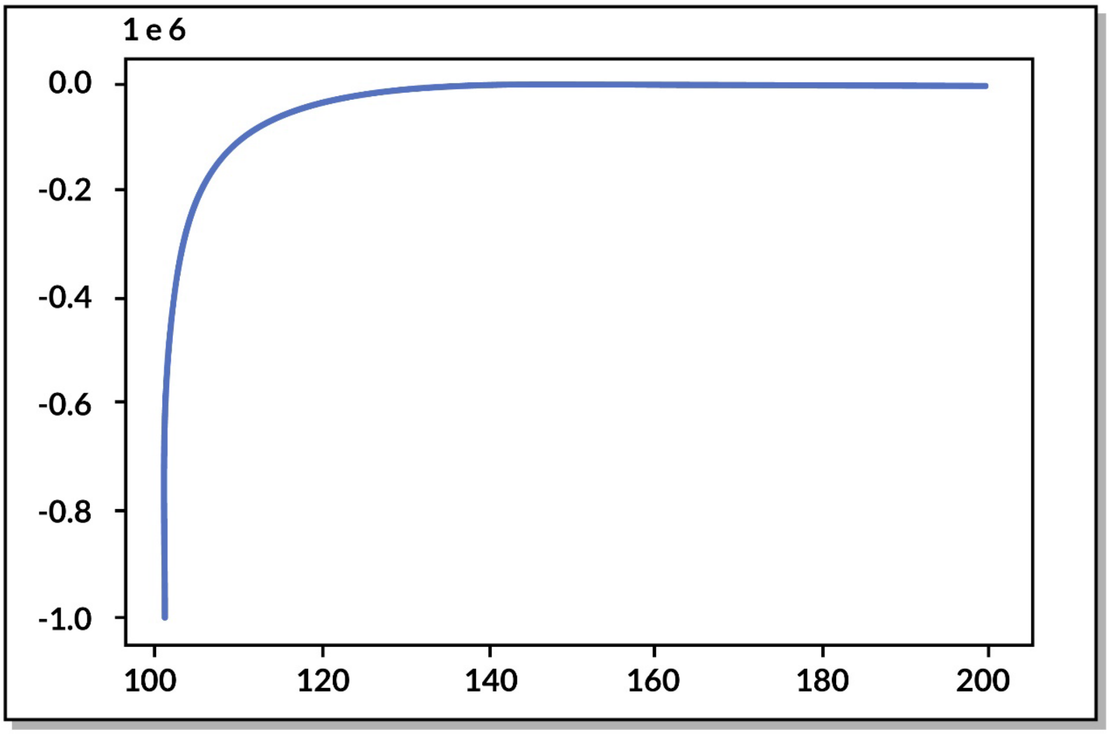


Sabemos que o resultado é um valor positivo muito grande, ou seja, infinito positivo. No entanto, o resultado apresentado pelo Python é um valor negativo.

Vamos representar graficamente a função

$$C(x) = 5000 + \frac{10000x}{100-x},$$

mas agora considerando valores maiores do que 100 para x:



onde os comandos utilizados foram:

```
import matplotlib.pyplot as plt
import numpy as np
x=np.linspace(100, 200, 100)
C=5000+10000*x/(100-x)
plt.plot(x,C)
plt.show()
```

O gráfico está abaixo do eixo *x* e, quando *x* se aproxima de 100, a função tende a menos infinito. Mas por que isto ocorreu? A resposta é bem simples e, para isso, precisamos falar um pouco sobre limites laterais.

Considerando a função

$$C(x) = 5000 + \frac{10000x}{100-x},$$

quando *x* se aproxima de 100 por valores menores do que 100, por exemplo, 99,999, 99,999999, 99,99999999 e assim por diante, temos um valor positivo para 100-x. Mas se *x* se aproxima de 100 por valores maiores do que 100, por exemplo 100,001, 100,000000001 e assim por diante, temos um valor negativo para 100-x.

Podemos dizer que o limite à esquerda de C(x) quando *x* se aproxima de 100 é ∞ e o limite à direita de C(x) quando *x* se aproxima de 100 é -∞.

O limite à esquerda é representado por

$$\lim_{x \to \infty^-} \left(5000 + \frac{10000x}{100 - x}\right)$$

e limite à direita é representado por

$$\lim_{x \to \infty^+} \left(5000 + \frac{10000x}{100-x}\right).$$

No Python, podemos identificar limites laterais acrescentando um parâmetro ao comando "limit". Quando temos um limite à esquerda, basta acrescentarmos "dir='-'", e quando temos um limite à direita, basta acrescentarmos "dir='+'". Desta maneira, o comando para calcularmos o limite à esquerda passa a ser

```
limit(C, x, 100, dir='-')
```

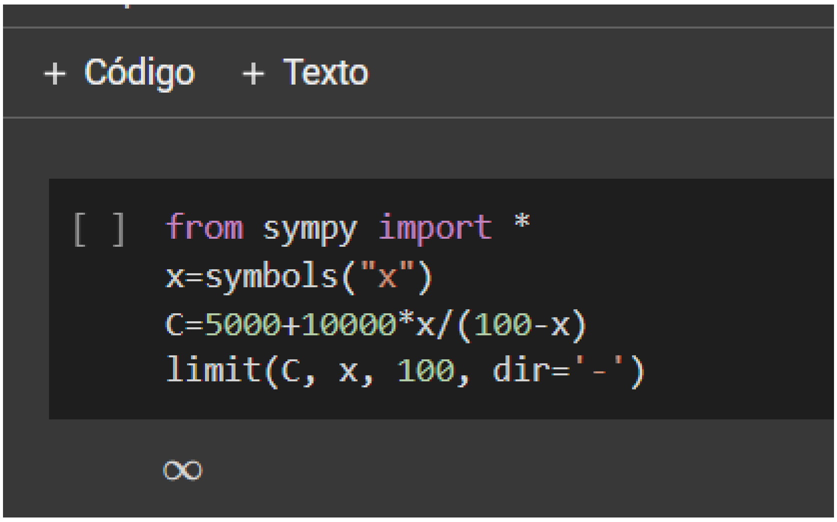
A sequência de comandos é:

from sympy import *

x=symbols("x")

C=5000+10000*x/(100-x)

limit(C, x, 100, dir='-')



Desta maneira, garantimos que o limite foi calculado com base em valores se aproximando de *x* pela esquerda.

Exemplo:

Calcule, por meio do Python, o seguinte limite:

$$\lim_{x \to 2} \frac{x^2 - 5x + 6}{x - 2}$$

Basta fazermos

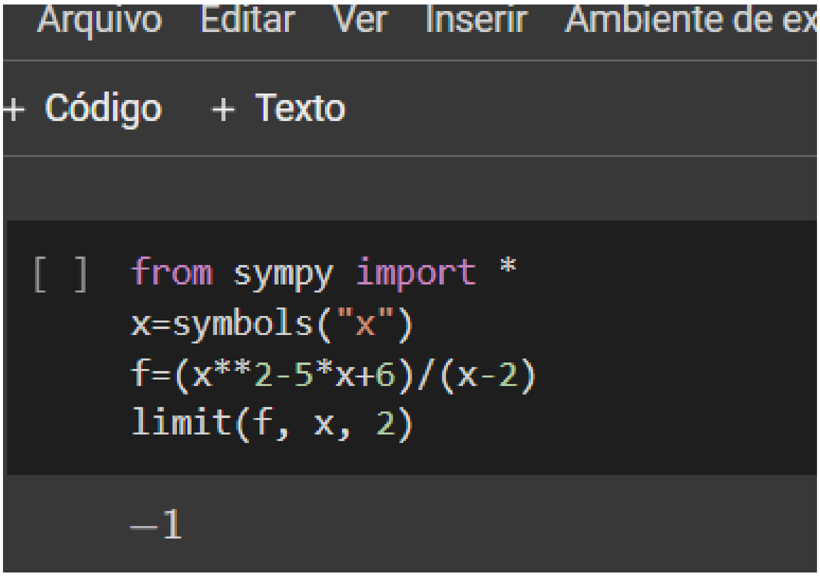
from sympy import *

x=symbols("x")

f=(x**2-5*x+6)/(x-2)

limit(f, x, 2)

e o resultado corresponde a -1.



Exemplo:

Considerando

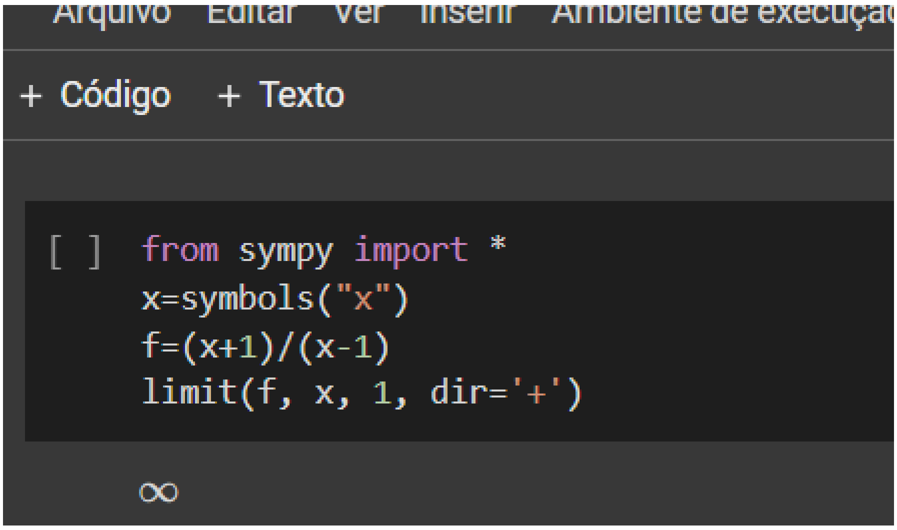
$$f(x) = \frac{x+1}{x-1},$$

calcule o limite de f(x) quando *x* tende a 1 pela direita.

Os comandos necessários são:

```
from sympy import *
x=symbols("x")
f=(x+1)/(x-1)
limit(f, x, 1, dir='+')
```

o que resulta em:

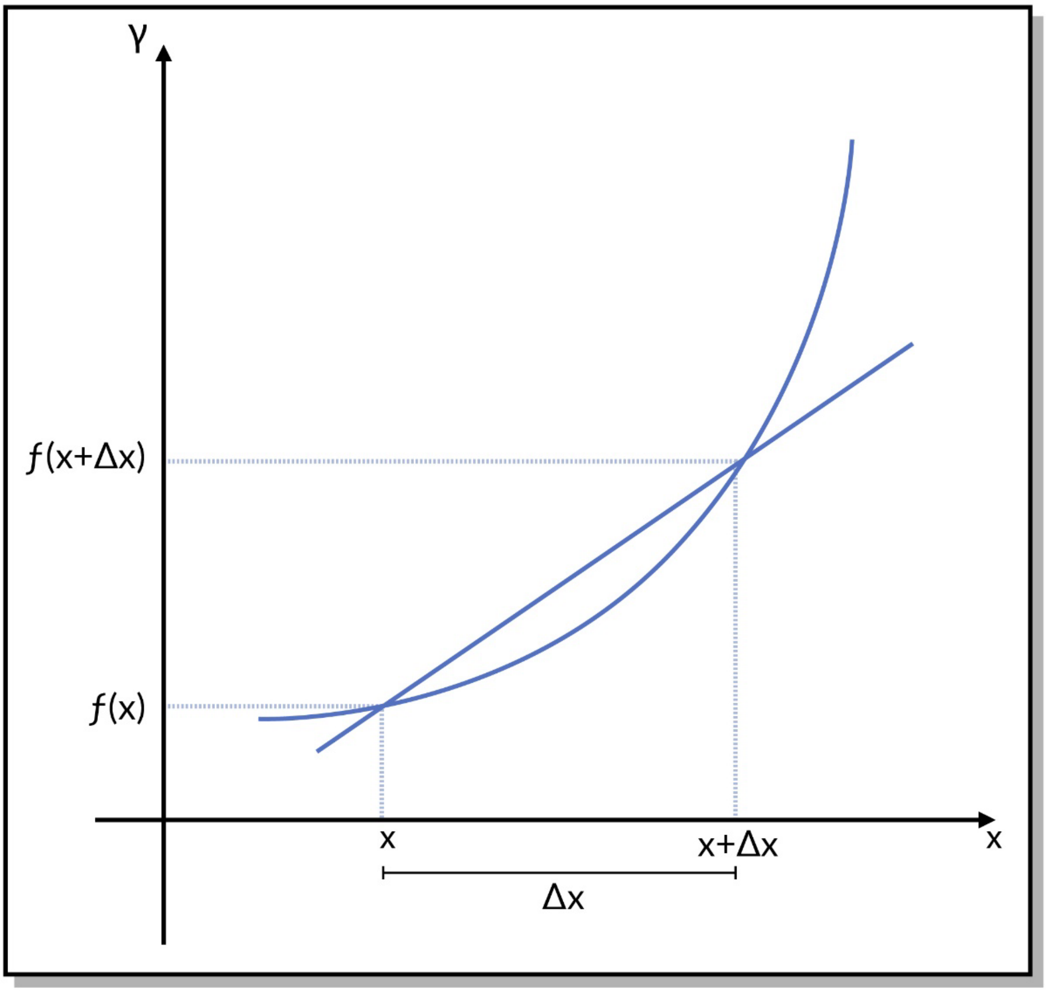


TEMA 3 – DERIVADAS

Quando tratamos de problemas relacionados a taxas de variação ou a movimentos, a derivação é utilizada. Utilizamos derivadas para obtermos máximos ou mínimos de funções ou para o traçado de curvas, algo muito útil em imagens obtidas por meio de computadores.

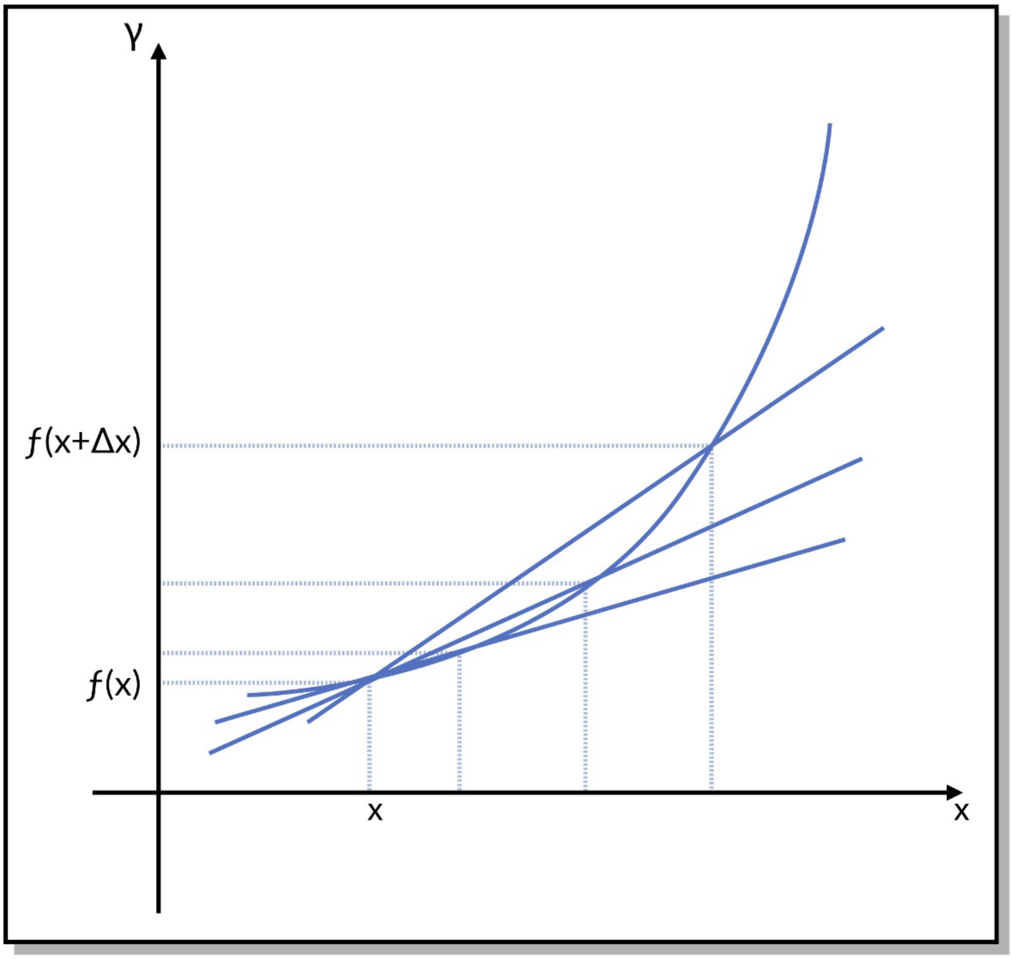
Por definição, uma derivada é dada por

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$



Geometricamente, a derivada corresponde à inclinação da reta tangente ao gráfico de *f* no ponto *x*.

Com base em dois pontos dados, temos uma reta secante à curva definida por eles, mas, quando $\Delta x \to 0$, estes pontos se aproximam, e assim a inclinação da reta secante fica cada vez mais próxima da inclinação da reta tangente à curva.



Na prática, não precisamos utilizar a definição de derivada para a resolução de problemas, pois com base na definição, diversas regras de derivação são apresentadas.

A derivada de $f(x) = x^2$, por exemplo, corresponde a $f'(x) = 2x$, e a derivada de $f(x) = x^3$, por exemplo, corresponde a $f'(x) = 3x^2$.

As regras de derivação facilitam a obtenção das respectivas derivadas.

TEMA 4 – REGRAS DE DERIVAÇÃO

Com base em regras de derivação, o cálculo da derivada passa a ser mais rápido e eficiente. Vamos estudar algumas dessas regras.

A derivada de uma função constante é igual a zero.

Sabemos que a definição de derivada corresponde a

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

Como a função *f* é uma constante, então f(x)=c. Sendo assim, f(x+Δx) = c. Isso ocorre porque, como a função é constante, sempre é igual a *c* para qualquer valor de *x*.

Logo,

$$f'(x) = \lim_{\Delta x \to 0} \frac{c - c}{\Delta x}$$

$$f'(x) = \lim_{\Delta x \to 0} \frac{0}{\Delta x}$$

$$f'(x) = \lim_{\Delta x \to 0} 0 = 0$$

Exemplo:

Considere a função f(x) = 1000. Calcule a respectiva derivada.

Como a derivada de uma constante é zero, se f(x) = 1000, então f'(x) = 0.

Quando temos funções relacionadas a potências, a regra de derivação é bem simples.

Por exemplo, com base na definição de derivada, podemos mostrar que se f(x) = x^2, então f'(x) = 2x.

Como $f(x) = x^2$, $f(x + \Delta x) = (x + \Delta x)^2 = x^2 + 2x\Delta x + \Delta x^2$

Logo,

$$f'(x) = \lim_{\Delta x \to 0} \frac{x^2 + 2x\Delta x + (\Delta x)^2 - x^2}{\Delta x}$$

$$f'(x) = \lim_{\Delta x \to 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x}$$

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta x(2x + \Delta x)}{\Delta x}$$

$$f'(x) = \lim_{\Delta x \to 0} (2x + \Delta x)$$

Sabendo que $\Delta x \to 0$

$$f'(x) = 2x$$

Da mesma maneira, se f(x) = x^3, então f'(x) = $3x^2$. Quando f(x) = x^4, f'(x) = $4x^3$ e assim por diante.

Logo, a regra da potência diz que se $f(x) = x^n$, então $f'(x) = nx^{n-1}$. E se $f(x) = cx^n$, então $f'(x) = cnx^{n-1}$.

A derivada da soma é a soma das derivadas, ou seja, se $f(x) = 2x^5+7x^3$, então $f(x) = 10x^4+21x^2$, ou seja, basta calcular a derivada de cada termo utilizando as regras de derivação.

Exemplo:

Calcule a derivada da seguinte função:

$$f(x) = x^7+5x^3-3x^{-4}.$$

Como a derivada de x^7 é dada por $7x^{7-1}$ que resulta em $7x^6$, a derivada de $5x^3$ resulta em $3.5x^{3-1} = 15x^2$ e a derivada de $-3x^{-4}$ corresponde a $(-4)(-3)x^{-4-1} = 12x^{-5}$, temos que se $f(x) = x^7+5x^3-3x^{-4}$, $f'(x) = 7x^6+15x^2+12x^{-5}$.

Na prática, podemos utilizar o Python para calcularmos derivadas de funções.

TEMA 5 – APLICAÇÕES DAS DERIVADAS E O USO DO PYTHON

Estudaremos agora algumas das mais diversas aplicações das derivadas, mas antes vamos aprender a calcular derivadas por meio do Python.

Precisamos definir a função e em seguida utilizar o comando “diff()” da biblioteca “sympy” do Python. Para isso, basta especificarmos qual é a função e qual é a variável de referência para a derivação.

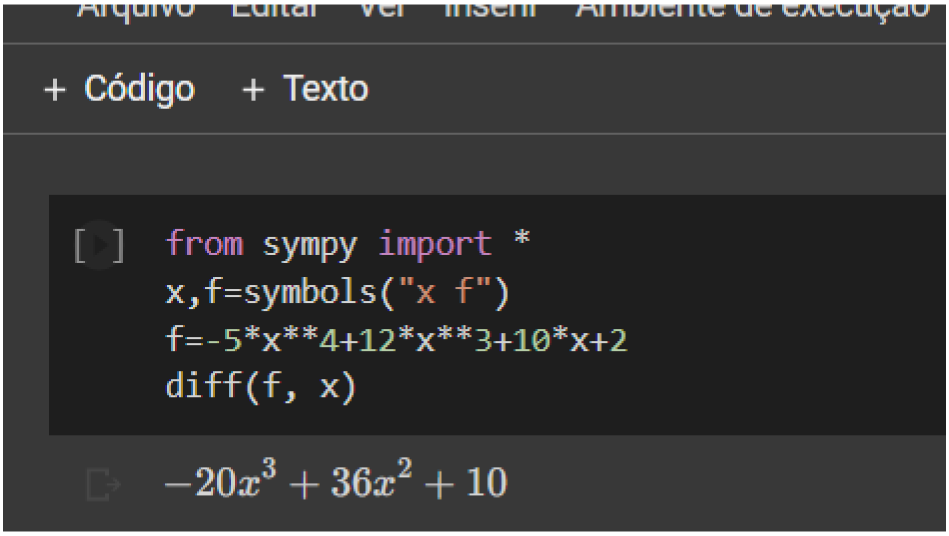
Se temos uma função “f” e queremos obter a respectiva derivada em relação a “x”, o comando é dado por “diff(f, x)”.

Exemplo:

Calcule, por meio do Python, a derivada primeira da função $f(x) = -5x^4+12x^3+10x+2$.

Vamos utilizar os seguintes comandos:

```
from sympy import *
x,f=symbols("x f")
f=-5*x**4+12*x**3+10*x+2
diff(f, x)
```



Logo, a derivada de $f(x) = -5x^4+12x^3+10x+2$ corresponde a

$$f'(x) = -20x^3+36x^2+10.$$

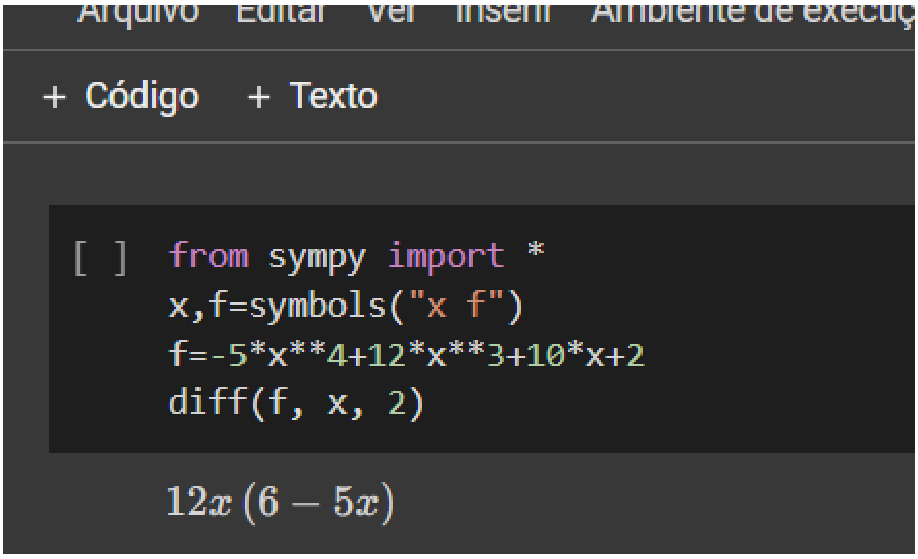
Podemos derivar sucessivamente as funções. Derivando a primeira vez, temos a derivada primeira f'. Se derivarmos novamente essa função, temos a derivada segunda f'' e assim por diante.

Exemplo:

Calcule, por meio do Python, a derivada segunda da função $f(x) = -5x^4+12x^3+10x+2$.

Vamos utilizar os seguintes comandos:

```
from sympy import *
x,f=symbols("x f")
f=-5*x**4+12*x**3+10*x+2
diff(f, x, 2)
```



Assim, a derivada segunda da função $f(x) = -5x^4+12x^3+10x+2$ corresponde a

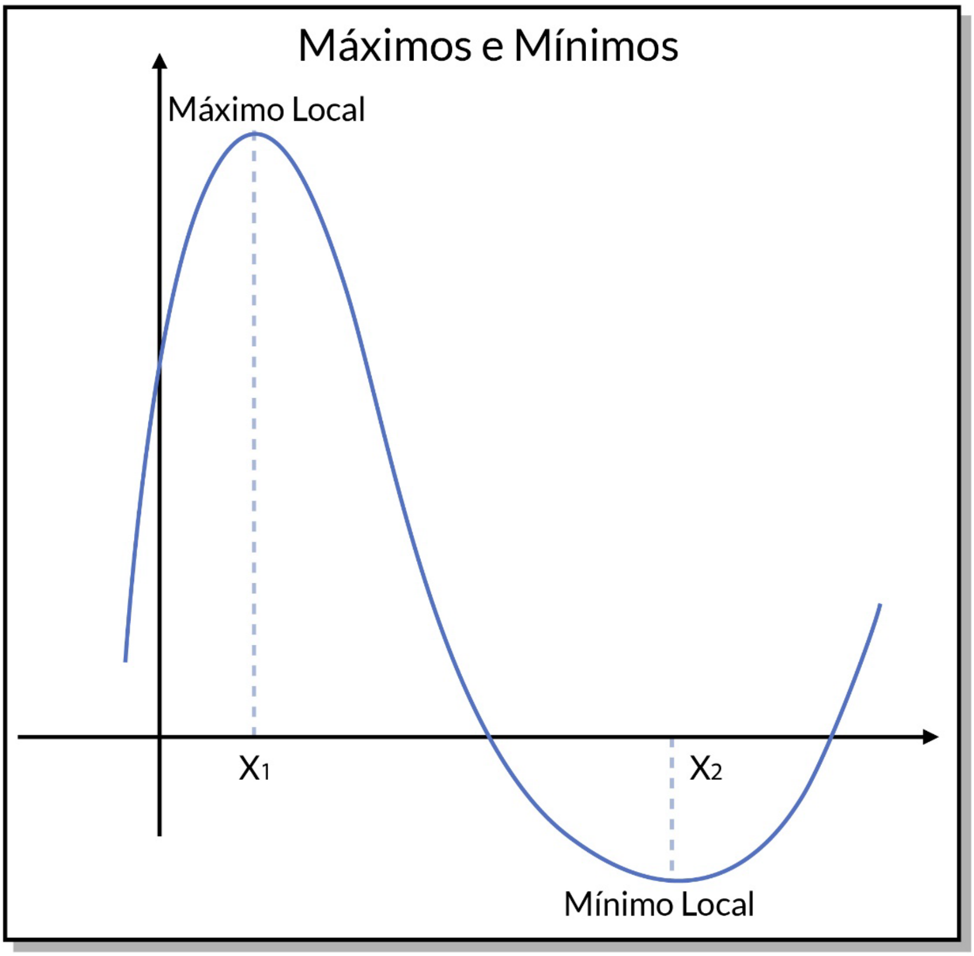
$$f''(x)=12x(6-5x).$$

De forma equivalente, podemos fazer a multiplicação de 12x pelos termos que estão entre parênteses e escrever esta derivada como sendo

$$f''(x)=72x-60x^2.$$

A derivada segunda de uma função está associada à concavidade da função em torno de um determinado ponto.

Quando a derivada segunda calculada no ponto resulta em um valor positivo, significa que a concavidade da função está voltada para cima em torno do ponto, e assim temos um ponto de mínimo local da função. E quando o resultado da derivada segunda calculada no ponto é negativo, a concavidade da função em torno do ponto está voltada para baixo, e assim temos um ponto de máximo local da função.



Uma das mais importantes aplicações da derivada consiste em encontrarmos máximos e mínimos de funções. O critério utilizado é que a derivada primeira em um ponto de máximo ou de mínimo é igual a zero.

É importante ressaltar que nem sempre derivada igual a zero indica um ponto de máximo ou de mínimo. Por isso é preciso utilizar o teste da derivada segunda que se baseia na concavidade da função em torno do ponto.

Vamos entender melhor acompanhando o exemplo a seguir:

A relação entre o preço de venda *x* de um modelo de aparelho de som e o lucro y referente à comercialização desse aparelho é dado pela função $y=-4x^2+4000x-200000$. Sendo assim, qual é o preço de venda que maximiza o lucro?

Como a função é dada por $y = -4x^2 + 4000x - 200000$, a respectiva derivada corresponde a $y' = -8x + 4000$.

Como queremos obter o ponto crítico (possível ponto de máximo ou de mínimo da função), basta igualarmos a derivada primeira a zero, ou seja, basta resolvermos a equação

$$-8x+4000=0$$

Isolando *x*, temos:

$$-8x=-4000$$
$$8x=4000$$
$$x=4000/8$$
$$x=500$$

Já sabemos que $x = 500$ é um candidato a ponto de máximo ou de mínimo local da função. Precisamos do teste da derivada segunda para sabermos se este ponto é minimizador ou maximizador local da função. Como a derivada de $y' = -8x + 4000$ é $y'' = -8$, podemos concluir que $x = 500$ é um maximizador local da função, pois a respectiva derivada segunda é negativa.

Assim, o preço que maximiza o lucro é R$ 500,00.

FINALIZANDO

No decorrer desta etapa, vimos que podemos utilizar limites para resolver diversos problemas reais utilizando o Python. Aprendemos o que são derivadas e como é possível obtermos derivadas de funções utilizando o comando “diff” do Python. Vimos também que as derivadas estão relacionadas a problemas de máximos e mínimos, taxa de variação, dentre outras aplicações.

REFERÊNCIAS

CASTANHERIA, N. P. **Matemática aplicada**. 3. ed. Curitiba: Ibpex, 2010.

DEMANA, F. D. et al. **Pré-cálculo**. 2. ed. São Paulo: Pearson, 2013.

FLEMING, D. M.; GONÇALVES, M. B. **Cálculo A**: função de uma variável. 2. ed. São Paulo: Pearson, 2007.

MÉTODOS QUANTITATIVOS

AULA 3

Prof. Ricardo Alexandre Deckmann Zanardini

CONVERSA INICIAL

Estudaremos os índices econômicos, elementos utilizados na análise de variações de preços, quantidades e valor de um determinado item ou de um conjunto formado por mais que um item. Com base na análise realizada, podemos conhecer melhor o respectivo cenário e ter parâmetros para auxiliar no processo de tomadas de decisões.

TEMA 1 – RELATIVOS

O índice mais simples é conhecido como *relativo*. Por meio do relativo, temos o índice de variação entre dois números associados a um determinado item. Assim, podemos realizar comparações entre preços, quantidades ou valores conhecidos em duas datas diferentes, uma delas sendo a data-base.

O relativo de preço é dado pela divisão entre o preço na data considerada (p_2) pelo preço na data-base (p_1):

$$R_p = \frac{p_2}{p_1}.$$

O relativo de quantidade é dado pela divisão entre a quantidade na data considerada (q_2) pela quantidade na data-base (q_1):

$$R_q = \frac{q_2}{q_1}.$$

O relativo de valor é dado pela divisão entre o valor na data considerada (v_2) pelo valor na data-base (v_1):

$$R_v = \frac{v_2}{v_1}.$$

Em métodos quantitativos, o valor é o produto entre preço e quantidade.

Vamos resolver alguns exemplos.

Exemplo:

Considerando que em junho de 2021 o litro de uma determinada marca de leite longa vida custava R$ 4,00 e que em junho de 2022 o mesmo item estava custando R$ 5,00, qual é o respectivo relativo de preço considerando 2021 como a data-base?

Para resolvermos o exemplo, temos que R$ 4,00 era o preço na data-base, ou seja, $p_1 = 4$, e que R$ 5,00 é o preço na data considerada, ou seja, $p_2 = 5$. Assim, o relativo de preço é dado por:

$$R_p = \frac{p_2}{p_1}$$

$$R_p = \frac{5}{4}$$

$$R_p = 1{,}25$$

Podemos concluir que em junho de 2022 o preço do leite longa vida dessa determinada marca corresponde a 1,25 vez o preço do mesmo leite em junho de 2021. Também podemos representar esse relativo na forma de porcentagem. Basta multiplicarmos 1,25 por 100, o que resulta em:

$$R_p = 125\%$$

Podemos dizer também que o preço em 2021 corresponde a 125% do preço em 2022 e que o aumento nesse período corresponde a 25%, pois 125% - 100% = 25%.

Exemplo:

Em junho de 2021, um estabelecimento comercial vendeu 12323 caixas de um litro de uma determinada marca de leite longa vida. Em junho de 2022, foram vendidas 11922 caixas do mesmo item. Com base nessas informações, qual é o respectivo relativo de quantidade considerando 2021 como a data-base?

Neste caso, $q_1 = 12323$ e $q_2 = 11922$, o que resulta em:

$$R_q = \frac{q_2}{q_1}$$

$$R_q = \frac{11922}{12323}$$

$$R_q = 0{,}967459$$

$$R_q = 96{,}7459\%$$

Arredondando com duas casas após a vírgula, $R_q = 96{,}75\%$. Assim, a quantidade desse item vendida em junho de 2022 corresponde a 96,75% da quantidade vendida em junho de 2021, uma redução de 3,25%.

Exemplo:

Em junho de 2021, um estabelecimento comercial vendeu 12323 caixas de um litro de uma determinada marca de leite longa vida a R$ 4,00 cada. Em junho de 2022, foram vendidas 11922 caixas do mesmo item a R$ 5,00 cada. Considerando essas informações, qual é o respectivo relativo de valor em que 2021 é a data-base?

Sabemos que o valor é o produto do preço pela quantidade. Assim, $v_1 = p_1 \,.\, q_1$ e $v_2 = p_2 \,.\, q_2$. Como $p_1 = 4$ e $q_1 = 12323$, $v_1 = (4)\,.\,(12323)$, $v_1 = 49292$ e $p_2 = 5$ e $q_2 = 11922$, $v_2 = (5)\,.\,(11922)$, $v_2 = 59610$, temos:

$$R_v = \frac{v_2}{v_1}$$

$$R_v = \frac{59610}{49292}$$

$$R_v = 1{,}209324$$

O relativo de valor corresponde a 1,209324, que equivale a 120,9324%. Como o aumento no valor é igual a 120,9324% - 100% = 20,9324%, em 2022 o valor associado ao litro dessa determinada marca de leite longa vida no estabelecimento comercial aumentou em 20,9324%.

Quando consideramos apenas um item, os relativos são importantes elementos de medição, mas, quando um conjunto é composto por mais que um item, podemos utilizar os índices agregativos simples.

Nesta etapa, estudaremos os índices de Bradstreet, de Sauerbeck, de Laspeyres e de Paasche.

TEMA 2 – ÍNDICES DE BRADSTREET

O índice de Bradstreet, também conhecido como *índice de Bradstreet-Dûtot*, é o mais elementar dos índices agregativos simples e se baseia nas médias aritméticas simples. O índice de Bradstreet de preço é dado pela divisão entre a soma dos preços na data considerada pela soma dos preços na data-base:

$$B_p = \frac{\sum_{i=1}^{n} p_2^i}{\sum_{i=1}^{n} p_1^i}$$

Desenvolvendo os somatórios, temos

$$B_p = \frac{\sum_{i=1}^{n} p_2^i}{\sum_{i=1}^{n} p_1^i} = \frac{p_2^1+p_2^2+p_2^3+\cdots+p_2^n}{p_1^1+p_1^2+p_1^3+\cdots+p_1^n}.$$

O índice de Bradstreet de quantidade é:

$$B_q = \frac{\sum_{i=1}^{n} q_2^i}{\sum_{i=1}^{n} q_1^i} = \frac{q_2^1 + q_2^2 + q_2^3 + \cdots + q_2^n}{q_1^1 + q_1^2 + q_1^3 + \cdots + q_1^n}$$

e o índice de Bradstreet de valor é:

$$B_v = \frac{\sum_{i=1}^{n} v_2^i}{\sum_{i=1}^{n} v_1^i} = \frac{v_2^1 + v_2^2 + v_2^3 + \cdots + v_2^n}{v_1^1 + v_1^2 + v_1^3 + \cdots + v_1^n}$$

onde $v_1 = p_1 \,.\, q_1$ e $v_2 = p_2 \,.\, q_2$.

Vamos resolver um exemplo.

Exemplo:

Uma mercearia está realizando um estudo sobre o impacto na variação do preço do açúcar, do arroz, da farinha de trigo e do feijão. Para isso, está considerando os preços e as respectivas quantidades comercializadas na data 1 e os preços e as quantidades referentes à data 2, 6 meses após a data 1.

Na tabela a seguir, são apresentados os preços e as quantidades na data 1 e na data 2 para cada um dos itens.

Tabela 1 – Preços e quantidades

Item	Preço 1	Quantidade 1	Preço 2	Quantidade 2
Açúcar	R$ 3,90	320	R$ 4,40	332
Arroz	R$ 5,50	410	R$ 6,10	380
Farinha de trigo	R$ 4,20	492	R$ 7,80	440
Feijão	R$ 6,30	370	R$ 8,40	400

Considerando esses dados, calcule os índices de Bradstreet de preço, de quantidade e de valor.

Para o índice de Bradstreet de preço, basta fazermos a soma dos preços na data 2 e dividirmos pela soma dos preços na data 1:

$$B_p = \frac{\sum_{i=1}^{n} p_2^i}{\sum_{i=1}^{n} p_1^i} = \frac{p_2^1 + p_2^2 + p_2^3 + \cdots + p_2^n}{p_1^1 + p_1^2 + p_1^3 + \cdots + p_1^n}$$

$$B_p = \frac{\sum_{i=1}^{n} p_2^i}{\sum_{i=1}^{n} p_1^i} = \frac{4{,}40 + 6{,}10 + 7{,}80 + 8{,}40}{3{,}90 + 5{,}50 + 4{,}20 + 6{,}30}$$

$$B_p = \frac{\sum_{i=1}^{n} p_2^i}{\sum_{i=1}^{n} p_1^i} = \frac{26{,}70}{19{,}90}$$

$$B_p = \frac{\sum_{i=1}^{n} p_2^i}{\sum_{i=1}^{n} p_1^i} = 1{,}341709$$

$$B_p = \frac{\sum_{i=1}^{n} p_2^i}{\sum_{i=1}^{n} p_1^i} = 134{,}17\%$$

O índice de Bradstreet de preço corresponde a 134,17%. Logo, o conjunto de itens teve um aumento de preço de 34,17%.

O índice de quantidade de Bradstreet corresponde à divisão da soma das quantidades na data 2 pela soma das quantidades na data 1:

$$B_q = \frac{\sum_{i=1}^{n} q_2^i}{\sum_{i=1}^{n} q_1^i} = \frac{q_2^1 + q_2^2 + q_2^3 + \cdots + q_2^n}{q_1^1 + q_1^2 + q_1^3 + \cdots + q_1^n}$$

$$B_q = \frac{\sum_{i=1}^{n} q_2^i}{\sum_{i=1}^{n} q_1^i} = \frac{332 + 380 + 440 + 400}{320 + 410 + 492 + 370}$$

$$B_q = \frac{\sum_{i=1}^{n} q_2^i}{\sum_{i=1}^{n} q_1^i} = \frac{1552}{1592}$$

$$B_q = \frac{\sum_{i=1}^{n} q_2^i}{\sum_{i=1}^{n} q_1^i} = 0{,}974874$$

$$B_q = \frac{\sum_{i=1}^{n} q_2^i}{\sum_{i=1}^{n} q_1^i} = 97{,}49\%$$

O índice de Bradstreet de quantidade indica uma redução nas vendas desses itens.

Para calcularmos o índice de Bradstreet de valor, precisamos multiplicar os preços na data 1 pelas respectivas quantidades na data 1:

Tabela 2 – Preços e quantidades na data 1

Item	Preço 1	Quantidade 1	Valor 1
Açúcar	R$ 3,90	320	R$ 1.248,00
Arroz	R$ 5,50	410	R$ 2.255,00

Farinha de trigo	R$ 4,20	492	R$ 2.066,40
Feijão	R$ 6,30	370	R$ 2.331,00

e multiplicar os preços na data 2 pelas respectivas quantidades na data 2:

Tabela 3 – Preços e quantidades na data 2

Item	Preço 2	Quantidade 2	Valor 2
Açúcar	R$ 4,40	332	R$ 1.460,80
Arroz	R$ 6,10	380	R$ 2.318,00
Farinha de trigo	R$ 7,80	440	R$ 3.432,00
Feijão	R$ 8,40	400	R$ 3.360,00

Tendo os valores na data 1 e na data 2, podemos calcular o respectivo índice de Bradstreet:

$$B_v = \frac{\sum_{i=1}^{n} v_2^i}{\sum_{i=1}^{n} v_1^i} = \frac{v_2^1 + v_2^2 + v_2^3 + \cdots + v_2^n}{v_1^1 + v_1^2 + v_1^3 + \cdots + v_1^n}$$

$$B_v = \frac{\sum_{i=1}^{n} v_2^i}{\sum_{i=1}^{n} v_1^i} = \frac{1460{,}80 + 2.18{,}00 + 3432{,}00 + 3360{,}00}{1248{,}00 + 2255{,}00 + 2066{,}40 + 2331{,}00}$$

$$B_v = \frac{\sum_{i=1}^{n} v_2^i}{\sum_{i=1}^{n} v_1^i} = \frac{10570{,}80}{7900{,}40}$$

$$B_v = \frac{\sum_{i=1}^{n} v_2^i}{\sum_{i=1}^{n} v_1^i} = 1{,}338008$$

$$B_v = \frac{\sum_{i=1}^{n} v_2^i}{\sum_{i=1}^{n} v_1^i} = 133{,}80\%$$

Mesmo com uma redução na quantidade, é possível identificar um aumento de 33,80% no valor associado a esses itens.

Os índices de Bradstreet estão sujeitos à soma de preços, de quantidades e de valores em diferentes unidades, tais como quilos, litros, dúzias, entre outros. Uma alternativa que não leva em consideração as unidades é utilizar um índice baseado nos respectivos relativos. Os índices de Sauerbeck são obtidos por meio dessa ideia.

TEMA 3 – ÍNDICES DE SAUERBECK

Os índices de preço, quantidade e valor de Sauerbeck levam em consideração os relativos de preço, de quantidade e de valor. Como os índices de Sauerbeck consideram os relativos, não precisamos nos preocupar com as unidades de medida adotadas para cada item.

Para o cálculo de cada um desses índices, podemos utilizar média aritmética, média harmônica ou média geométrica. Sendo assim, há três formas diferentes para calcularmos cada um desses índices, e a escolha é feita de acordo com quem está realizando o estudo.

Nesta etapa, abordaremos os índices aritméticos de Sauerbeck. Para calcular os índices aritméticos de preço, quantidade e valor, precisamos calcular os relativos de preço (r_p), de quantidade (r_q) e de valor (r_v) de cada item. Em seguida, somamos esses relativos e dividimos o resultado de cada soma pelo número total de itens *n*.

Índice aritmético de Sauerbeck de preço:

$$S_p^A = \frac{\sum_{i=1}^{n} r_p^i}{n}$$

Índice aritmético de Sauerbeck de quantidade:

$$S_q^A = \frac{\sum_{i=1}^{n} r_q^i}{n}$$

Índice aritmético de Sauerbeck de valor:

$$S_v^A = \frac{\sum_{i=1}^{n} r_v^i}{n}$$

Exemplo:

Uma mercearia está realizando um estudo sobre o impacto na variação do preço do açúcar, do arroz, da farinha de trigo e do feijão. Para isso, está considerando os preços e as respectivas quantidades comercializadas na data 1 e os preços e as quantidades referentes na data 2, 6 meses após a data 1.

Na tabela a seguir, são apresentados os preços e as quantidades na data 1 e na data 2 para cada um dos itens.

Tabela 4 – Preços e quantidades

Item	Preço 1	Quantidade 1	Preço 2	Quantidade 2
Açúcar	R$ 3,90	320	R$ 4,40	332
Arroz	R$ 5,50	410	R$ 6,10	380
Farinha de trigo	R$ 4,20	492	R$ 7,80	440
Feijão	R$ 6,30	370	R$ 8,40	400

Considerando esses dados, calcule os índices aritméticos de Sauerbeck de preço, de quantidade e de valor.

Para calcularmos o índice aritmético de preço de Sauerbeck, precisamos dos relativos de preço de cada item. Precisamos então dividir cada preço na data 2 pelo respectivo preço na data 1, o que resulta em:

Tabela 5 – Divisão de preços

Item	Preço 1	Preço 2	p_2/p_1
Açúcar	R$ 3,90	R$ 4,40	1,13
Arroz	R$ 5,50	R$ 6,10	1,11
Farinha de trigo	R$ 4,20	R$ 7,80	1,86
Feijão	R$ 6,30	R$ 8,40	1,33

O índice aritmético de Sauerbeck de preço é calculado pela divisão da soma dos relativos de preço pelo total de itens:

$$S_p^A = \frac{\sum_{i=1}^{n} r_p^i}{n}$$

$$S_p^A = \frac{1{,}13 + 1{,}11 + 1{,}86 + 1{,}33}{4}$$

$$S_p^A = \frac{5{,}43}{4}$$

$$S_p^A = 1{,}356943$$

$$S_p^A = 135{,}69\%$$

O índice aritmético de Sauerbeck de preço indica um aumento de 35,69% no conjunto de itens da mercearia.

Para calcularmos o índice aritmético de Sauerbeck de quantidade, precisamos dividir a soma dos relativos de quantidade pelo total de itens.

Logo, o primeiro passo é calcularmos os relativos de quantidade:

Tabela 6 – Relativos de quantidade

Item	Quantidade 1	Quantidade 2	q_2/q_1
Açúcar	320	332	1,04
Arroz	410	380	0,93
Farinha de trigo	492	440	0,89
Feijão	370	400	1,08

Índice aritmético de Sauerbeck de quantidade:

$$S_q^A = \frac{\sum_{i=1}^{n} r_q^i}{n}$$

$$S_q^A = \frac{1{,}04 + 0{,}93 + 0{,}89 + 1{,}08}{4}$$

$$S_q^A = \frac{3{,}94}{4}$$

$$S_q^A = 0{,}984930$$

$$S_q^A = 98{,}49\%$$

Com base no índice aritmético de Sauerbeck de quantidade, temos que na data 2 a quantidade corresponde a 98,49% da quantidade na data 1, o que indica uma redução.

O índice aritmético de Sauerbeck de valor corresponde à divisão da soma dos relativos de valor pelo número total de itens. Os relativos de valor são:

Tabela 7 – Relativos de valor

Item	Valor 1	Valor 2	v_2/v_1
Açúcar	R$ 1.248,00	R$ 1.460,80	1,17
Arroz	R$ 2.255,00	R$ 2.318,00	1,03
Farinha de trigo	R$ 2.066,40	R$ 3.432,00	1,66
Feijão	R$ 2.331,00	R$ 3.360,00	1,44

Assim, o índice aritmético de Sauerbeck de valor é dado por:

$$S_v^A = \frac{\sum_{i=1}^{n} r_v^i}{n}$$

$$S_v^A = \frac{1{,}17 + 1{,}03 + 1{,}66 + 1{,}44}{4}$$

$$S_v^A = \frac{5{,}30}{4}$$

$$S_v^A = 1{,}325188$$

$$S_v^A = 132{,}52\%$$

Com base nesse índice, é possível observar que houve um aumento de 32,52% no valor, pois temos que o valor na data 2 corresponde a 132,52% do valor na data 1.

Os índices de Sauerbeck se baseiam em um mesmo peso, ou seja, em uma mesma importância para cada item. No entanto, em muitas situações, é preciso considerar a importância de cada item no conjunto de dados. Quando isto ocorre, precisamos de índices ponderados.

TEMA 4 – ÍNDICES DE LASPEYRES

O índice de Laspeyres é um índice ponderado, pois leva em consideração a participação de cada item no conjunto de informações. É conhecido também como *índice da época base*, pois os pesos são definidos com relação à data-base.

Para calcularmos o índice de preço de Laspeyres, precisamos multiplicar os preços da data atual pelas quantidades da data-base, somar estes valores e em seguida dividi-los pela soma das multiplicações dos preços pelas quantidades, ambos sendo considerados na data-base, ou seja, precisamos do somatório de $p_2 \cdot q_1$ e do somatório de $p_1 \cdot q_1$. O índice de preço de Laspeyres é dado por:

$$L_p = \frac{\sum_{i=1}^{n}(p_2^i q_1^i)}{\sum_{i=1}^{n}(p_1^i q_1^i)}.$$

O índice de Laspeyres de quantidade é calculado por meio da fórmula:

$$L_q = \frac{\sum_{i=1}^{n}(p_1^i q_2^i)}{\sum_{i=1}^{n}(p_1^i q_1^i)}$$

E o índice de valor é calculado pela fórmula:

$$L_v = \frac{\sum_{i=1}^{n}(p_2^i q_2^i)}{\sum_{i=1}^{n}(p_1^i q_1^i)}.$$

Para calcularmos, então, os índices de Laspeyres de preço, quantidade e valor, precisamos fazer as multiplicações $p_1 . q_1$, $p_1 . q_2$, $p_2 . q_1$ e $p_2 . q_2$ para cada item.

Exemplo:

Uma mercearia está realizando um estudo sobre o impacto na variação do preço do açúcar, do arroz, da farinha de trigo e do feijão. Para isso, está considerando os preços e as respectivas quantidades comercializadas na data 1 e os preços e as quantidades referentes à data 2, 6 meses após a data 1.

Na tabela a seguir, são apresentados os preços e as quantidades na data 1 e na data 2 para cada um dos itens.

Tabela 8 – Preços e quantidades

Item	Preço 1	Quantidade 1	Preço 2	Quantidade 2
Açúcar	R$ 3,90	320	R$ 4,40	332
Arroz	R$ 5,50	410	R$ 6,10	380
Farinha de trigo	R$ 4,20	492	R$ 7,80	440
Feijão	R$ 6,30	370	R$ 8,40	400

Considerando esses dados, calcule os índices de Laspeyres de preço, de quantidade e de valor.

Primeiro, precisamos fazer as seguintes multiplicações: $p_1 . q_1$, $p_1 . q_2$, $p_2 . q_1$ e $p_2 . q_2$:

Tabela 9 – Multiplicações

Preço 1	Quantidade 1	p1 . q1
R$ 3,90	320	R$ 1.248,00
R$ 5,50	410	R$ 2.255,00

R$ 4,20	492	R$ 2.066,40
R$ 6,30	370	R$ 2.331,00

Preço 1	Quantidade 2	p1 . q2
R$ 3,90	332	R$ 1.294,80
R$ 5,50	380	R$ 2.090,00
R$ 4,20	440	R$ 1.848,00
R$ 6,30	400	R$ 2.520,00

Preço 2	Quantidade 1	p2 . q1
R$ 4,40	320	R$ 1.408,00
R$ 6,10	410	R$ 2.501,00
R$ 7,80	492	R$ 3.837,60
R$ 8,40	370	R$ 3.108,00

Preço 2	Quantidade 2	p2 . q2
R$ 4,40	332	R$ 1.460,80
R$ 6,10	380	R$ 2.318,00
R$ 7,80	440	R$ 3.432,00
R$ 8,40	400	R$ 3.360,00

Agora que já temos as multiplicações necessárias, podemos calcular os índices de Laspeyres de preço, de quantidade e de valor.

Índice de Laspeyres de preço:

$$L_p = \frac{\sum_{i=1}^{n}(p_2^i q_1^i)}{\sum_{i=1}^{n}(p_1^i q_1^i)}$$

$$L_p = \frac{1408{,}00 + 2501{,}00 + 3837{,}60 + 3108{,}00}{1248{,}00 + 2255{,}00 + 2066{,}40 + 2331{,}00}$$

$$L_p = \frac{10854{,}60}{7900{,}40}$$

$$L_p = 1{,}373930$$

$$L_p = 137{,}39\%$$

Índice de Laspeyres de quantidade:

$$L_q = \frac{\sum_{i=1}^{n}\left(p_1^i q_2^i\right)}{\sum_{i=1}^{n}\left(p_1^i q_1^i\right)}$$

$$L_q = \frac{1294{,}80 + 2090{,}00 + 1848{,}00 + 2520{,}00}{1248{,}00 + 2255{,}00 + 2066{,}40 + 2331{,}00}$$

$$L_q = \frac{7752{,}80}{7900{,}40}$$

$$L_q = 0{,}981317$$

$$L_q = 98{,}13\%$$

Índice de Laspeyres de valor:

$$L_v = \frac{\sum_{i=1}^{n}\left(p_2^i q_2^i\right)}{\sum_{i=1}^{n}\left(p_1^i q_1^i\right)}$$

$$L_v = \frac{1460{,}80 + 2318{,}00 + 3432{,}00 + 3360{,}00}{1248{,}00 + 2255{,}00 + 2066{,}40 + 2331{,}00}$$

$$L_v = \frac{10570{,}80}{7900{,}40}$$

$$L_v = 1{,}338008$$

$$L_v = 133{,}80\%$$

Podemos concluir que, no conjunto de itens considerados, o preço na data 2 corresponde a 137,39% do preço na data 1, implicando em um aumento de 37,39%. A quantidade na data 2 corresponde a 98,13% da quantidade na data 1. Fazendo 100% - 98,13%, podemos observar que a redução foi de 1,87%. Em relação ao valor, o índice de Laspeyres indica que na data 2 temos 133,80% do valor referente à data 1, o que corresponde a um aumento de 33,80%.

É importante ressaltar que, no índice de preços de Laspeyres, temos $\sum_{i=1}^{n}\left(p_2^i q_1^i\right)$ no numerador, o que corresponde às quantidades da data-base multiplicadas pelos preços na data atual. No denominador temos a expressão $\sum_{i=1}^{n}\left(p_1^i q_1^i\right)$, que corresponde às quantidades e aos preços na data-base, ou seja é o total referente à data-base. Sendo assim, estamos comparando a variação de preços desse conjunto de itens em duas épocas diferentes mantendo as quantidades da data-base. No caso do índice de Laspeyres de quantidade, fixamos os preços e calculamos a variação da quantidade nas duas datas. Por esse motivo, o numerador corresponde a $\sum_{i=1}^{n}(p_1^i q_2^i)$, e o denominador corresponde a $\sum_{i=1}^{n}(p_1^i q_1^i)$. Quanto ao índice de Laspeyres de valor, a medida corresponde à variação tanto da quantidade quanto do preço. Sendo assim, temos a expressão $\sum_{i=1}^{n}(p_2^i q_2^i)$ no numerador e a expressão $\sum_{i=1}^{n}(p_1^i q_1^i)$ no denominador.

TEMA 5 – ÍNDICES DE PAASCHE

No índice de Paasche, conhecido também como *índice da época atual*, calculamos a média harmônica ponderada de relativos. Neste índice, os pesos estão relacionados aos preços e às quantidades que se referem à data atual.

Para calcularmos o índice de Paasche de preço, utilizamos a fórmula:

$$P_p = \frac{\sum_{i=1}^{n}(p_2^i q_2^i)}{\sum_{i=1}^{n}(p_1^i q_2^i)}.$$

Observe que, para o índice de Paasche de preço, estamos considerando a relação entre preços e quantidades na data atual com as quantidades atuais e os preços na data-base. Assim, esse índice de preço fornece a porcentagem de quanto será gasto para adquirirmos as mesmas quantidades da data-base com preços atuais quando comparadas com as quantidades e os preços da data-base.

O índice de Paasche de quantidade é dado por

$$P_q = \frac{\sum_{i=1}^{n}(p_2^i q_2^i)}{\sum_{i=1}^{n}(p_2^i q_1^i)},$$

pois estamos analisando a variação da quantidade em relação aos preços atuais. No numerador temos o valor atual e no denominador, o valor necessário para adquirirmos as quantidades da data anterior.

O índice de Paasche de valor é:

$$P_v = \frac{\sum_{i=1}^{n}(p_2^i q_2^i)}{\sum_{i=1}^{n}(p_1^i q_1^i)}.$$

É um índice destinado à comparação da variação de preço e de quantidade em relação à data 1 e à data 2.

Exemplo:

Uma mercearia está realizando um estudo sobre o impacto na variação do preço do açúcar, do arroz, da farinha de trigo e do feijão. Para isso, está considerando os preços e as respectivas quantidades comercializadas na data 1 e os preços e as quantidades referentes à data 2, 6 meses após a data 1.

Na tabela a seguir, são apresentados os preços e as quantidades na data 1 e na data 2 para cada um dos itens.

Tabela 10 – Preços e quantidades

Item	Preço 1	Quantidade 1	Preço 2	Quantidade 2
Açúcar	R$ 3,90	320	R$ 4,40	332
Arroz	R$ 5,50	410	R$ 6,10	380
Farinha de trigo	R$ 4,20	492	R$ 7,80	440
Feijão	R$ 6,30	370	R$ 8,40	400

Considerando esses dados, calcule os índices de Paasche de preço, de quantidade e de valor.

Multiplicando $p_1 . q_1$, $p_1 . q_2$, $p_2 . q_1$ e $p_2 . q_2$, temos:

Preço 1	Quantidade 1	p1 . q1
R$ 3,90	320	R$ 1.248,00
R$ 5,50	410	R$ 2.255,00
R$ 4,20	492	R$ 2.066,40
R$ 6,30	370	R$ 2.331,00

Preço 1	Quantidade 2	p1 . q2
R$ 3,90	332	R$ 1.294,80
R$ 5,50	380	R$ 2.090,00
R$ 4,20	440	R$ 1.848,00
R$ 6,30	400	R$ 2.520,00

Preço 2	Quantidade 1	p2 . q1
R$ 4,40	320	R$ 1.408,00
R$ 6,10	410	R$ 2.501,00
R$ 7,80	492	R$ 3.837,60

R$ 8,40	370	R$ 3.108,00

Preço 2	Quantidade 2	p2 . q2
R$ 4,40	332	R$ 1.460,80
R$ 6,10	380	R$ 2.318,00
R$ 7,80	440	R$ 3.432,00
R$ 8,40	400	R$ 3.360,00

Agora que já efetuamos as multiplicações, podemos calcular os índices de Paasche de preço, de quantidade e de valor.

Índice de Paasche de preço:

$$P_p = \frac{\sum_{i=1}^{n}(p_2^i q_2^i)}{\sum_{i=1}^{n}(p_1^i q_2^i)}$$

$$P_p = \frac{1460,80 + 2318,00 + 3432,00 + 3360,00}{1294,80 + 2090,00 + 1848,00 + 2520,00}$$

$$P_p = \frac{10570,80}{7752,80}$$

$$P_p = 1,363481$$

$$P_p = 136,35\%$$

Índice de Paasche de quantidade:

$$P_q = \frac{\sum_{i=1}^{n}(p_2^i q_2^i)}{\sum_{i=1}^{n}(p_2^i q_1^i)}$$

$$P_q = \frac{1460,80 + 2318,00 + 3432,00 + 3360,00}{1408,00 + 2501,00 + 3837,60 + 3108,00}$$

$$P_q = \frac{10570,80}{10854,60}$$

$$P_q = 0,973854$$

$$P_q = 97,39\%$$

Índice de Paasche de valor:

$$P_v = \frac{\sum_{i=1}^{n}(p_2^i q_2^i)}{\sum_{i=1}^{n}(p_1^i q_1^i)}$$

$$P_v = \frac{1460{,}80 + 2318{,}00 + 3432{,}00 + 3360{,}00}{1248{,}00 + 2255{,}00 + 2066{,}40 + 2331{,}00}$$

$$P_v = \frac{10570{,}80}{7900{,}40}$$

$$P_v = 1{,}338008$$

$$P_v = 133{,}80$$

Com base nos índices de Paasche, temos que o preço na data 2 corresponde a 136,35% do preço na data 1, o que corresponde a um aumento de 36,35%. A quantidade na data 2 é 97,39% da quantidade na data 1, uma redução de 2,61%. Quanto ao valor, na data 2 temos 133,80% do valor referente à data 1, o que corresponde a um aumento no valor associado aos itens de 33,80%.

FINALIZANDO

Chegamos ao final desta etapa. Estudamos relativos e aprendemos que é possível utilizar relativos para compararmos variações entre dois números referentes a um determinado item. Podemos considerar uma variação de preço, de quantidade ou de valor. Vimos que, no caso de realizarmos um estudo contendo um conjunto formado por dois ou mais itens, é necessário utilizar índices agregativos. Inicialmente estudamos o índice de Bradstreet, que considera a divisão entre as somas dos preços, das quantidades ou dos valores em duas datas distintas. Vimos também os índices de Sauerbeck, obtidos por meio dos relativos de preço, de quantidade ou de valor. Aprendemos que quando precisamos de índices ponderados, os índices que levam em consideração a importância de cada item no conjunto, é possível utilizarmos os índices de Laspeyres ou de Paasche. Os índices de Laspeyres são conhecidos como *índices da época base*, e os de Paasche, como índices da época atual. Por meio desses índices, é possível estudarmos variações que ocorrem em preços, quantidades ou valores.

REFERÊNCIAS

CASTANHERIA, N. P. **Matemática aplicada**. 3. ed. Curitiba: Ibpex, 2010.

DEMANA, F. D. et al. **Pré-cálculo**. 2. ed. São Paulo: Pearson, 2013.

FLEMING, D. M.; GONÇALVES, M. B.; **Cálculo A**: função de uma variável. 2. ed. São Paulo: Pearson, 2007.

MÉTODOS QUANTITATIVOS

AULA 4

Prof. Ricardo Alexandre Deckmann Zanardini

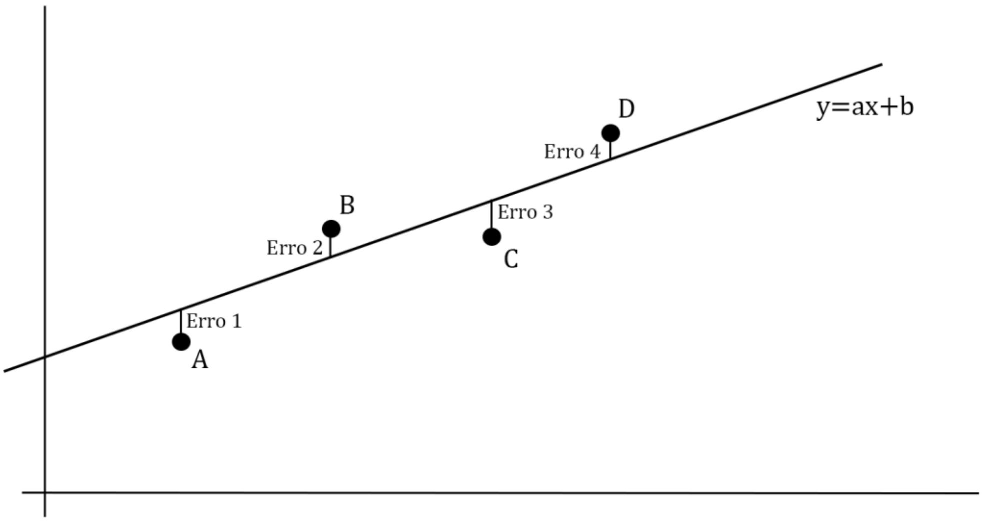
CONVERSA INICIAL

Estudaremos, nesta etapa, a regressão linear, um assunto muito importante para a resolução de diversos problemas práticos. Aprenderemos a construir um diagrama de dispersão, obter a reta de regressão e calcular o coeficiente de correlação de Pearson, que indica o grau de correlação dos dados.

TEMA 1 – REGRESSÃO LINEAR

A regressão linear está associada à obtenção de uma função linear $y = ax + b$ que melhor se ajusta a um conjunto de pontos dados. O critério de obtenção dessa função consiste em minimizar os desvios quadráticos entre os pontos e a função linear. Sendo assim, queremos uma função que melhor se aproxima dos pontos sem, necessariamente, passar pelos pontos dados.

Figura 1 – Regressão linear



Fonte: Zanardini, [S.d.].

A partir da regressão linear, podemos realizar estudos relacionados às vendas, à produção, ao consumo de diversas fontes de energia, aos tempos relacionados à execução de uma determinada tarefa, entre muitas outras aplicações importantes.

Podemos obter a reta de regressão por meio de uma fórmula destinada ao coeficiente angular **a** e por meio de uma fórmula destinada ao coeficiente linear **b**.

Também é possível utilizarmos o Python para obtermos a reta de regressão relacionada a um conjunto de pontos.

Por meio de um conjunto de pontos e, consequentemente, da reta de regressão, é possível fazer uma análise do comportamento dos dados. Podemos identificar se há um crescimento ou decrescimento, além de que podemos saber qual é a respectiva taxa de crescimento ou de decrescimento. Com base nas informações, é possível também fazer algumas estimativas ou previsões. É claro que toda previsão está sujeita a erros, mas é uma forma de obtermos informações consistentes que servem de auxílio no processo de tomada de decisões.

Antes de pensarmos em como podemos obter a reta de regressão, aprenderemos a fazer a representação gráfica do conjunto de pontos, o que é chamado de diagrama de dispersão. Aprenderemos também a calcular o coeficiente de correlação de Pearson com base em um conjunto de dados. Esse coeficiente indica o nível de correlação entre os dados e também informa se a correlação é positiva ou negativa. Assim, podemos saber, com base nesse coeficiente, se a reta de regressão é crescente ou decrescente.

A seguir, veremos o que é o diagrama de dispersão e como podemos construir esses diagramas.

TEMA 2 – DIAGRAMA DE DISPERSÃO

Uma forma muito útil de representarmos um conjunto de dados na forma de pares ordenados (x, y) é o uso de um diagrama de dispersão.

Por meio desse diagrama, podemos entender de forma visual a relação entre as variáveis do problema.

Vamos utilizar um exemplo para compreendermos melhor o que é um diagrama de dispersão e como podemos construir um por meio do Python.

2.1 Exemplo

A produção de automóveis em um determinado país, nos seis primeiros meses do ano, é apresentada na Tabela 1 a seguir. Com base nesses dados, construa o diagrama de dispersão respectivo.

Tabela 1 – Produção de automóveis nos seis primeiros meses do ano

Janeiro	Fevereiro	Março	Abril	Maio	Junho
182 mil	174 mil	193 mil	207 mil	218 mil	197 mil

Fonte: Zanardini, [S.d.].

O diagrama de dispersão é formado por pontos da forma (x, y). O primeiro passo é identificarmos as variáveis independentes (x) e as variáveis dependentes (y). Como as produções dependem de cada um dos meses, estes são as variáveis independentes (x), e as respectivas produções mensais são as variáveis dependentes (y).

Para construirmos o diagrama de dispersão por meio do Python, inicialmente vamos importar as bibliotecas *matplotlib* e *numpy*. Em seguida, vamos definir um *array* para os valores associados a **x** (meses) e um outro *array* para os valores associados a **y** (produções mensais). Apenas para a criação do diagrama de dispersão não é necessário usar o comando *array* para **x** e **y**, mas futuramente aprenderemos a obter a reta de regressão por meio do Python e, nesse caso, será necessário criarmos um *array* para **x** e um *array* para **y**. Assim, já vamos nos ambientando com a forma.

O *array* referente a **x** é:

x=np.array([1, 2, 3, 4, 5, 6])

Já o *array* referente a **y** é:

y=np.array([182, 174, 193, 207, 218, 197])

Para termos a representação de pontos isolados e não de linhas, precisamos acrescentar o parâmetro "'o'" ao comando *plot*:

plt.plot(x, y, 'o').

Para termos um gráfico que mantenha as devidas proporções dos dados em relação ao eixo y, precisamos definir os limites desse eixo. O limite inferior é 0 e, o limite superior, um valor acima do maior valor associado a **y**. Escolhemos o valor 250 para o limite superior do eixo y, mas poderíamos ter escolhido um outro valor qualquer maior ou igual a 218, que é o maior valor referente a **y**. O comando para definir os limites do eixo y é:

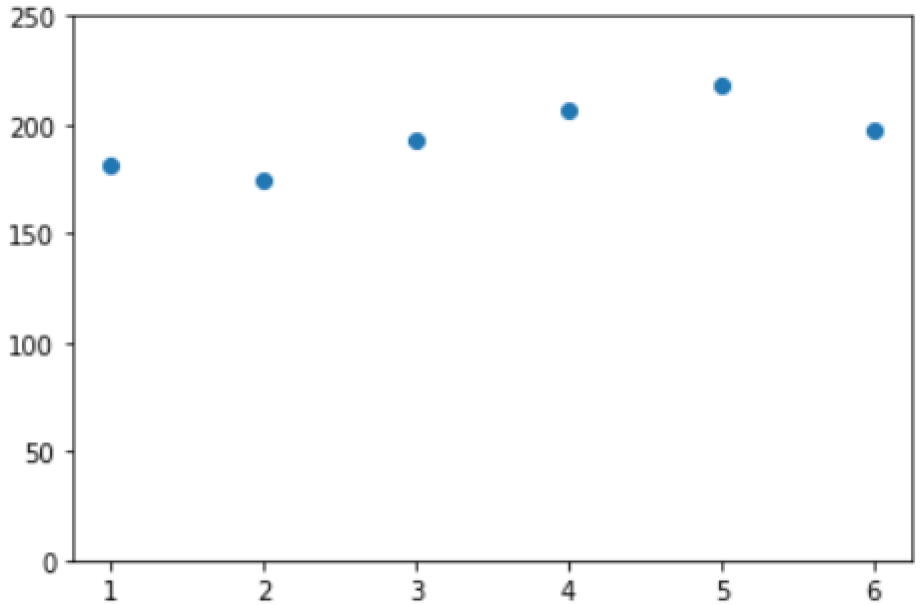
plt.ylim(0, 250)

Assim, toda a sequência de comandos destinados à construção do diagrama de dispersão é:

- import matplotlib.pyplot as plt
- import numpy as np
- x=np.array([1, 2, 3, 4, 5, 6])
- y=np.array([182, 174, 193, 207, 218, 197])
- plt.plot(x, y, 'o')
- plt.ylim(0, 250)
- plt.show()

Após aplicar esses comandos, obtemos o diagrama de dispersão disposto na Figura 2 a seguir.

Figura 2 – Diagrama de dispersão



Fonte: Zanardini, [S.d.].

O diagrama de dispersão facilita a compreensão do comportamento da produção em função do tempo. Também é possível observarmos que os pontos não estão alinhados. No entanto, nesse caso, podemos considerar uma linha reta que melhor se aproxima dos pontos, a reta de regressão.

TEMA 3 – RETA DE REGRESSÃO

A reta de regressão é um importante instrumento para o estudo de diversos fenômenos. Com base em um conjunto de pontos da forma (x, y), caso haja correlação linear entre os dados, podemos obter uma reta que melhor se aproxima desses pontos. Essa reta é chamada de “reta de regressão”. Os coeficientes **a** e **b** da reta de regressão y = ax + b são dados, respectivamente, pelas seguintes equações.

$$a = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

e

$$b = \frac{\sum y}{n} - a\frac{\sum x}{n}$$

Como podemos obter a respectiva reta com base em um conjunto de pontos? Para compreendermos melhor, vamos resolver um exemplo de forma detalhada.

3.1 Exemplo

A produção de automóveis em um determinado país, nos seis primeiros meses do ano, é apresentada na Tabela 2 a seguir. Com base nesses dados, obtenha a reta de regressão respectiva e faça uma estimativa para a produção no mês de julho.

Tabela 2 – Produção de automóveis nos seis primeiros meses do ano

Janeiro	**Fevereiro**	**Março**	**Abril**	**Maio**	**Junho**
182 mil	174 mil	193 mil	207 mil	218 mil	197 mil

Fonte: Zanardini, [S.d.].

Sabemos que a equação reduzida da reta de regressão é da forma y = ax + b, em que:

$$a = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

e

$$b = \frac{\sum y}{n} - a\frac{\sum x}{n}$$

Podemos perceber que, para obtermos os coeficientes da reta, precisamos do somatório de **x**, do somatório de **y**, do somatório do produto de **x** por **y**, do somatório de **x²** e do quadrado do somatório de **x**, além do valor de **n**, que é o número total de pontos do problema.

Podemos organizar esses valores em uma tabela composta pelos valores de **x**, de **y**, de **xy** e de **x²**, e com seus respectivos somatórios (Tabela 3).

Tabela 3 – Somatórios

x	y	xy	x^2
1	182	182	1
2	174	348	4
3	193	579	9
4	207	828	16
5	218	1090	25
6	197	1182	36
21	**1171**	**4209**	**91**

Fonte: Zanardini, [S.d.].

Temos, então, n = 6, $\sum x = 21$, $\sum y = 1171$, $\sum xy = 4209$, $\sum x^2 = 91$ e $(\sum x)^2 = 441$.

Vamos, inicialmente, calcular o valor de **a**:

$$a = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{(6)(4209) - (21)(1171)}{(6)(91) - (21)^2}$$

$$a = \frac{25254 - 24591}{546 - 441}$$

$$a = \frac{663}{105}$$

$$a = 6{,}314285714$$

$$a = 6{,}314$$

Para obtermos o valor de **b**, basta fazermos:

$$b = \frac{\sum y}{n} - a\frac{\sum x}{n}$$

$$b = \frac{1171}{6} - 6{,}314\frac{21}{6}$$

$$b = 195{,}167 - 6{,}314\,.\,3{,}5$$

$$b = 195{,}167 - 22{,}099$$

$$b = 173{,}068$$

Podemos concluir que a equação da reta de regressão é dada por:

y = 6,314x + 173,068

É importante ressaltar que os coeficientes **a** e **b** estão em milhares, ou seja, 6,314 corresponde a 6.314 unidades e 173,068 corresponde a 173.068 unidades produzidas.

Com base na equação da reta de regressão, podemos ter importantes informações. Uma delas está associada ao coeficiente angular **a**. Como a = 6,314, por mais que os níveis mensais de produção oscilem, temos um crescimento mensal médio de 6.314 unidades.

Supondo que a tendência se mantenha, podemos fazer uma estimativa para a produção referente ao sétimo mês. É claro que previsões não vão necessariamente acontecer, mas fornecem uma noção do que se pode esperar.

Para essa estimativa, basta substituirmos **x** por 7 na equação y = 6,314 x + 173,068, pois é o número associado ao sétimo mês.

Fazendo isto, temos:

y = 6,314x + 173,068

y = 6,314(7) + 173,068

y = 44,198 + 173,068

y = 217,266

Logo, a previsão de produção para o sétimo mês é de 217.266 unidades.

Podemos utilizar o Python para obtermos a reta de regressão. Em particular, utilizaremos o pacote *stats* da biblioteca *scipy*. Para importarmos o pacote, utilizaremos a seguinte sintaxe:

from scipy import stats

Em seguida, precisaremos importar a biblioteca *numpy* e chamá-la de **np**:

import numpy as np

Vamos definir **x** e **y** como sendo *arrays* contendo os meses na forma numérica (x) e as respectivas produções (y):

x=np.array([1, 2, 3, 4, 5, 6])

y=np.array([182, 174, 193, 207, 218, 197])

Para obtermos os termos relacionados à reta de regressão, basta escrevermos o seguinte:

a,b,c,d,e=stats.linregress(x,y)

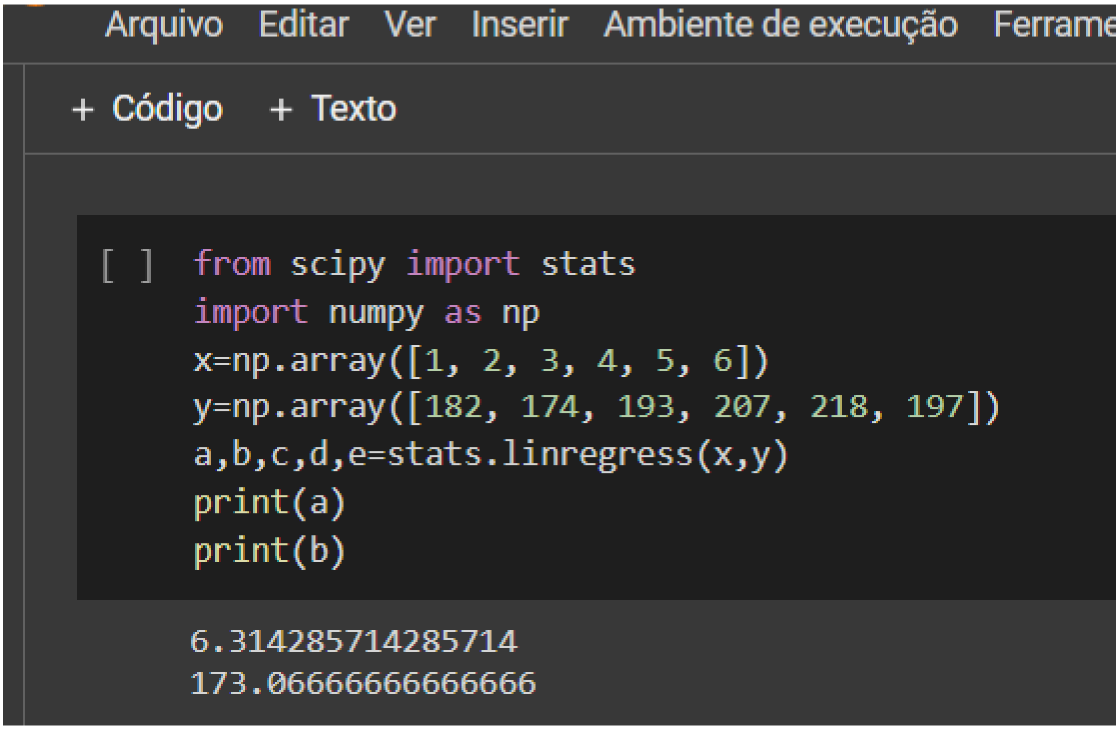
O comando *linregress* informa o valor de cinco parâmetros: **a**, **b**, **c**, **d** e **e**. Nestes, **a** é o coeficiente angular da reta, **b** é o coeficiente linear, **c** é o coeficiente de correlação de Pearson – que será estudado a seguir –, **d** é um parâmetro e **e** é o erro. A nós, interessam os valores de **a** e de **b**. Assim, precisamos apresentar esses valores fazendo print(a) e print(b).

A sequência total de comandos é:

- from scipy import stats
- import numpy as np
- x=np.array([1, 2, 3, 4, 5, 6])
- y=np.array([182, 174, 193, 207, 218, 197])
- a,b,c,d,e=stats.linregress(x,y)
- print(a)
- print(b)

O resultado obtido está disposto na Figura 3 a seguir.

Figura 3 – Resultado



Fonte: Zanardini, [S.d.].

Assim, por meio do Python e considerando três casas decimais, obtivemos a = 6,314 e b = 173,067.

Também é possível utilizarmos o comando *print* de forma mais elaborada para escrevermos a equação da reta de regressão e não apenas os valores de **a** e de **b**.

Utilizaremos uma *f-string*, uma forma simples de escrevermos frases e adicionarmos variáveis a elas.

A sintaxe é a seguinte:

```
print(f'Reta de regressão: y={a:.3f}x+{b:.3f}')
```

Nela, os valores de **a** e de **b** são apresentados nas posições em que temos {a:.3f} e {b:.3f}, respectivamente. Para que sejam apresentadas três casas decimais, foram adicionados os termos “:.3f” após **a** e **b**. Para apresentarmos a equação com duas casas decimais, por exemplo, basta adicionarmos “:.2f” após **a** e após **b**.

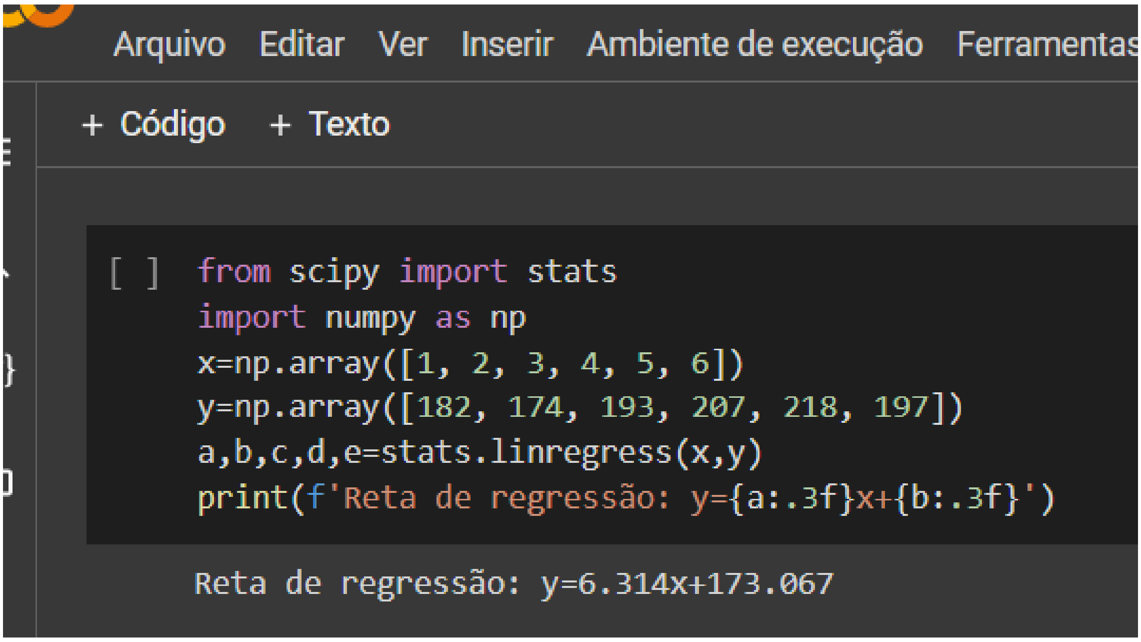
A sequência total de comandos agora é:

- from scipy import stats
- import numpy as np
- x=np.array([1, 2, 3, 4, 5, 6])

- y=np.array([182, 174, 193, 207, 218, 197])
- a,b,c,d,e=stats.linregress(x,y)
- print(f'Reta de regressão: y={a:.3f}x+{b:.3f}')

O resultado obtido está disposto na Figura 4 a seguir.

Figura 4 – Resultado



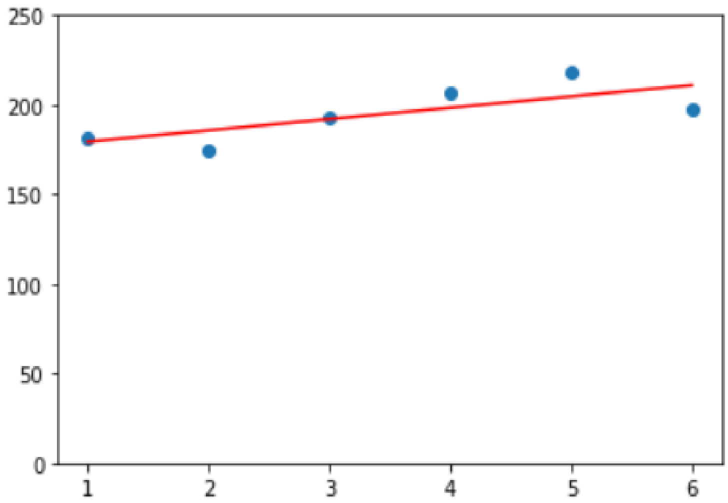
Fonte: Zanardini, [S.d.].

O uso de uma *f-string* é opcional, mas o resultado fica melhor de ser visualizado e é obtido de forma bastante simples.

Também é possível representar a reta de regressão obtida juntamente com os dados originais do problema. Basta utilizarmos os seguintes comandos:

- from scipy import stats
- import matplotlib.pyplot as plt
- import numpy as np
- x=np.array([1, 2, 3, 4, 5, 6])
- y=np.array([182, 174, 193, 207, 218, 197])
- a,b,c,d,e=stats.linregress(x,y)
- plt.plot(x, y, 'o')
- f=a*x+b
- plt.plot(x, f, 'r')
- plt.ylim(0, 250)
- plt.show()

Figura 5 – Reta de regressão



Fonte: Zanardini, [S.d.].

Utilizamos duas vezes o comando *plot*. O primeiro está relacionado aos dados originais e o segundo à reta de regressão. Para obtermos os respectivos valores de **y** referentes à reta, calculamos f = a * x + b e plotamos os respectivos dados por meio do comando plt.plot(x, f, 'r'), em que o parâmetro **‘r’** foi utilizado para a cor vermelha (*red*) da reta.

Aprendemos, até então, como devemos proceder para obtermos a reta de regressão por meio de um conjunto de pontos. Agora já podemos também saber o grau de proximidade dos pontos em relação à reta, ou seja, se a reta de regressão está bastante próxima ou menos próxima dos pontos por meio do coeficiente de correlação de Pearson.

TEMA 4 – COEFICIENTE DE CORRELAÇÃO DE PEARSON

Para um conjunto de pontos dados, nem sempre é uma reta que melhor se ajusta aos pontos. Mesmo que a melhor aproximação seja feita por meio de uma reta, é necessário saber se ela está suficientemente próxima ou não do conjunto de pontos. Precisamos, então, de um parâmetro para sabermos se há ou não correlação linear entre os dados e qual é o nível de proximidade dos pontos em relação à reta.

Pensando nisso, o coeficiente de correlação de Pearson é importante para podermos fazer uma análise mais criteriosa.

Podemos, então, denotar o coeficiente de Pearson **c**. O coeficiente de correlação de Pearson é obtido por meio da seguinte fórmula:

$$c = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

É importante observar que o numerador e alguns termos do denominador da fórmula são obtidos de forma equivalente quando calculamos o coeficiente angular da reta de regressão.

A variação do coeficiente de correlação de Pearson é de –1 a 1. Quanto mais próximo de 1, mais forte é a correlação entre os pontos e a função é crescente. Se o coeficiente está mais próximo de –1, a correlação também é forte, mas a função é decrescente. Quando c = 1 ou c = –1, temos que a reta passa exatamente sobre os pontos dados.

No entanto, quanto mais próximo de 0 estiver o coeficiente de correlação de Pearson, mais fraca é a correlação. Se c = 0, então não há uma correlação linear entre os dados do problema.

Figura 6 – Variação do coeficiente de correlação de Pearson

Fonte: Zanardini, [S.d.].

Em muitas obras, o coeficiente de correlação de Pearson é representado pela letra **r**.

Mas como podemos calcular o coeficiente de Pearson em um problema?

A forma é muito parecida com o procedimento que realizamos para obter o coeficiente angular da reta de regressão.

4.1 Exemplo

A produção de automóveis em um determinado país, nos seis primeiros meses do ano, é apresentada na Tabela 4 a seguir. Com base nesses dados, obtenha o respectivo coeficiente de correlação de Pearson.

Tabela 4 – Produção de automóveis nos seis primeiros meses do ano

Janeiro	Fevereiro	Março	Abril	Maio	Junho
182 mil	174 mil	193 mil	207 mil	218 mil	197 mil

Fonte: Zanardini, [S.d.].

Para resolvermos o problema, vamos utilizar a seguinte fórmula:

$$c = \frac{n \sum xy - \sum x \sum y}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}}$$

Além do somatório de **x**, do somatório de **y**, do somatório do produto de **x** por **y**, do somatório de **x²**, do quadrado do somatório de **x** e do valor de **n**, precisamos agora do somatório de **y²** e do quadrado do somatório de **y**.

Assim, temos a Tabela 5 a seguir.

Tabela 5 – Somatórios

x	y	xy	x^2	y^2
1	182	182	1	33124
2	174	348	4	30276
3	193	579	9	37249
4	207	828	16	42849
5	218	1090	25	47524
6	197	1182	36	38809
21	**1171**	**4209**	**91**	**229831**

Fonte: Zanardini, [S.d.].

Com base nos dados da tabela, n = 6, $\sum x = 21$, $\sum y = 1171$, $\sum xy = 4209$, $\sum x^2 = 91$, $(\sum x)^2 = 441$, $\sum y^2 = 229831$ e $(\sum y)^2 = 1371241$.

Basta, agora, substituirmos os valores na seguinte fórmula:

$$c = \frac{n \sum xy - \sum x \sum y}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}}$$

Após, devemos realizar os respectivos cálculos:

$$c = \frac{n \sum xy - \sum x \sum y}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}}$$

$$c = \frac{(6)(4209) - (21)(1171)}{\sqrt{((6)(91) - (21)^2)((6)(229831) - (1171)^2)}}$$

$$c = \frac{25254 - 24591}{\sqrt{(546 - 441)(1378986 - 1371241)}}$$

$$c = \frac{663}{\sqrt{(105)(7745)}}$$

$$c = \frac{663}{\sqrt{813225}}$$

$$c = \frac{663}{901{,}7898868}$$

$$c = 0{,}7352045$$

$$c = 73{,}52\%$$

Podemos concluir que o coeficiente de correlação de Pearson corresponde a 0,7352045. É um valor positivo e está entre 0,5 e 1. Nesse caso, podemos dizer que a reta que melhor se aproxima dos pontos é crescente e que a correlação é relativamente forte. Notemos que, na forma de porcentagem, podemos dizer que a correlação é de 73,52%. Para isso, basta multiplicarmos 0,7352045 por 100.

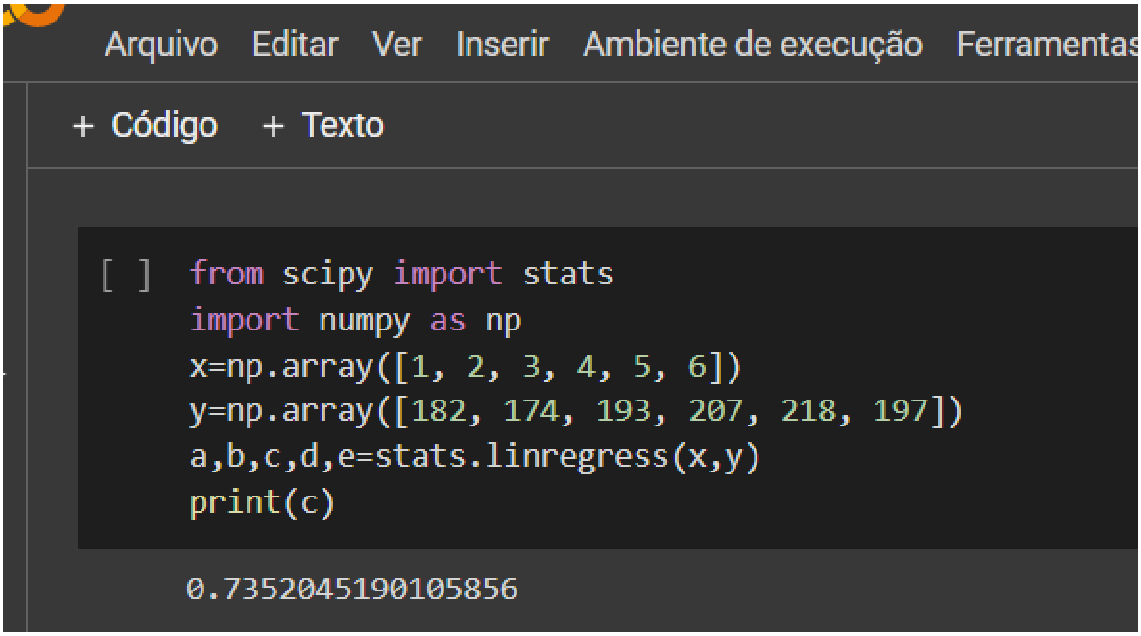
Podemos obter o coeficiente de correlação de Pearson utilizando a função *linregress* desenvolvida para o Python. O procedimento é bastante similar ao que foi feito para a obtenção da reta de regressão.

A sequência de comandos é:

- from scipy import stats
- import numpy as np
- x=np.array([1, 2, 3, 4, 5, 6])
- y=np.array([182, 174, 193, 207, 218, 197])
- a,b,c,d,e=stats.linregress(x,y)
- print(c)

O resultado obtido está disposto na Figura 7 a seguir.

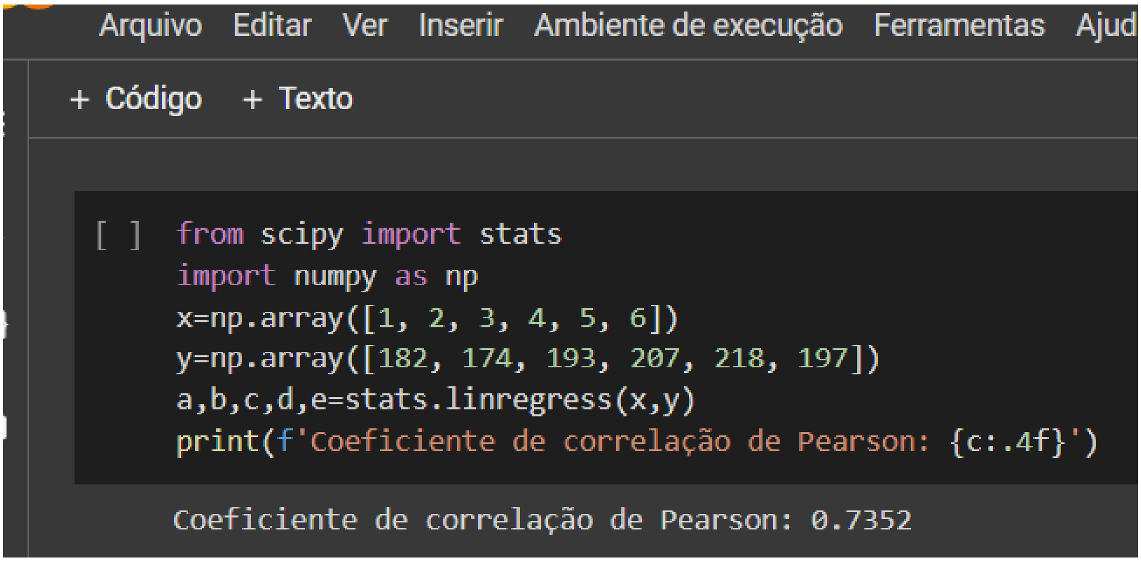
Figura 7 – Resultado



Fonte: Zanardini, [S.d.].

Para uma melhor visualização do resultado, podemos substituir o comando print(c) por print(f'Coeficiente de correlação de Pearson: {c:.4f}'). O resultado obtido está disposto na Figura 8 a seguir.

Figura 8 – Resultado



Fonte: Zanardini, [S.d.].

Na sequência, teremos mais exemplos de situações em que é possível utilizarmos a regressão linear e o coeficiente de correlação de Pearson.

TEMA 5 – APLICAÇÕES

Vamos abordar alguns exemplos de aplicações relacionadas à regressão linear e ao coeficiente de correlação de Pearson.

5.1 Exemplo 1

O consumo de energia elétrica de uma determinada residência varia mensalmente. A seguir, na Tabela 6, estão dispostos os dados sobre o consumo ao longo do ano. Com base nos valores apresentados, faça o diagrama de dispersão, obtenha a reta de regressão e o coeficiente de correlação de Pearson e identifique se durante o ano ocorreu aumento ou redução no consumo de energia elétrica na residência em questão. Faça também o gráfico da reta de regressão juntamente com os pontos referentes aos dados do problema.

Tabela 6 – Consumo ao longo do ano

Mês	**Consumo (kWh)**
Janeiro	212
Fevereiro	182
Março	207
Abril	211
Maio	191
Junho	204
Julho	201
Agosto	189
Setembro	181
Outubro	220
Novembro	207
Dezembro	199

Fonte: Zanardini, [S.d.].

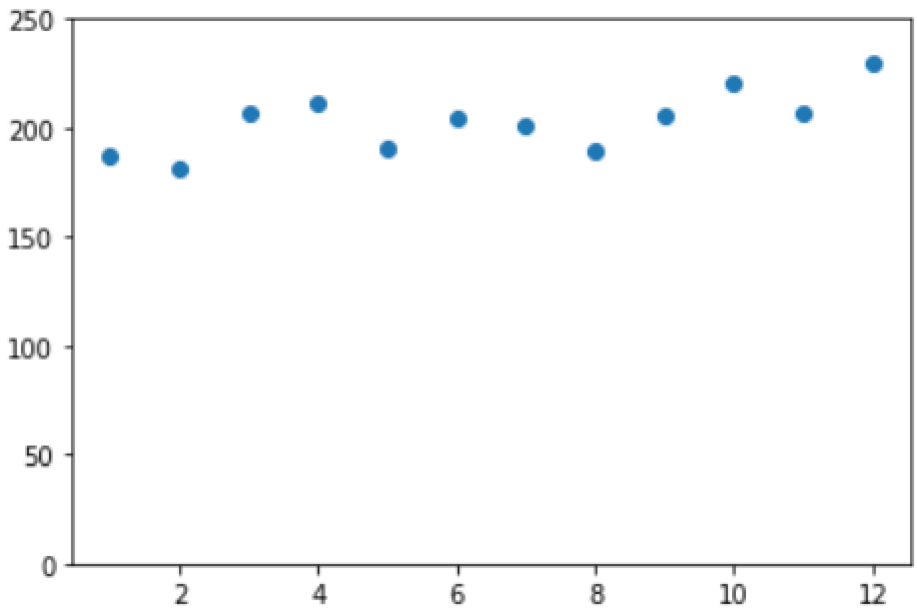
Para o diagrama de dispersão, temos os seguintes comandos:

- import matplotlib.pyplot as plt
- import numpy as np

- x=np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12])
- y=np.array([187, 182, 207, 211, 191, 204, 201, 189, 205, 220, 207, 230])
- plt.plot(x, y, 'o')
- plt.ylim(0, 250)
- plt.show()

Isso gera o seguinte diagrama, disposto na Figura 9 a seguir.

Figura 9 – Diagrama



Fonte: Zanardini, [S.d.].

Os comandos utilizados para a obtenção da reta de regressão e do coeficiente de correlação de Pearson são:

- from scipy import stats
- import numpy as np
- x=np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12])
- y=np.array([187, 182, 207, 211, 191, 204, 201, 189, 205, 220, 207, 230])
- a,b,c,d,e=stats.linregress(x,y)
- print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')
- print(f'Coeficiente de correlação de Pearson: {c:.4f}')

Figura 10 – Resultado

```python
from scipy import stats
import numpy as np
x=np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12])
y=np.array([187, 182, 207, 211, 191, 204, 201, 189, 205, 220, 207, 230])
a,b,c,d,e=stats.linregress(x,y)
print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')
print(f'Coeficiente de correlação de Pearson: {c:.4f}')
```

```
Reta de regressão: y=2.62x+185.79
Coeficiente de correlação de Pearson: 0.6748
```

Fonte: Zanardini, [S.d.].

Podemos concluir que a equação reduzida da reta de regressão associada ao problema é a seguinte:

$$y = 2{,}62x + 185{,}79$$

Ainda, o coeficiente de correlação de Pearson é igual a 0,6748, o que corresponde a 67,48%.

Os resultados obtidos indicam que, no decorrer do ano, houve um aumento no consumo de energia elétrica na residência. A taxa de crescimento é de 2,62 kWh.

5.2 Exemplo 2

A variação de custo, em reais, de um determinado artigo está associada à respectiva quantidade produzida, conforme pode ser observado na Tabela 7 a seguir. Obtenha a reta que melhor se ajusta ao seguinte conjunto de pontos e o respectivo coeficiente de correlação de Pearson. Por meio da reta de regressão, faça uma estimativa para o custo referente à produção de 210 unidades.

Tabela 7 – Variação de custo, em reais

Quantidade (x)	100	120	140	160	180	200
Custo (y)	1.010	1.080	1.112	1.160	1.200	1.245

Fonte: Zanardini, [S.d.].

Os comandos necessários são:

- from scipy import stats
- import numpy as np

- x=np.array([100, 120, 140, 160, 180, 200])
- y=np.array([1010, 1080, 1112, 1160, 1200, 1245])
- a,b,c,d,e=stats.linregress(x,y)
- print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')
- print(f'Coeficiente de correlação de Pearson: {c:.4f}')

Assim, a reta de regressão é dada por y = 2,26x + 795,29 e o coeficiente de correlação de Person é igual a 0,9954, o que equivale a 99,54%.

Figura 11 – Resultado

```
from scipy import stats
import numpy as np
x=np.array([100, 120, 140, 160, 180, 200])
y=np.array([1010, 1080, 1112, 1160, 1200, 1245])
a,b,c,d,e=stats.linregress(x,y)
print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')
print(f'Coeficiente de correlação de Pearson: {c:.4f}')
```

```
Reta de regressão: y=2.26x+795.29
Coeficiente de correlação de Pearson: 0.9954
```

Fonte: Zanardini, [S.d.].

Para termos uma estimativa de produção de 210 unidades, basta multiplicarmos 210 por **a** e somarmos o resultado com o valor de **b**:

$$y = 2{,}26x + 795{,}29$$

$$y = 2{,}26(210) + 795{,}29$$

$$y = 474{,}6 + 795{,}29$$

$$y = 1269{,}89$$

A estimativa é de que o custo para a produção de 210 unidades seja de aproximadamente R$ 1.269,89.

5.3 Exemplo 3

O consumo de combustível de um determinado automóvel varia de acordo com a velocidade com a qual o automóvel se desloca. A relação entre consumo e velocidade é apresentada na Tabela 8 a seguir. Qual é a reta que melhor se ajusta aos dados do problema? Faça a respectiva representação gráfica.

Tabela 8 – Relação entre consumo e velocidade

Velocidade (km/h)	Consumo (km/l)
50	12,5
60	11,9
70	11,4
80	10,7
90	9,9

Fonte: Zanardini, [S.d.].

Para obtermos a equação reduzida da reta de regressão por meio do Python, basta fazermos o seguinte:

- from scipy import stats
- import numpy as np
- x=np.array([50, 60, 70, 80, 90])
- y=np.array([12.5, 11.9, 11.4, 10.7, 9.9])
- a,b,c,d,e=stats.linregress(x,y)
- print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')

Logo, a equação da reta de regressão é dada por:

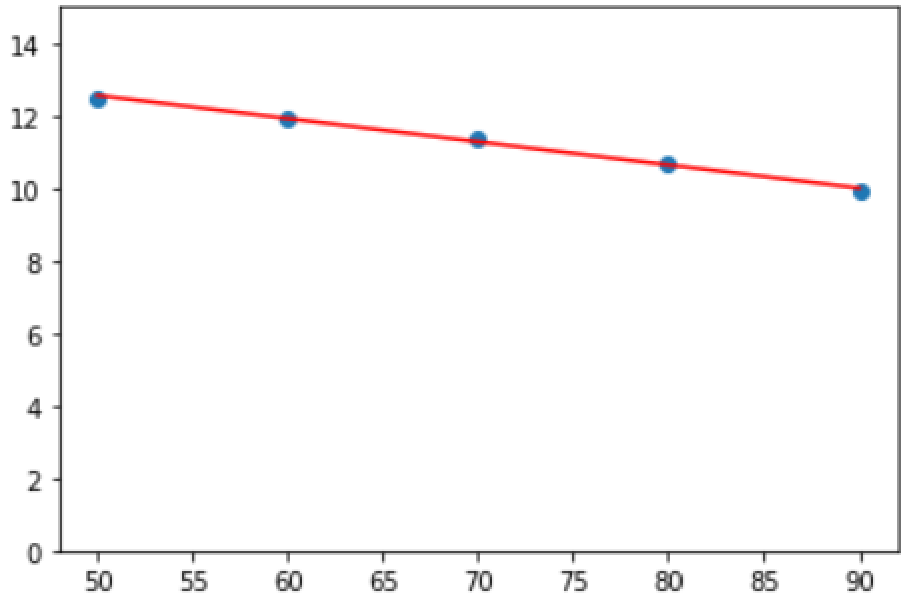
$$Y = -0,06x + 15,76$$

Para obtermos o gráfico, basta fazermos:

- from scipy import stats
- import numpy as np
- x=np.array([50, 60, 70, 80, 90])
- y=np.array([12.5, 11.9, 11.4, 10.7, 9.9])
- a,b,c,d,e=stats.linregress(x,y)
- plt.plot(x, y, 'o')
- f=a*x+b
- plt.plot(x, f, 'r')
- plt.ylim(0, 15)

- plt.show()

Figura 11 – Reta de regressão



Fonte: Zanardini, [S.d.].

Nesse caso, quanto maior a velocidade, menor é a distância percorrida a cada litro de combustível.

FINALIZANDO

Nesta etapa, aprendemos o que é diagrama de dispersão, regressão linear e coeficiente de correlação de Pearson. Aprendemos também a obter os resultados analiticamente e também por meio do Python. Por fim, tivemos contato com algumas de suas muitas aplicações.

REFERÊNCIAS

CASTANHERIA, N. P. **Matemática aplicada**. 3 ed. Curitiba: IBPEX, 2010.

DEMANA, F. D. et al. **Pré-cálculo**. 2. ed. São Paulo: Pearson, 2013.

FLEMING, D. M.; GONÇALVES, M. B. **Cálculo A**: função de uma variável. 2. ed. São Paulo: Pearson, 2007.

MÉTODOS QUANTITATIVOS

AULA 5

Prof. Ricardo Alexandre Deckmann Zanardini

CONVERSA INICIAL

Nesta etapa, abordaremos importantes aspectos relacionados à interpolação polinomial. Utilizando os conhecimentos associados à interpolação, podemos modelar diversos problemas e também elaborar algumas estimativas. Aprenderemos a obter um polinômio interpolador com base em um conjunto de pontos e estudaremos diversas aplicações.

TEMA 1 – POLINÔMIOS

Como nosso objetivo é tratar da interpolação polinomial, é importante falarmos um pouco a respeito dos polinômios.

Um polinômio é uma expressão que tem a forma

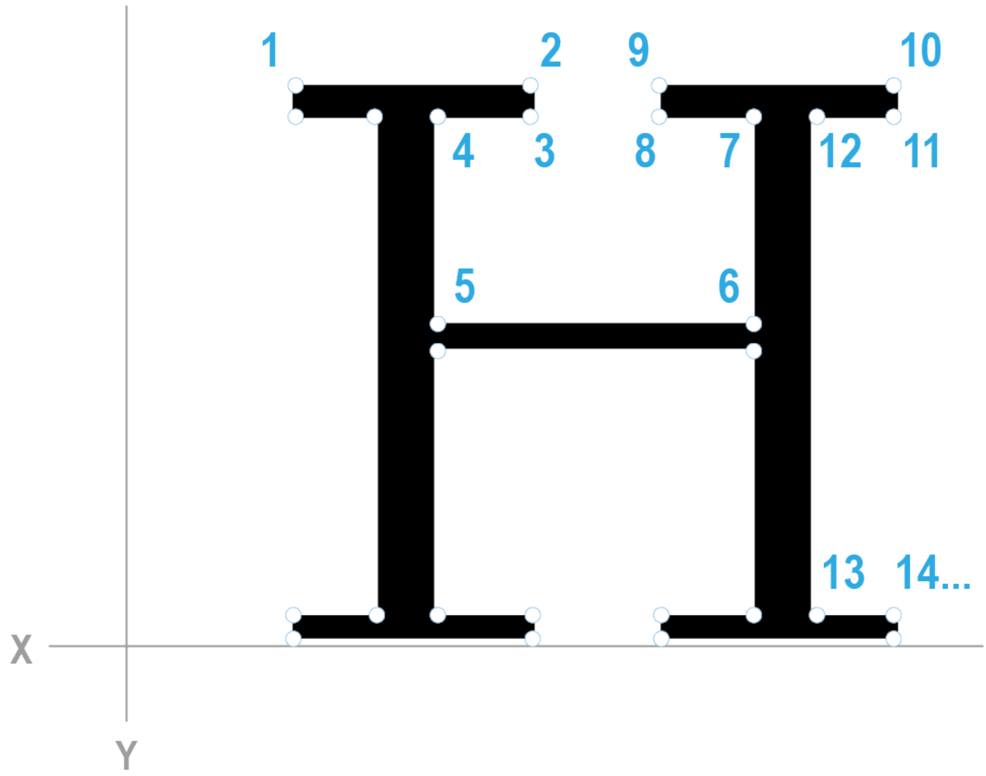
$$p(x)=a_nx^n+a_{n-1}x^{n-1}+...+a_2x^2+a_1x+a_0$$

onde a_n, a_{n-1}, a_2, a_1, a_0 são os coeficientes do polinômio. Para que possamos garantir que o polinômio tem grau n, temos a condição de que a_n seja diferente de zero.

Há diversas situações onde os polinômios estão associados à modelagem e à resolução de problemas.

Podemos, por exemplo, gerar imagens por meio de polinômios, as quais são obtidas a partir de pontos específicos e chamadas de *imagens vetoriais*. As fontes que vemos na tela do celular, *tablet* ou do computador, por exemplo, são geradas por polinômios interpoladores.

Figura 1 – Exemplo de imagem vetorial



Crédito: Tarannum Ansari/Shutterstock.

É um processo que ocupa pouco espaço na memória e facilita, além da construção da fonte, da obtenção de diferentes possibilidades de tamanho ou de espessura.

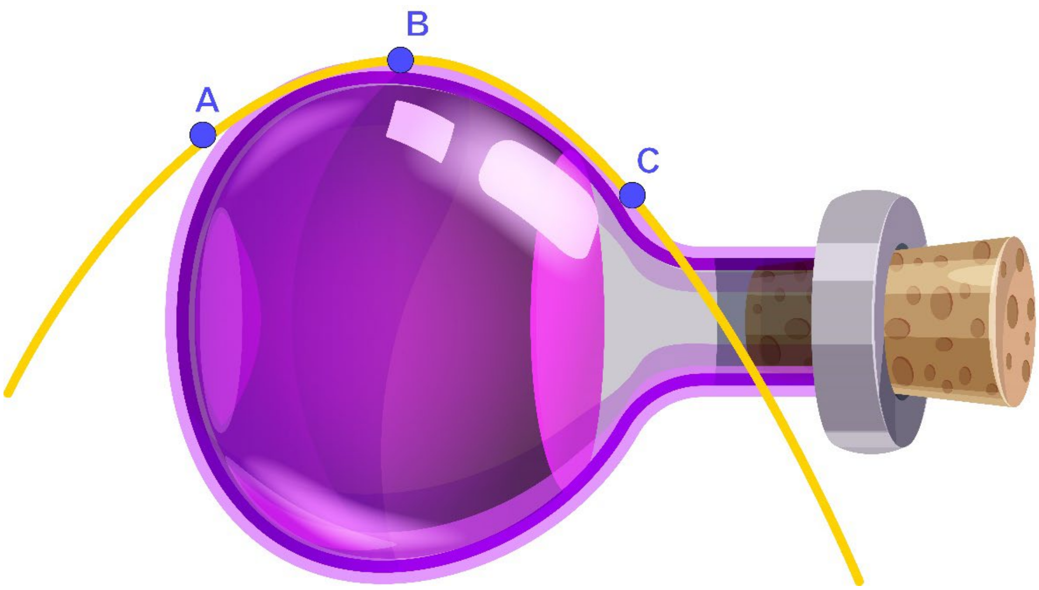
Polinômios podem ser utilizados na modelagem de fenômenos físicos tais como o movimento de um objeto sendo lançado ao ar ou a trajetória relacionada a um salto.

Figura 2 – Trajetória relacionada a um salto

Crédito: Wyssu/Adobe Stock.

Outra aplicação importante está relacionada à estimativa do volume de embalagens que possuem diferentes formas. Observando o projeto de um frasco, por exemplo, é possível estimar a capacidade da embalagem por meio das técnicas adequadas e de um polinômio que melhor se ajusta ao formato do frasco.

Figura 3 – Formato de um frasco



Crédito: Vector Up/Shutterstock.

Também é possível utilizarmos polinômios para relacionarmos a variação do lucro da venda de um determinado produto com a respectiva variação de preço, dentre muitas outras aplicações. A seguir, aprenderemos algumas formas de obtenção de um polinômio que interpola um conjunto de pontos.

TEMA 2 – INTERPOLAÇÃO POLINOMIAL

Quando temos um conjunto finito de pontos, podemos encontrar uma função f(x) que passa por todos eles. Essa função é chamada de *função interpoladora*. Dentre diversos tipos de funções interpoladoras, as funções polinomiais são muito comuns e podem ser úteis na resolução de diversos problemas.

Dados $x_0, x_1, \ldots, x_n$ números distintos e $y_0, y_1, \ldots, y_n$ os respectivos valores funcionais associados a eles, o polinômio $p(x_i) = y_i$, $i = 0, 1, \ldots, n$ que interpola estes pontos pode ser encontrado.

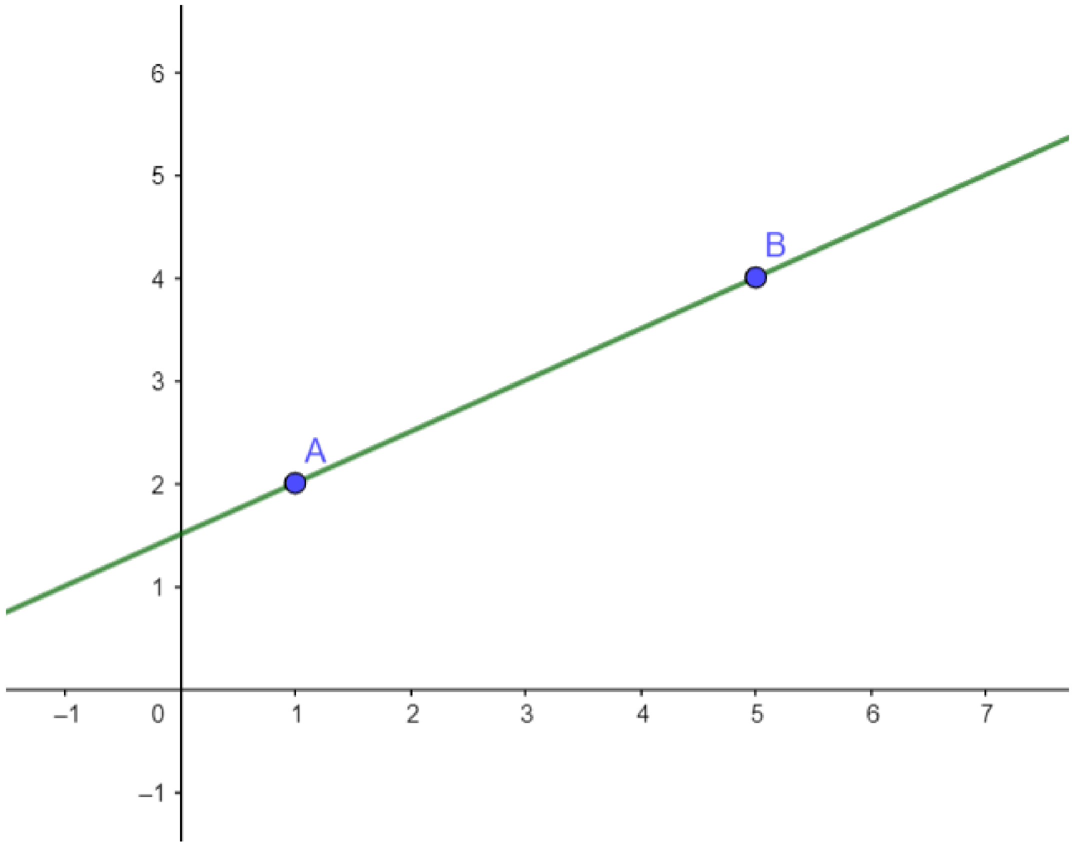
Basicamente, podemos escrever

$$p(x_i) = a_0 + a_1x + \ldots + a_nx^n$$

para um polinômio de grau n com $p(x_i) = y_i$.

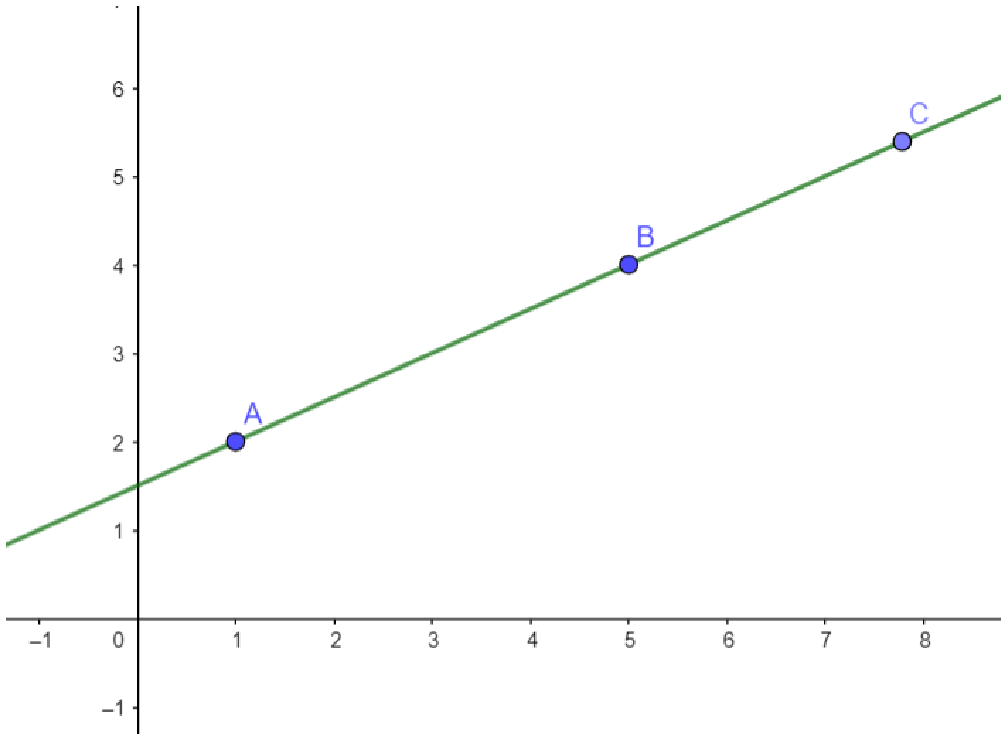
Assim, se temos n+1 pontos, o polinômio interpolador tem, no máximo, grau n. Por exemplo, para dois pontos distintos, temos um polinômio de grau 1 que corresponde a uma reta.

Figura 4 – Polinômio de grau 1 que corresponde a uma reta



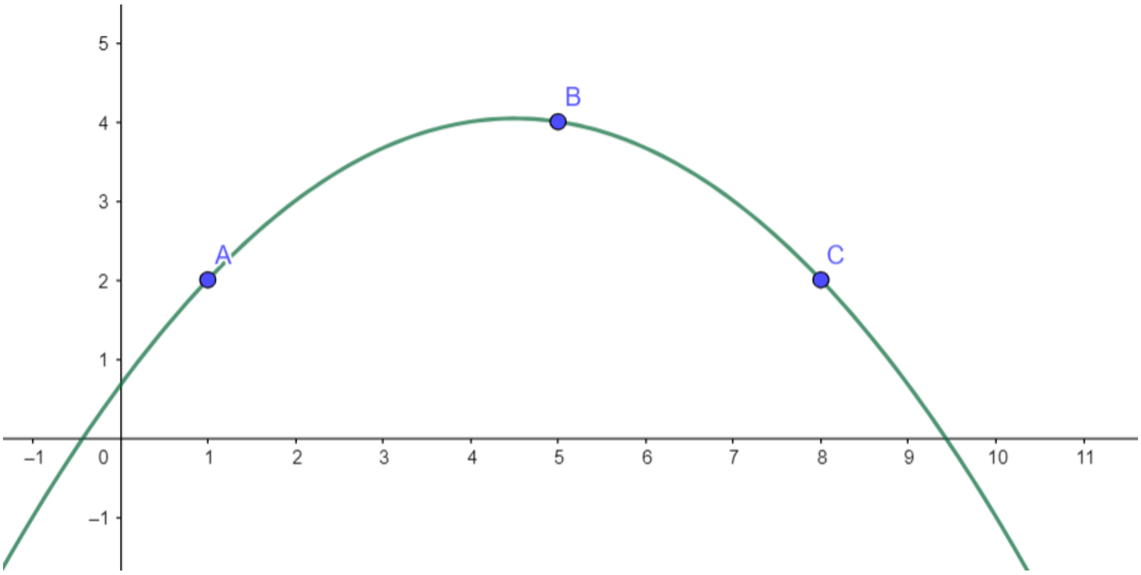
Para três pontos distintos, podemos ter uma reta interpolando os pontos caso eles estejam alinhados ou uma função de grau 2.

Figura 5 – Reta interpolando os pontos



ou

Figura 6 – Função de grau 2



Se tivermos quatro pontos, o polinômio interpolador tem grau menor ou igual a 3 e assim por diante.

Substituindo cada ponto (x_0, y_0), (x_1, y_1), …, (x_n, y_n) na expressão

$$a_0 + a_1x + \ldots + a_nx^n = y_i,$$

temos o seguinte sistema de equações lineares

$$a_0 + a_1x_0 + a_2x_0^2 + \ldots a_nx_0^n = y_0$$

$$a_0 + a_1x_1 + a_2x_1^2 + \ldots a_nx_1^n = y_1$$

$$a_0 + a_1x_2 + a_2x_2^2 + \ldots a_nx_2^n = y_2$$

.

.

.

$$a_0 + a_1x_n + a_2x_n^2 + \ldots a_nx_n^n = y_n$$

Além da resolução do sistema linear, há vários métodos destinados à obtenção de um polinômio que interpola um conjunto de pontos.

Antes de aprendermos a forma de obtermos um polinômio interpolador por meio da resolução de um sistema linear, vamos estudar o método de

Lagrange, uma forma simples e eficiente utilizada na resolução de problemas de interpolação onde não é preciso resolver um sistema de equações lineares.

Pelo método de Lagrange, podemos escrever

$$p_n(x) = \sum_{i=0}^{n} y_i l_i(x)$$

onde

$$l_i(x) = \prod_{j=0, j\neq i}^{n} \left(\frac{x - x_j}{x_i - x_j}\right)$$

O termo Π indica o produtório, ou seja, para cada i=0, 1, …, n, temos as respectivas multiplicações com j variando de 0 a n, sendo que j não pode ser igual a i.

A obtenção de um polinômio interpolador de Lagrange é bastante simples e veremos a seguir um exemplo para ajudar na compreensão.

Exemplo:

Dados os pontos A(1, 7), B(3, 23) e C(5, 47), obtenha o polinômio de Lagrange que interpola estes pontos.

Nesse caso, temos

$x_0=1$, $y_0=7$, $x_1=3$, $y_1=23$, $x_2=5$ e $y_2=47$.

Como são dados três pontos, o polinômio interpolador tem grau menor ou igual a 2.

Logo

$$p_n(x) = \sum_{i=0}^{2} y_i l_i(x)$$

com

$$l_i(x) = \prod_{j=0, j\neq i}^{2} \left(\frac{x - x_j}{x_i - x_j}\right)$$

Para obtermos $p_n(x)$, o primeiro passo é calcularmos cada um dos termos

$$l_i(x) = \prod_{j=0, j\neq i}^{2} \left(\frac{x - x_j}{x_i - x_j}\right)$$

com i=0, i=1 e i=2.

Para i=0, temos

$$l_0(x) = \prod_{j=0, j\neq 0}^{2} \left(\frac{x - x_j}{x_i - x_j}\right)$$

Neste caso, i=0 e j será igual a 1 e igual a 2, o que resulta em

$$l_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}$$

$$l_0(x) = \frac{(x - 3)(x - 5)}{(1 - 3)(1 - 5)}$$

$$l_0(x) = \frac{x^2 - 5x - 3x + 15}{(-2)(-4)}$$

$$l_0(x) = \frac{x^2 - 8x + 15}{8}$$

Para i=1, temos

$$l_1(x) = \prod_{j=0, j\neq 1}^{2} \left(\frac{x - x_j}{x_i - x_j}\right)$$

Neste caso, i=1 e j será igual a 0 e igual a 2, o que resulta em

$$l_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}$$

$$l_1(x) = \frac{(x - 1)(x - 5)}{(3 - 1)(3 - 5)}$$

$$l_1(x) = \frac{x^2 - 5x - x + 5}{(2)(-2)}$$

$$l_1(x) = \frac{x^2 - 6x + 5}{-4}$$

Para i=2, temos

$$l_2(x) = \prod_{j=0, j\neq 2}^{2} \left(\frac{x - x_j}{x_i - x_j}\right)$$

Neste caso, i=2 e j será igual a 0 e igual a 1, o que resulta em

$$l_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}$$

$$l_2(x) = \frac{(x - 1)(x - 3)}{(5 - 1)(5 - 3)}$$

$$l_2(x) = \frac{x^2 - 3x - x + 3}{(4)(2)}$$

$$l_2(x) = \frac{x^2 - 4x + 3}{8}$$

Agora que temos os termos l_0, l_1 e l_2, podemos fazer o somatório

$$p_n(x) = \sum_{i=0}^{2} y_i l_i(x)$$

lembrando que y_0=7, y_1=23 e y_2=47. Assim, temos

$$p_n(x) = 7.\frac{x^2 - 8x + 15}{8} + 23.\frac{x^2 - 6x + 5}{-4} + 47.\frac{x^2 - 4x + 3}{8}$$

$$p_n(x) = \frac{7x^2 - 56x + 105}{8} - \frac{23x^2 - 138x + 115}{4} + \frac{47x^2 - 188x + 141}{8}$$

$$p_n(x) = \frac{7x^2 - 56x + 105}{8} - \frac{46x^2 - 276x + 230}{8} + \frac{47x^2 - 188x + 141}{8}$$

$$p_n(x) = \frac{7x^2 - 56x + 105 - (46x^2 - 276x + 230) + 47x^2 - 188x + 141}{8}$$

$$p_n(x) = \frac{7x^2 - 56x + 105 - 46x^2 + 276x - 230 + 47x^2 - 188x + 141}{8}$$

$$p_n(x) = \frac{8x^2 + 32x + 16}{8}$$

$$p_n(x) = x^2 + 4x + 2$$

Podemos concluir que o polinômio de Lagrange que interpola os pontos A(1, 7), B(3, 23) e C(5, 47) corresponde a

$$p_n(x) = x^2 + 4x + 2.$$

Quando temos poucos pontos, é possível resolvermos o sistema de equações para obtermos o polinômio interpolador.

TEMA 3 – RESOLUÇÃO DE SISTEMA LINEAR

Quando temos dois pontos distintos, o polinômio interpolador tem grau 1, e, quando temos três pontos, o polinômio interpolador tem grau menor ou igual a 2.

Nesses casos, é simples obtermos o polinômio interpolador por meio da resolução de um sistema linear associado ao problema.

Exemplo:

Uma empresa produz cabos HDMI e precisa obter a função que relaciona o lucro mensal com o preço de venda de um determinado modelo de cabo. Os custos fixos associados a esse produto correspondem a R$ 1.200,00 por mês. Quando o cabo é comercializado por R$ 20,00, o lucro mensal corresponde a R$ 600,00. Para o preço de R$ 22,00, o lucro mensal é de R$ 648,00. Obtenha o polinômio quadrático que relaciona o lucro y com o preço de venda x.

Como os custos fixos associados a esse cabo correspondem a R$ 1.200,00, para x=0, temos y=-1200.

Logo, o primeiro ponto é A(0, -1200).

Se o preço é de R$ 20,00, o lucro é de R$ 600,00 e se o preço é de R$ 22,00, o lucro é de R$ 648,00.

Assim, temos os pontos

B(20, 600) e C(22, 648).

Vamos substituir cada um dos pontos na expressão $y=ax^2+bx+c$ para encontrarmos os coeficientes a, b e c.

Para o ponto A(0, -1200), temos:

$$y=ax^2+bx+c$$

$$-1200=a(0)^2+b(0)+c$$

$$-1200=0+0+c$$

$$-1200=c$$

$$c=-1200$$

Para o ponto B(20, 600), temos:

$$y=ax^2+bx+c$$

$$600=a(20)^2+b(20)+(-1200)$$

$$600=400a+20b-1200$$

$$600+1200=400a+20b$$

$$1800=400a+20b$$

$$400a+20b=1800$$

Para o ponto C(22, 648), temos:

$$y=ax^2+bx+c$$

$$648=a(22)^2+b(22)+(-1200)$$

$$648=484a+22b-1200$$

$$648+1200=484a+22b$$

$$1848=484a+22b$$

$$484a+22b=1848$$

Precisamos resolver agora o sistema de equações

$$400a+20b=1800$$

$$484a+22b=1848$$

para obtermos os valores de “a” e de “b”.

Para simplificarmos, podemos dividir a primeira equação por 20 e a segunda equação por 22, o que resulta em

$$20a+b=90$$

$$22a+b=84$$

Multiplicando a segunda equação por -1, temos:

$$20a+b=90$$

$$-22a-b=-84$$

Somando termo a termo, temos:

$$20a-22a=-2a,\ b-b=0 \text{ e } 90-84=6$$

Logo

$$-2a=6$$

$$a=6/(-2)$$

$$a=-3$$

Vamos substituir o valor de “a” na equação 20a+b=90 para calcularmos o valor de “b”:

$$20a+b=90$$

$$20(-3)+b=90$$

$$-60+b=90$$

$$b=90+60$$

$$b=150$$

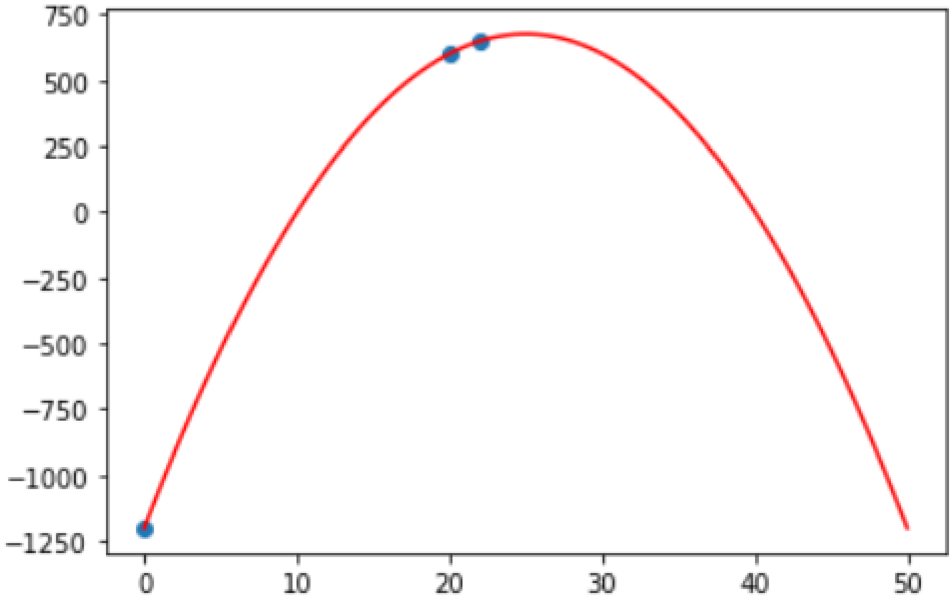
Logo, a=-3, b=150 e c=-1200

Portanto, o polinômio quadrático que relaciona o lucro com o preço é

$$y=-3x^2+150x-1200.$$

Graficamente, temos:

Figura 7 – Gráfico de um polinômio quadrático que relaciona o lucro com o preço



Os comandos utilizados para a construção do gráfico foram:

```
import matplotlib.pyplot as plt

import numpy as np

x1=[0, 20, 22]

x2=np.linspace(0, 50, 100)

y1=[-1200, 600, 648]

y2=-3*x2**2+150*x2-1200

plt.plot(x1, y1, 'o')

plt.plot(x2, y2, 'r')

plt.show()
```

Podemos utilizar o Python para obtermos de uma forma bastante simples um polinômio que interpola um conjunto de dados.

TEMA 4 – USO DO PYTHON NA INTERPOLAÇÃO POLINOMIAL

Vimos que é possível obtermos analiticamente polinômios que interpolam conjuntos de pontos. Também é possível utilizarmos a função “lagrange()” da biblioteca “ScyPy” para obtermos um polinômio que interpola um conjunto de pontos.

A sequência de comandos é:

```
from scipy.interpolate import *

x=[]

y=[]

p=lagrange(x, y)

print(p)
```

bastando apenas informar os valores de x e os respectivos valores de y relacionados ao problema a ser resolvido.

Exemplo:

Uma empresa produz cabos HDMI e precisa obter a função que relaciona o lucro mensal com o preço de venda de um determinado modelo de cabo. Os custos fixos associados a esse produto correspondem a R$ 1.200,00 por mês. Quando o cabo é comercializado por R$ 20,00, o lucro mensal corresponde a R$ 600,00. Para o preço de R$ 22,00, o lucro mensal é de R$ 648,00. Obtenha o polinômio quadrático que relaciona o lucro y com o preço de venda x.

Como os custos fixos associados a esse cabo correspondem a R$ 1.200,00, para x=0, temos y=-1200. Se o preço é de R$ 20,00, o lucro é de R$ 600,00 e se o preço é de R$ 22,00, o lucro é de R$ 648,00.

Logo, precisamos interpolar os pontos A(0, -1200), B(20, 600) e C(22, 648).

Vamos digitar em uma célula de código do Google Colab os seguintes comandos:

```
from scipy.interpolate import *

x=[0, 20, 22]

y=[-1200, 600, 648]

p=lagrange(x, y)

print(p)
```

Fazendo isso, temos:

```
[ ]  from scipy.interpolate import *
     x=[0, 20, 22]
     y=[-1200, 600, 648]
     p=lagrange(x, y)
     print(p)

        2
     -3 x + 150 x - 1200
```

Podemos concluir que o polinômio quadrático que relaciona o lucro mensal com o preço do determinado modelo de cabo HDMI é

$$y=-3x^2+150x-1200.$$

Veremos a seguir alguns dos diversos problemas práticos em que a interpolação é utilizada.

TEMA 5 – APLICAÇÕES

A seguir, alguns exemplos relacionados ao uso da interpolação polinomial em problemas práticos.

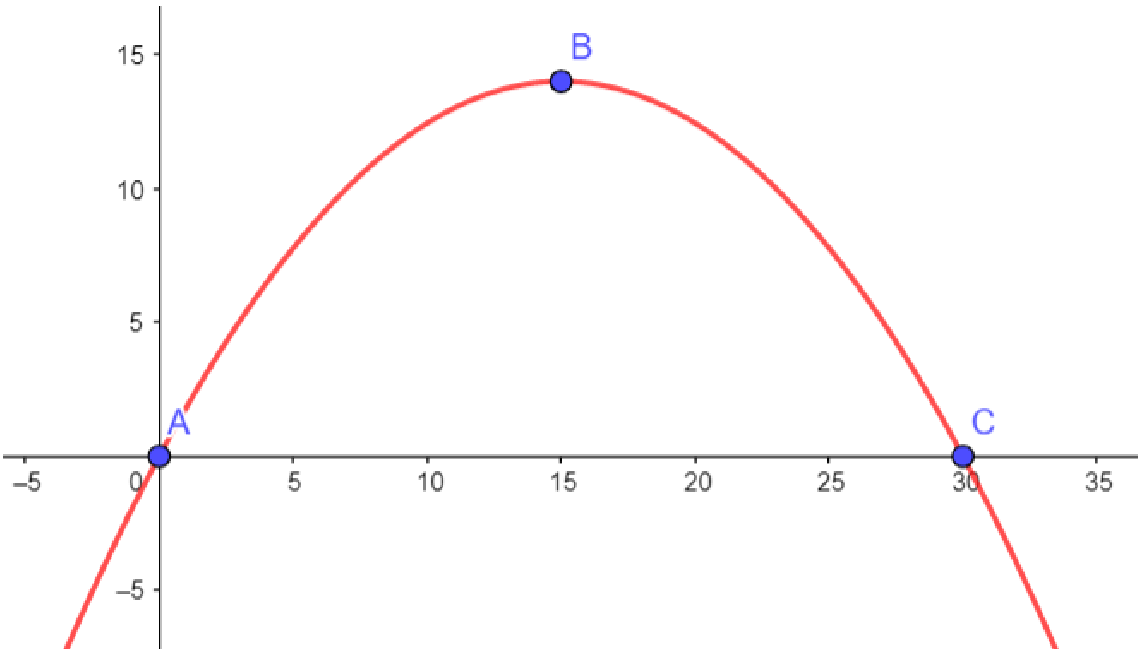
Exemplo 1:

A estrutura de uma determinada ponte está baseada em uma parábola. Sabendo que as raízes da parábola estão nos pontos A(0, 0) e C(30, 0) e que o vértice da parábola está no ponto B(15, 14), qual é o polinômio que representa essa estrutura?

Figura 8 – Estrutura de ponte baseada em uma parábola

Crédito: Tshooter/Shutterstock

Precisamos do polinômio que interpola os pontos A(0, 0), B(15, 14) e C(30, 0).



Assim, por meio do Python, temos os seguintes comandos:

from scipy.interpolate import *

x=[0, 15, 30]

y=[0, 14, 0]

p=lagrange(x, y)

print(p)

```
[ ]  from scipy.interpolate import *
     x=[0, 15, 30]
     y=[0, 14, 0]
     p=lagrange(x, y)
     print(p)

               2
     -0.06222 x + 1.867 x
```

O polinômio associado à estrutura da ponte é $p(x)=-0{,}06222x^2+1{,}867x$. Com base nesse polinômio, é possível obtermos, por exemplo, a altura da estrutura em algum determinado ponto.

Exemplo 2:

Uma indústria produz camisetas de algodão. Quando cada camiseta é vendida por R$ 12,00, o lucro mensal é de R$ 13.980,00. Quando cada camiseta é vendida por R$ 15,00, o lucro mensal é de R$ 19.350,00 e quando cada camiseta é vendida por R$ 19,00, o lucro mensal é de R$ 18.770,00. Obtenha o polinômio quadrático que relaciona o preço de venda de cada camiseta com o lucro mensal e em seguida determine o preço que maximiza o lucro mensal da indústria.

Utilizando-se dos dados do problema, precisamos interpolar os pontos A(12, 13980), B(15, 19350) e C(19, 18770).

Para encontrarmos o polinômio por meio do Python, basta fazermos:

from scipy.interpolate import *

x=[12, 15, 19]

y=[13980, 19350, 18770]

p=lagrange(x, y)

print(p)

O resultado obtido é

```
[ ]  from scipy.interpolate import *
     x=[12, 15, 19]
     y=[13980, 19350, 18770]
     p=lagrange(x, y)
     print(p)

            2
     -276.4 x + 9254 x - 5.726e+04
```

Logo, o polinômio que relaciona o lucro mensal com o preço de venda de cada camiseta é

$$p(x)=-276{,}4x^2+9254x-57260.$$

Note que o termo -5,726e+04 corresponde a $-5{,}726 \cdot 10^4$ que é igual a -57260.

Para sabermos o preço que maximiza o lucro, basta considerarmos o x do vértice da função quadrática, ou seja, basta substituirmos a=-276,4 e b=9254 na fórmula

$$x_v = \frac{-b}{2a},$$

o que resulta em

$$x_v = \frac{-b}{2a}$$

$$x_v = \frac{-9254}{2(-276{,}4)}$$

$$x_v = \frac{-9254}{-552{,}8}$$

$$x_v = 16{,}74$$

Logo, o lucro máximo da indústria ocorre quando cada camiseta é comercializada por R$ 16,74.

Exemplo 3:

A concentração média de partículas por milhão (ppm) de um certo poluente na atmosfera está se alterando com o passar do tempo.

Na tabela a seguir, são apresentados os dados obtidos no decorrer de 5 anos de estudos.

Ano	1	2	3	4	5
Concentração	320	380	395	410	450

Qual é o polinômio que interpola estes dados?

Considerando os valores 1, 2, 3, 4 e 5 para x e os respectivos valores 320, 380, 395, 410 e 450 para y, temos:

```
from scipy.interpolate import *

x=[1, 2, 3, 4, 5]

y=[320, 380, 395, 410, 450]

p=lagrange(x, y)

print(p)
```

```
[ ]  from scipy.interpolate import *
     x=[1, 2, 3, 4, 5]
     y=[320, 380, 395, 410, 450]
     p=lagrange(x, y)
     print(p)

             4         3         2
-0.8333 x + 15.83 x - 96.67 x + 251.7 x + 150
```

O polinômio interpolador, neste caso, é

$$p(x)=-0,833x^4+15,83x^3-96,67x^2+251,7x+150.$$

FINALIZANDO

No decorrer da etapa, aprendemos o que são polinômios e como podemos resolver problemas práticos utilizando-os. Aprendemos a forma da interpolação de Lagrange, a interpolação por meio da resolução de sistemas lineares e também por meio do Python. Finalizamos estudando algumas das diversas aplicações reais relacionadas à interpolação polinomial.

REFERÊNCIAS

CASTANHEIRA, N. P. **Matemática aplicada**. 3. ed. Curitiba: Ibpex, 2010.

DEMANA, F. D. et al. **Pré-cálculo.** 2. ed. São Paulo: Pearson, 2013.

FLEMING, D. M.; GONÇALVES, M. B. **Cálculo A**: função de uma variável. 2. ed. São Paulo: Pearson, 2007.

MÉTODOS QUANTITATIVOS

AULA 6

Prof. Ricardo Alexandre Deckmann Zanardini

CONVERSA INICIAL

Nosso objetivo agora é tratarmos das séries temporais. Aprenderemos conceitos relacionados às séries temporais, bem como a importância e a utilização dos conhecimentos em situações reais. Nesta etapa, estudaremos também as médias móveis, assim como aplicações relacionadas ao assunto.

TEMA 1 – SÉRIES TEMPORAIS

Quando nos referimos a séries temporais, estamos tratando de uma distribuição estatística, em que a principal característica é a de que a variável em questão está organizada em função de determinados períodos de tempo.

As séries temporais são comumente utilizadas em problemas associados ao processamento de sinais de áudio, vídeo, imagens, comunicação sem fio, sistemas de controle, problemas relacionados à matemática financeira, matemática aplicada, estatística, marketing e muito mais.

Quando um determinado conjunto de dados consiste em uma série temporal, é importante manter a ordem de observação dos dados, pois eles precisam estar associados às respectivas datas de ocorrência.

É importante ressaltar que, no caso de séries temporais, não é possível organizar os dados em ordem crescente ou decrescente, pois nosso objetivo é verificar variações e tendências desses dados em relação ao tempo.

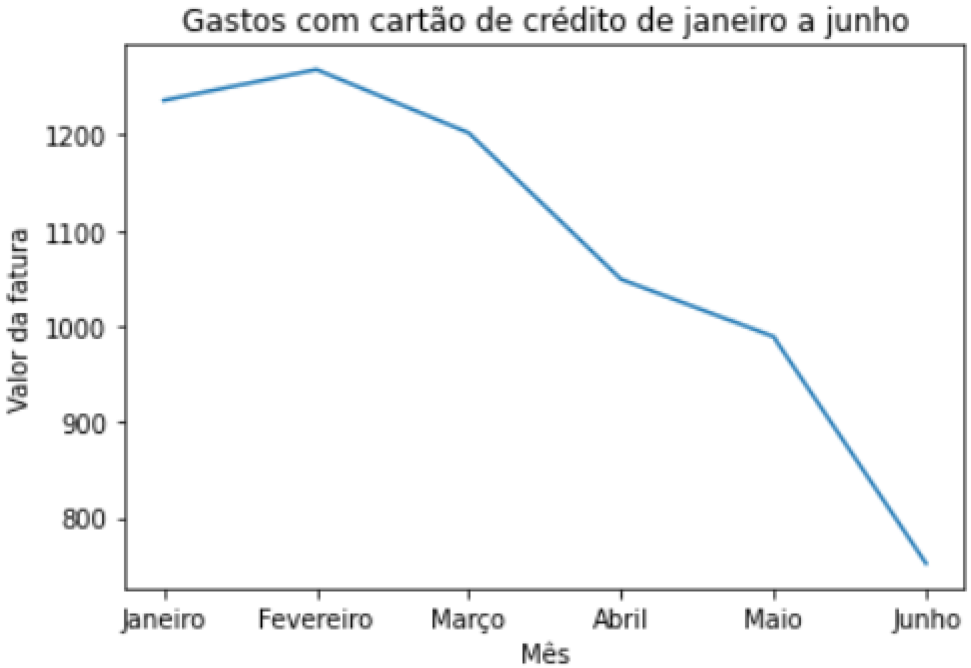
Exemplo:

Os gastos mensais com cartão de crédito de uma pessoa, de janeiro a junho de um determinado ano, são apresentados a seguir.

Mês	Janeiro	Fevereiro	Março	Abril	Maio	Junho
Valor	R$ 1.236,12	R$ 1.268,39	R$ 1.202,10	R$ 1.049,02	R$ 988,87	R$ 752,26

Faça uma representação gráfica dos dados e descreva o comportamento da série.

Temos uma série de dados associados ao tempo, cuja representação gráfica é:



Os comandos utilizados para a construção do gráfico foram:

```
import matplotlib.pyplot as plt
x=['Janeiro', 'Fevereiro', 'Março', 'Abril', 'Maio', 'Junho']
y=[1236.12, 1268.39, 1202.10, 1049.02, 988.87, 752.26]
plt.plot(x,y)
plt.title('Gastos com cartão de crédito de janeiro a junho')
plt.xlabel('Mês')
plt.ylabel('Valor da fatura')
plt.show()
```

De janeiro a fevereiro, ocorreu um aumento no valor dos gastos com cartão de crédito. Após o mês de fevereiro, mensalmente, há uma redução no valor da fatura.

Exemplo:

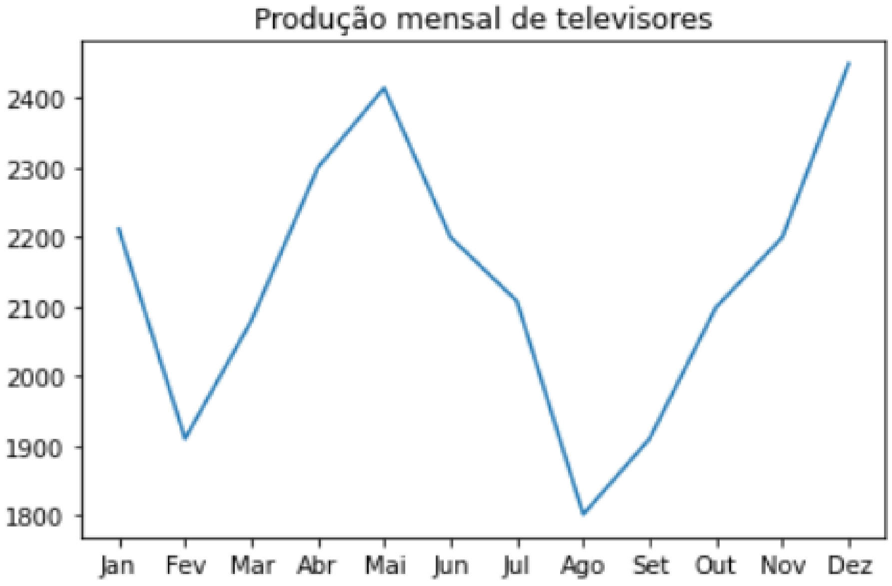
Na tabela a seguir, temos a produção mensal de televisores de uma determinada indústria.

Mês	Produção
Janeiro	2212
Fevereiro	1910
Março	2080
Abril	2300
Maio	2415
Junho	2200
Julho	2108
Agosto	1801
Setembro	1910
Outubro	2099
Novembro	2200
Dezembro	2450

Por meio do Python, construa o respectivo gráfico de linhas utilizando os comandos.

```
import matplotlib.pyplot as plt
x=['Jan', 'Fev', 'Mar', 'Abr', 'Mai', 'Jun', 'Jul', 'Ago', 'Set', 'Out', 'Nov', 'Dez']
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
plt.plot(x,y)
plt.title('Produção mensal de televisores')
plt.show()
```

Temos o gráfico a seguir:



Há muitos outros exemplos de séries temporais, tais como as quantidades de acidentes anuais em uma determinada rodovia, a cotação diária do dólar, a inflação de um país medida a cada ano, entre muitos outros.

Por outro lado, se pensarmos, por exemplo, em uma pesquisa de preços de um certo modelo de televisor em diferentes lojas, teremos os respectivos valores, mas não temos uma série temporal, pois temos os preços do refrigerador em diferentes lojas, mas não temos o acompanhamento desses preços com o passar do tempo. No entanto, se o estudo está relacionado ao acompanhamento do preço de um determinado televisor em relação ao passar de tempo, temos um exemplo de série temporal.

No estudo das séries temporais, temos alguns importantes elementos relacionados ao comportamento da variável, que é objeto de estudo.

Um deles é a **tendência**. Por meio da tendência, é possível observar se há crescimento ou decrescimento da variável.

Temos também a **estacionalidade**. Esse comportamento ocorre quando não há mudanças em relação à variável. Nesse caso, não ocorre crescimento e nem decrescimento. A variável apresenta um equilíbrio estável em torno de um valor constante.

Outro comportamento é a **sazonalidade**. É um comportamento muito corriqueiro quando há mudanças, tais como crescimento e decrescimento que ocorrem em determinadas épocas.

Além desses elementos, há situações em que ocorrem **períodos cíclicos**. Esse tipo de comportamento se refere às mudanças que se repetem em determinados períodos de tempo.

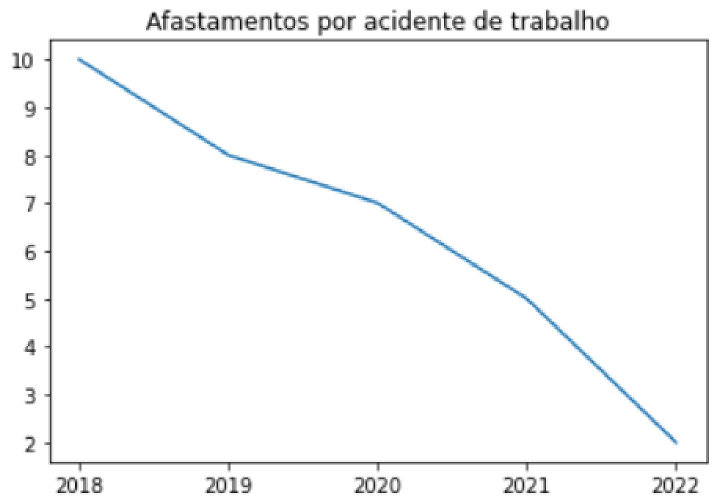
Por exemplo, podemos imaginar uma loja de materiais escolares que tem, a cada mês, um determinado volume de vendas e identifica que todos os anos, nos meses de janeiro e fevereiro, há um aumento considerável nas vendas em relação aos demais meses do ano.

Exemplo:

Na tabela a seguir, são apresentados os números relacionados aos afastamentos por acidente de trabalho em uma determinada empresa.

Ano	Acidentes
2018	10
2019	8
2020	7
2021	5
2022	2

Construa o respectivo gráfico e analise o comportamento da série de dados. É possível observar que o número de afastamentos por acidente de trabalho diminuiu com o passar do tempo. Podemos dizer então que há uma tendência de decrescimento.



Os comandos utilizados para a construção do gráfico foram:

```
import matplotlib.pyplot as plt
x=['2018', '2019', '2020', '2021', '2022']
y=[10, 8, 7, 5, 2]
plt.plot(x,y)
plt.title('Afastamentos por acidente de trabalho')
plt.show()
```

TEMA 2 – O USO DA REGRESSÃO LINEAR NO ESTUDO DE SÉRIES TEMPORAIS

Para obtermos estimativas ou analisarmos melhor o comportamento de séries temporais, podemos, muitas vezes, utilizar a regressão linear. Muitas vezes, é preciso saber se há uma correlação entre os dados, qual é a taxa de crescimento ou de decrescimento ou também fazer estimativas a partir de informações conhecidas. Caso os dados estejam relativamente próximos de uma reta, o uso da regressão linear, nesses casos, é uma opção.

Exemplo:

Por meio do Python, obtenha a reta de regressão e o coeficiente de correlação de Pearson relacionados ao conjunto de pontos a seguir:

Mês	**Produção**
Janeiro	12000
Fevereiro	13000
Março	14000
Abril	13000
Maio	15000
Junho	15000
Julho	17000
Agosto	17000
Setembro	17000

Outubro	18000
Novembro	19000
Dezembro	20000

Utilizando os comandos

from scipy import stats

import numpy as np

x=np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12])

y=np.array([12000, 13000, 14000, 13000, 15000, 15000, 17000, 17000, 17000, 18000, 19000, 20000])

a,b,c,d,e=stats.linregress(x,y)

print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')

print(f'Coeficiente de correlação de Pearson: {c:.4f}')

```
from scipy import stats
import numpy as np
x=np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12])
y=np.array([12000, 13000, 14000, 13000, 15000, 15000, 17000, 17000, 17000, 18000, 19000, 20000])
a,b,c,d,e=stats.linregress(x,y)
print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')
print(f'Coeficiente de correlação de Pearson: {c:.4f}')

Reta de regressão: y=692.31x+11333.33
Coeficiente de correlação de Pearson: 0.9779
```

https://docs.scipy.org/doc/scipy/tutorial/stats.html

Temos que a reta de regressão é dada por y=692,31x+11333,33. É possível observar que a estimativa de crescimento mensal é de aproximadamente 692 unidades. Como o coeficiente corresponde a 0,9779, ou seja, 97,79%, a correlação é forte.

Mas nem sempre a reta de regressão é a que melhor se ajusta a uma série temporal.

No exemplo a seguir, veremos um caso em que a correlação entre os dados é fraca e, dessa forma, a regressão linear não é a melhor opção de abordagem para uma análise mais detalhada dos dados.

Exemplo:

Na tabela a seguir, temos a produção mensal de televisores de uma determinada indústria.

Mês	Produção
Janeiro	2212
Fevereiro	1910
Março	2080
Abril	2300
Maio	2415
Junho	2200
Julho	2108
Agosto	1801
Setembro	1910
Outubro	2099
Novembro	2200
Dezembro	2450

Obtenha, a partir dos dados apresentados, a respectiva reta de regressão e o coeficiente de correlação de Pearson. Faça o gráfico dos dados originais e da reta de regressão.

Para resolvermos o problema, vamos utilizar os comandos a seguir:

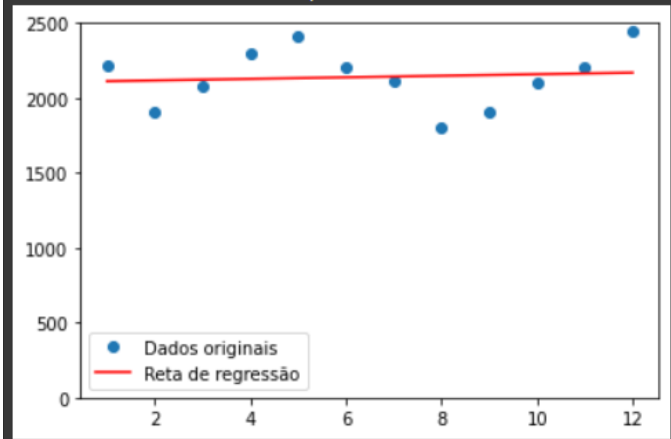
```
from scipy import stats
import matplotlib.pyplot as plt
import numpy as np
x=np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12])
y=np.array([2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099,
2200, 2450])
a,b,c,d,e=stats.linregress(x,y)
print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')
print(f'Coeficiente de correlação de Pearson: {c:.4f}')
plt.plot(x, y, 'o', label='Dados originais')
f=a*x+b
```

```
plt.plot(x, f, 'r', label='Reta de regressão')
plt.ylim(0, 2500)
plt.legend()
plt.show()
```

O que resulta em:

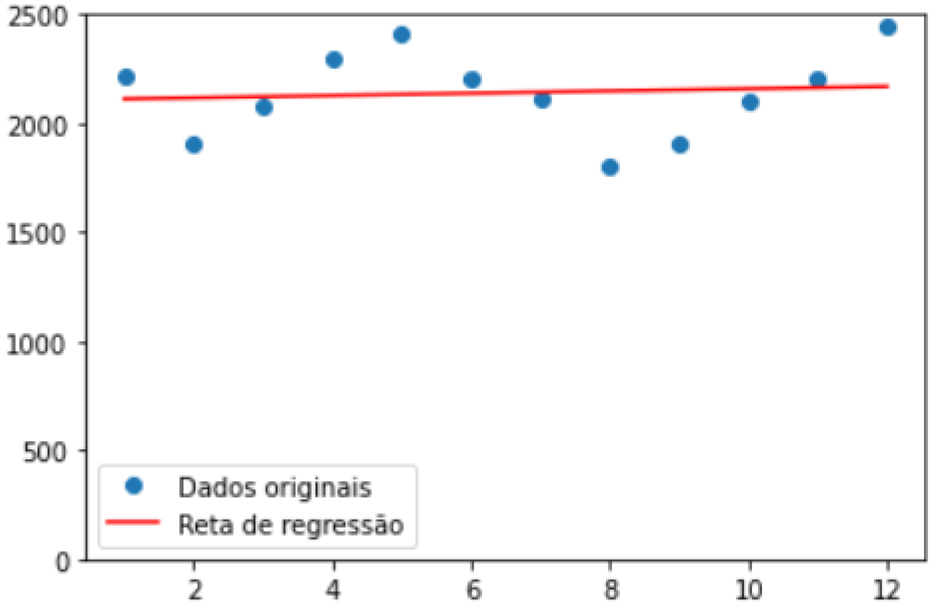
```
from scipy import stats
import matplotlib.pyplot as plt
import numpy as np
x=np.array([1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12])
y=np.array([2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450])
a,b,c,d,e=stats.linregress(x,y)
print(f'Reta de regressão: y={a:.2f}x+{b:.2f}')
print(f'Coeficiente de correlação de Pearson: {c:.4f}')
plt.plot(x, y, 'o', label='Dados originais')
f=a*x+b
plt.plot(x, f, 'r', label='Reta de regressão')
plt.ylim(0, 2500)
plt.legend()
plt.show()
```

```
Reta de regressão: y=5.16x+2106.85
Coeficiente de correlação de Pearson: 0.0934
```



Para este problema, temos um coeficiente de correlação de Pearson igual a 9,34%, o que indica uma correlação fraca.

Se observarmos os dados, veremos que eles apresentam um comportamento que oscila, ou seja, ora são crescentes, ora são decrescentes.



Nesse caso, o ideal é fazer o ajuste por outro tipo de função.

TEMA 3 – AJUSTES NÃO LINEARES PARA SÉRIES TEMPORAIS

Quando não há um comportamento linear para uma série temporal, é preciso fazer o ajuste de uma curva por meio de uma função não linear.

Por se tratar de um tipo bastante simples, e muito útil, iremos aprender uma forma de obtermos um polinômio que melhor se ajusta a um conjunto de dados por meio do Python. Se o grau "n" do polinômio é igual a 1, temos uma reta de regressão. Para um polinômio de grau 2, iremos obter uma parábola que melhor se ajusta aos dados. Podemos ter também um polinômio de grau 3, de grau 4 e assim por diante.

Mas como é possível saber qual é o melhor polinômio para cada caso?

Após a obtenção do polinômio, podemos calcular o respectivo coeficiente de determinação, representado por r^2, e que indica a proximidade que há entre os dados originais e o polinômio que está sendo utilizado. Quanto mais próximo de 1 (100%), melhor é o ajuste relacionado ao polinômio escolhido.

Na prática, para sabermos qual é o melhor polinômio para os dados que temos, basta escolhermos diferentes valores para "n" e verificarmos para qual deles o valor de r^2 é maior.

O comando que gera o polinômio é proveniente da biblioteca numpy:

```
p=numpy.poly1d(numpy.polyfit(x,y,n))
```

O termo “p” é o polinômio, “x” e “y” são os dados do problema e “n” é o grau que escolhemos para o polinômio.

Para obtermos o respectivo coeficiente de determinação, precisamos importar a função r2_score da biblioteca sklearn.metrics:

```
from sklearn.metrics import r2_score
```

Onde a sintaxe para o cálculo do coeficiente é:

```
r2=r2_score(y,p(x))
```

TEMA 4 – OBTENÇÃO DE AJUSTES POLINOMIAIS

Agora que já sabemos a forma de obtermos polinômios e de calcularmos os respectivos coeficientes de determinação por meio do Python, vamos aprender a resolver problemas relacionados a este importante assunto.

Exemplo:

Na tabela a seguir, temos a produção mensal de televisores de uma determinada indústria.

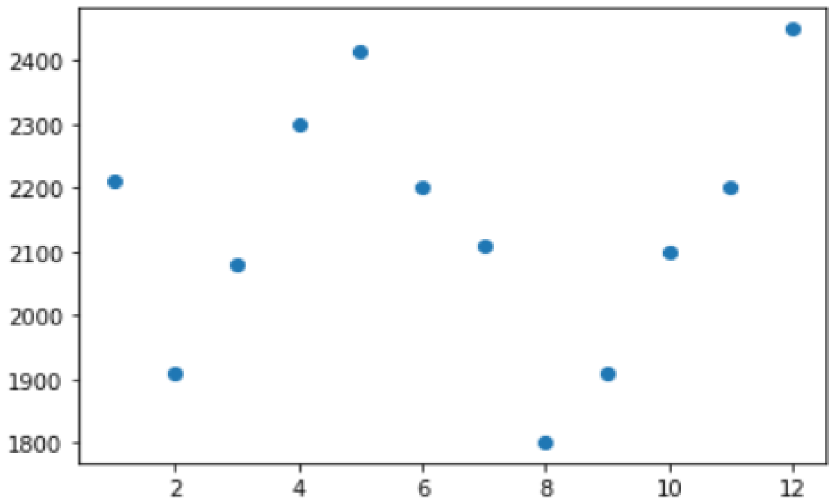
Mês	Produção
Janeiro	2212
Fevereiro	1910
Março	2080
Abril	2300
Maio	2415
Junho	2200
Julho	2108
Agosto	1801
Setembro	1910
Outubro	2099
Novembro	2200
Dezembro	2450

Faça o gráfico dos dados do problema. Em seguida, obtenha polinômios que melhor se aproximam dos dados e os respectivos coeficientes de determinação para n=2, 3, 4, e assim por diante.

O gráfico associado aos dados do problema é obtido a partir dos comandos a seguir:

```
import matplotlib.pyplot as plt
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
plt.plot(x,y,'o')
plt.show()
```

Resultando em:



Podemos observar que há uma oscilação dos valores de y associados à variação do tempo.

Para obtermos os polinômios de grau 2, 3, 4 e assim por diante, construirmos os respectivos gráficos e obtermos os respectivos coeficientes de determinação, precisamos importar as bibliotecas e função a seguir:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
```

Em seguida, vamos precisar dos dados do problema e do grau do polinômio:

```
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
```

n=2

Para construirmos o gráfico do polinômio, precisamos definir valores para “x”. Nesse caso, iremos armazenar na variável x2 os valores destinados à construção do gráfico do polinômio:

x2=numpy.linspace(1, 12, 100)

Para obtermos o polinômio e visualizarmos o resultado, basta fazermos:

p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)

O cálculo do coeficiente de determinação com a respectiva visualização é dado por:

r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')

Finalmente, para a construção dos gráficos dos dados originais e do polinômio que se ajusta aos dados, temos:

plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()

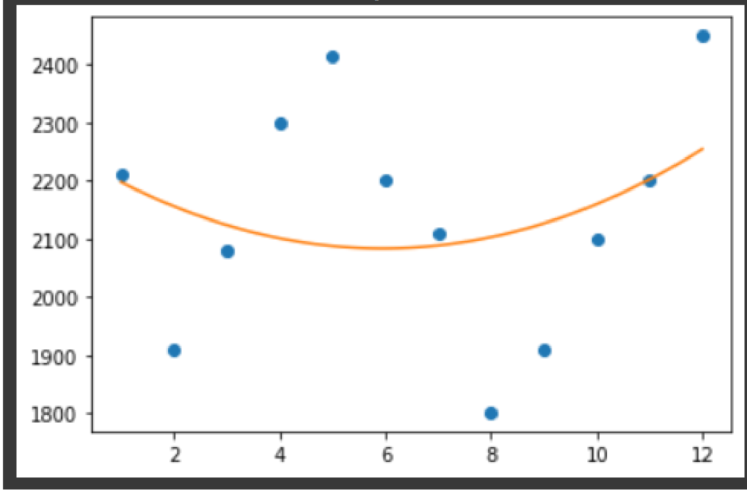
A sequência completa dos comandos é:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=2
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

Os resultados obtidos são:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=2
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

```
       2
4.677 x - 55.63 x + 2249
Coeficiente de determinação: 0.0756
```



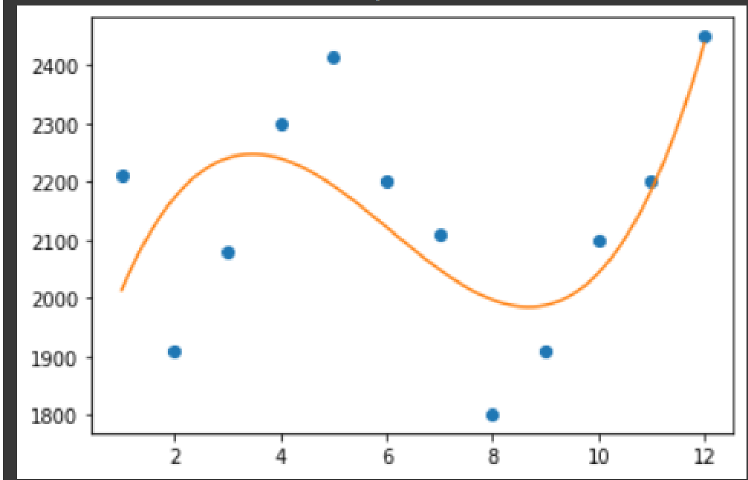
Podemos observar que o polinômio p(x)=4,577x^2-55,63x+2249 de grau 2 não é o ideal para este problema, pois temos um coeficiente de determinação de 7,56% e, por meio do gráfico, podemos perceber também que há muitos pontos distantes da curva.

Vamos repetir o processo utilizando agora um polinômio de grau 3.

Basta substituirmos 2 por 3 ao atribuirmos o valor à variável “n”, ou seja, basta fazermos n=3. Os demais comandos são os mesmos.

```python
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=3
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

```
      3         2
3.71 x - 67.66 x + 335.7 x + 1742
Coeficiente de determinação: 0.4405
```

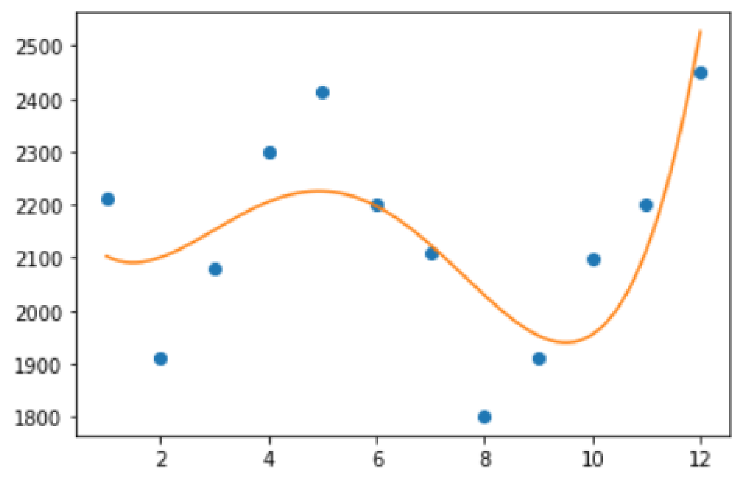


Agora, temos o polinômio $p(x)=3{,}71x^3-67{,}66x^2+335{,}7x+1742$ com um coeficiente de determinação igual a 0,4405, o que corresponde a 44,05%. Já apresenta um ajuste melhor, mas ainda não está suficientemente próximo dos dados do problema.

Para n=4, temos:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=4
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

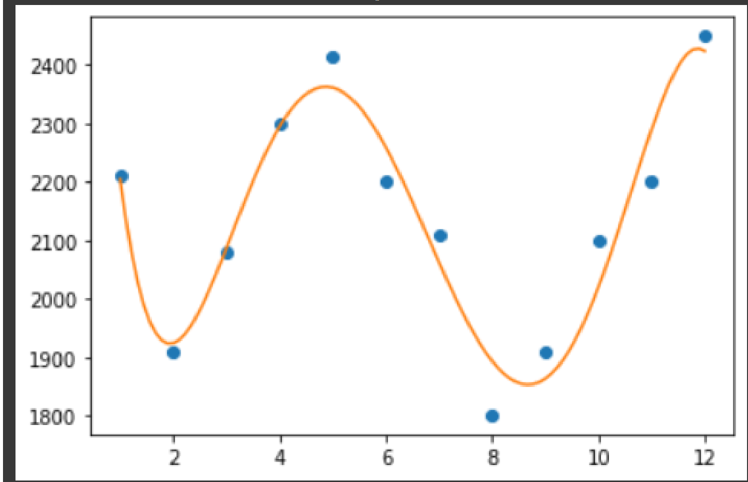
```
        4         3        2
0.7814 x - 16.61 x + 107 x - 218.6 x + 2230
Coeficiente de determinação: 0.5721
```



Para n=5:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=5
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

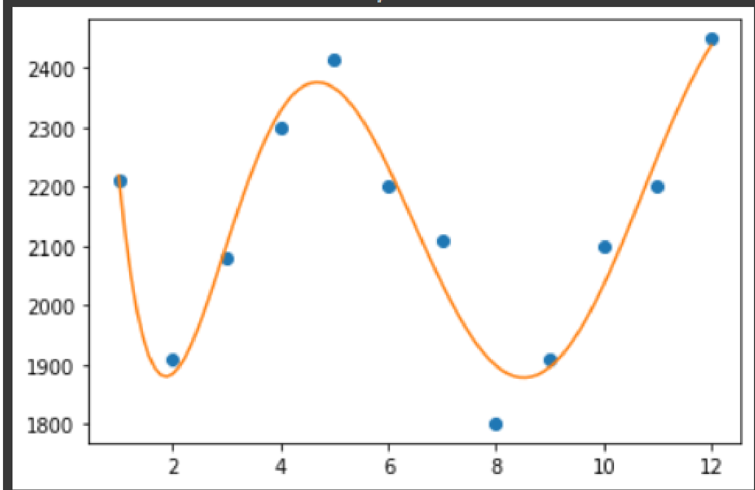
```
         5          4          3         2
-0.4642 x + 15.87 x - 195.1 x + 1037 x - 2251 x + 3598
Coeficiente de determinação: 0.9211
```



Para n=6:

```python
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=6
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

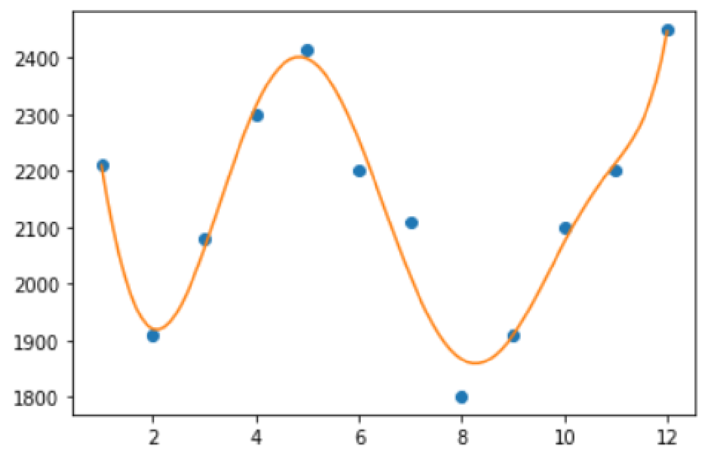
```
         6         5         4        3        2
0.03945 x - 2.003 x + 39.07 x - 365 x + 1658 x - 3280 x + 4168
Coeficiente de determinação: 0.9382
```



Para n=7:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=7
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

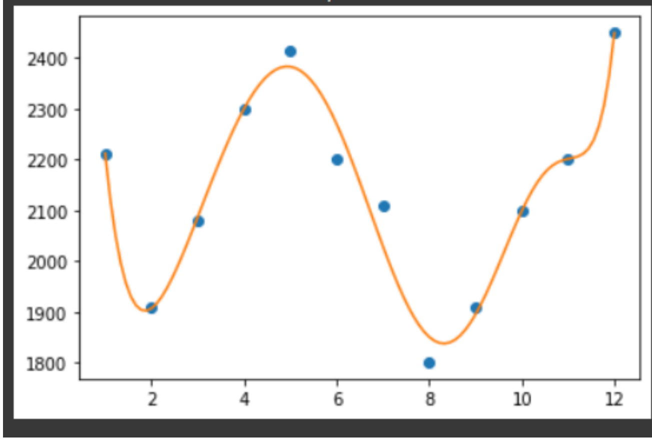
```
         7          6         5         4         3         2
0.01793 x - 0.7763 x + 12.97 x - 102.7 x + 372.1 x - 398.1 x - 515.1 x + 2842
Coeficiente de determinação: 0.9593
```



Para n=8:

```python
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=8
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

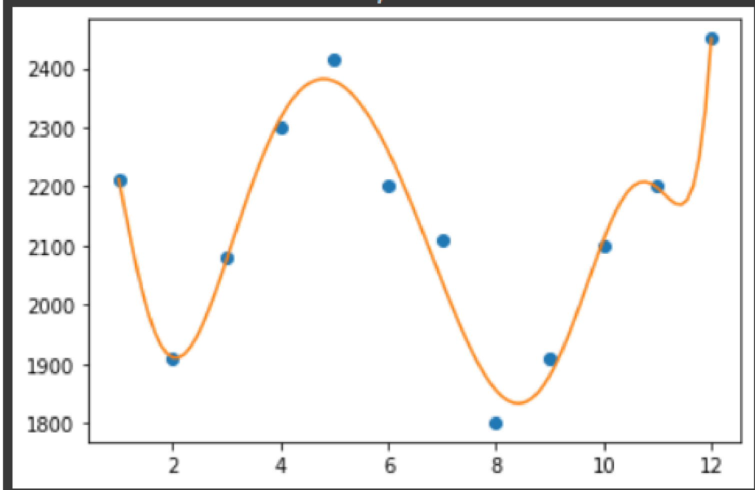
```
           8          7         6         5         4        3        2
0.004503 x - 0.2162 x + 4.288 x - 46.04 x + 298.1 x - 1232 x + 3230 x - 4661 x + 4619
Coeficiente de determinação: 0.9660
```



Para n=9:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=9
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

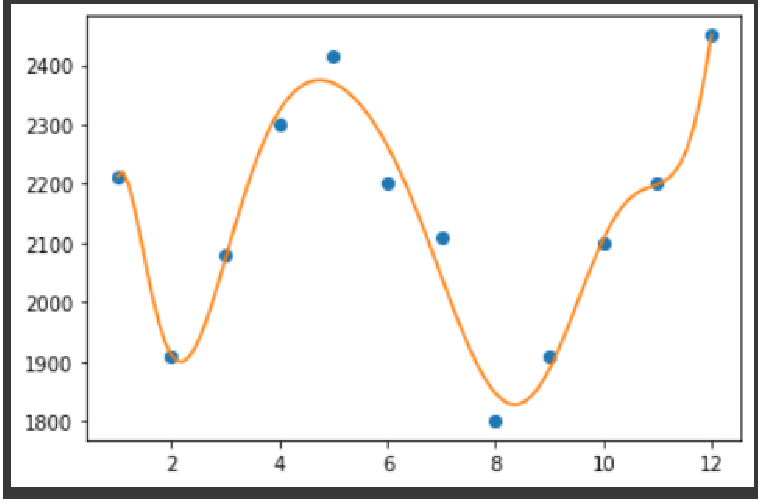
```
          9          8        7         6         5         4        3
0.001328 x - 0.0732 x + 1.72 x - 22.52 x + 179.8 x - 890.8 x + 2639 x
         2
 - 4184 x + 2783 x + 1705
Coeficiente de determinação: 0.9683
```



Para n=10:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=10
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

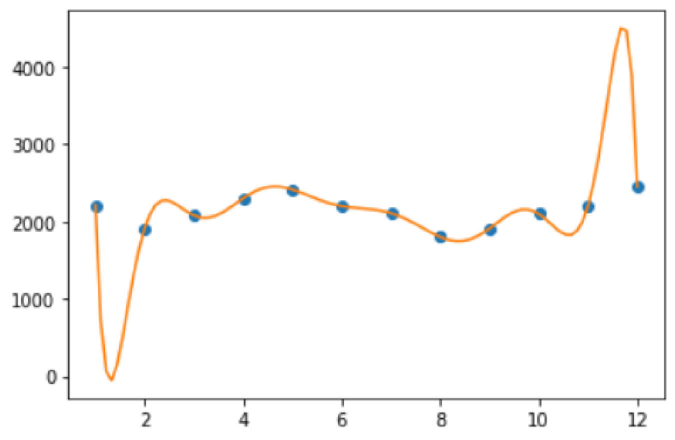
```
            10           9           8          7          6         5
-0.0004559 x  + 0.03096 x - 0.9094 x + 15.15 x - 157.8 x + 1068 x
         4             3             2
 - 4712 x + 1.318e+04 x - 2.183e+04 x + 1.874e+04 x - 4086
Coeficiente de determinação: 0.9690
```



Para n=11:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
n=11
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

```
           11          10         9         8        7        6
-0.002447 x   + 0.1745 x   - 5.484 x + 99.94 x - 1170 x + 9194 x
              5              4              3              2
 - 4.928e+04 x + 1.791e+05 x - 4.296e+05 x + 6.417e+05 x - 5.31e+05 x + 1.831e+05
Coeficiente de determinação: 1.0000
```



Para este problema, quanto maior o grau do polinômio, maior foi o coeficiente de determinação. Quando fizemos n=11, o coeficiente de determinação foi igual a 1 (100%), pois o polinômio obtido passou por todos os pontos dados (interpolação) em vez de fazer uma aproximação (regressão). Isso ocorre porque, quando interpolamos n+1 pontos, temos um polinômio de grau n.

Na prática, o grau do polinômio escolhido depende da escolha de quem está fazendo o estudo.

Em muitos casos, um polinômio quadrático é bastante simples e já é suficiente para um coeficiente de determinação próximo de 1.

Exemplo:

Uma pessoa deposita mensalmente diferentes valores em uma conta-poupança, cuja remuneração mensal também é variável. A tabela a seguir apresenta o saldo mensal dessa conta-poupança durante um ano:

Mês	**Saldo**
Janeiro	R$ 5.348,05
Fevereiro	R$ 12.013,83
Março	R$ 15.436,57
Abril	R$ 23.910,14
Maio	R$ 29.015,10
Junho	R$ 33.006,99
Julho	R$ 43.044,98
Agosto	R$ 51.230,33
Setembro	R$ 55.010,76
Outubro	R$ 62.667,77
Novembro	R$ 73.098,99
Dezembro	R$ 81.084,06

Obtenha um polinômio de grau 2 que melhor se ajusta aos dados apresentados. Obtenha também o coeficiente de determinação e o respectivo gráfico.

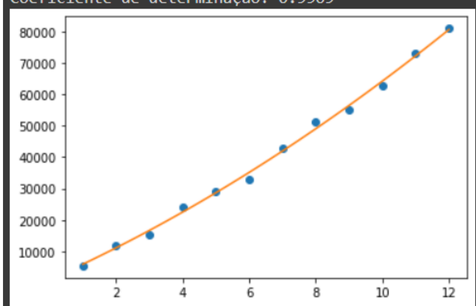
Para resolvermos o problema, basta considerarmos a sequência de comandos a seguir:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[5348.05, 12013.83, 15436.57, 23910.14, 29015.10, 33006.99, 43044.98, 51230.33, 55010.76, 62667.77, 73098.99, 81084.06]
n=2
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

Assim, obtemos os resultados a seguir:

```
import numpy
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score
x=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y=[5348.05, 12013.83, 15436.57, 23910.14, 29015.10, 33006.99, 43044.98, 51230.33, 55010.76, 62667.77, 73098.99, 81084.06]
n=2
x2=numpy.linspace(1, 12, 100)
p=numpy.poly1d(numpy.polyfit(x,y,n))
print(p)
r2=r2_score(y,p(x))
print(f'Coeficiente de determinação: {r2:.4f}')
plt.plot(x,y,'o')
plt.plot(x2,p(x2))
plt.show()
```

```
       2
157.1 x + 4761 x + 951.8
Coeficiente de determinação: 0.9969
```



O polinômio que melhor se ajusta ao conjunto de pontos é $p(x)=157{,}1x^2+4761x+951{,}8$, com um coeficiente de determinação igual a 0,9969, o que corresponde a 99,69%.

Graficamente, podemos perceber que a curva obtida está bastante próxima dos dados do problema.

TEMA 5 – MÉDIA MÓVEL

Além da regressão linear ou da regressão não linear, é muito comum o uso do método da média móvel para a análise de séries temporais. Nesse caso, consideramos uma sequência de médias de valores do problema a ser analisado em um determinado período de tempo. A sequência de médias apresenta uma tendência do comportamento do fato que está sendo observado.

A quantidade de valores utilizados para o cálculo da média móvel é fixa. Assim, a cada nova observação inserida no cálculo da média móvel, a observação mais antiga é retirada. O procedimento é repetido para todo o conjunto de pontos, gerando, assim, uma sequência de médias móveis que tem como objetivo suavizar a representação da série temporal e, com isso, auxiliar no estudo da série.

Se considerarmos o conjunto de n dados observados e calcularmos a sequência de médias de k dos dados, temos:

$$MM_1 = \frac{y_1 + y_2 + \ldots + y_k}{k}$$

$$MM_2 = \frac{y_2 + y_3 + \ldots + y_{k+1}}{k}$$

$$MM_3 = \frac{y_3 + y_4 + \ldots + y_{k+2}}{k}$$

e assim por diante.

Exemplo:

Na tabela a seguir, temos a produção mensal de televisores de uma determinada indústria.

Mês	Produção
Janeiro	2212
Fevereiro	1910
Março	2080
Abril	2300
Maio	2415
Junho	2200
Julho	2108
Agosto	1801
Setembro	1910
Outubro	2099
Novembro	2200
Dezembro	2450

Calcule as médias móveis para k=5.

Para a primeira média móvel, precisamos somar os valores associados à produção dos cinco primeiros meses e, em seguida, dividir o resultado por 5:

$$MM_1 = \frac{y_1 + y_2 + \ldots + y_k}{k}$$

$$MM_1 = \frac{2212 + 1910 + 2080 + 2300 + 2415}{5}$$

$$MM_1 = \frac{10917}{5}$$

$$MM_1 = 2183{,}4$$

Para calcularmos a segunda média móvel, precisamos somar as produções do segundo ao sexto mês e dividir o resultado por 5:

$$MM_2 = \frac{y_2 + y_3 + \ldots + y_{k+1}}{k}$$

$$MM_2 = \frac{1910 + 2080 + 2300 + 2415 + 2200}{5}$$

$$MM_2 = \frac{10905}{5}$$

$$MM_2 = 2181$$

A terceira média móvel é dada por:

$$MM_3 = \frac{y_3 + y_4 + \ldots + y_{k+2}}{k}$$

$$MM_3 = \frac{2080 + 2300 + 2415 + 2200 + 2108}{5}$$

$$MM_3 = \frac{11103}{5}$$

$$MM_3 = 2220{,}6$$

Para calcularmos MM_4, basta fazermos:

$$MM_4 = \frac{y_4 + y_5 + \ldots + y_{k+3}}{k}$$

$$MM_4 = \frac{2300 + 2415 + 2200 + 2108 + 1801}{5}$$

$$MM_4 = \frac{10824}{5}$$

$$MM_4 = 2164{,}8$$

As demais médias móveis são:

$$MM_5 = \frac{y_5 + y_6 + \ldots + y_{k+4}}{k}$$

$$MM_5 = \frac{2415 + 2200 + 2108 + 1801 + 1910}{5}$$

$$MM_5 = \frac{10434}{5}$$

$$MM_5 = 2086{,}8$$

$$MM_6 = \frac{y_6 + y_7 + \ldots + y_{k+5}}{k}$$

$$MM_6 = \frac{2200 + 2108 + 1801 + 1910 + 2099}{5}$$

$$MM_6 = \frac{10118}{5}$$

$$MM_6 = 2023{,}6$$

$$MM_7 = \frac{y_7 + y_8 + \ldots + y_{k+6}}{k}$$

$$MM_7 = \frac{2108 + 1801 + 1910 + 2099 + 2200}{5}$$

$$MM_7 = \frac{10118}{5}$$

$$MM_7 = 2023{,}6$$

$$MM_8 = \frac{y_8 + y_9 + \ldots + y_{k+7}}{k}$$

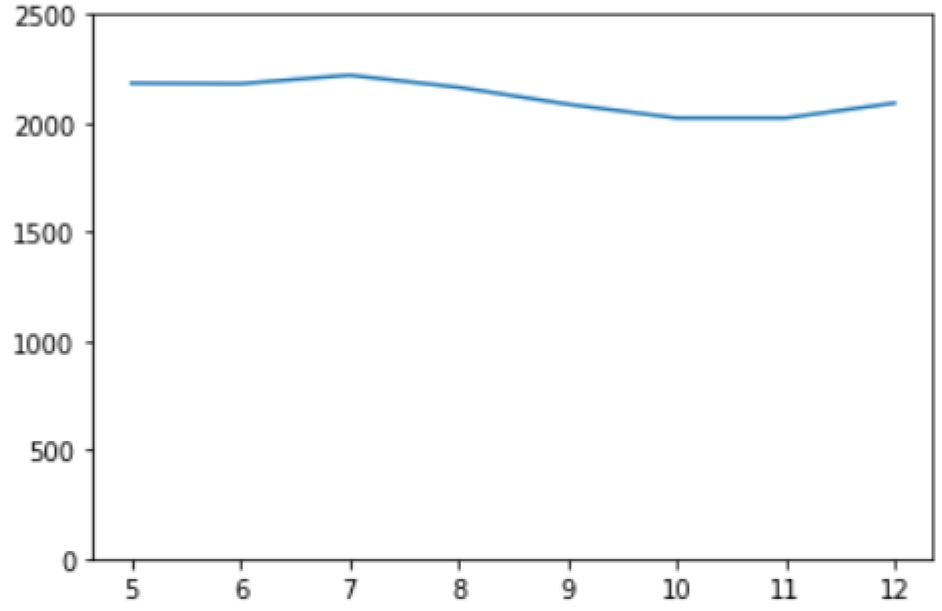
$$MM_8 = \frac{1801 + 1910 + 2099 + 2200 + 2450}{5}$$

$$MM_8 = \frac{10460}{5}$$

$$MM_8 = 2092$$

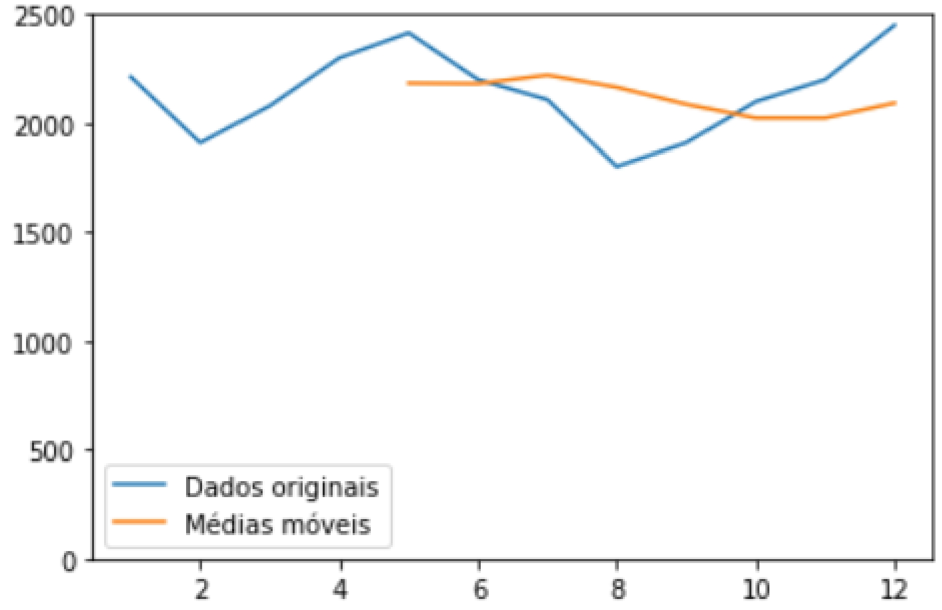
Agora que temos as médias móveis, podemos representá-las graficamente, lembrando que, para a primeira média móvel, consideramos os dados referentes aos cinco primeiros meses e, assim, ela está associada ao mês de maio, a segunda média móvel está associada ao mês de junho e assim por diante.

```
import matplotlib.pyplot as plt
x=[5, 6, 7, 8, 9, 10, 11, 12]
y=[2183.4, 2181, 2220.6, 2164.8, 2086.8, 2023.6, 2023.6, 2092]
plt.plot(x,y)
plt.ylim(0, 2500)
plt.show()
```



No gráfico a seguir, temos a comparação entre os dados originais e a média móvel.

```
import matplotlib.pyplot as plt
x1=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]
y1=[2212, 1910, 2080, 2300, 2415, 2200, 2108, 1801, 1910, 2099, 2200, 2450]
x2=[5, 6, 7, 8, 9, 10, 11, 12]
y2=[2183.4, 2181, 2220.6, 2164.8, 2086.8, 2023.6, 2023.6, 2092]
plt.plot(x1,y1,label='Dados originais')
plt.plot(x2,y2,label='Médias móveis')
plt.ylim(0, 2500)
plt.legend()
plt.show()
```



Comparando os dois gráficos, podemos perceber que a sequência de médias móveis suavizou a representação da série temporal em comparação com o gráfico obtido a partir dos dados originais. É importante ressaltar que os períodos de crescimento e de decrescimento foram alterados. Assim, é importante considerar que, ao trabalharmos com médias móveis, temos representações aproximadas quando comparadas com os dados originais.

FINALIZANDO

Chegamos ao final da nossa etapa de métodos quantitativos e também do nosso estudo. Aqui, falamos as séries temporais. Vimos que muitos fenômenos que ocorrem em nossas vidas estão diretamente relacionados com o passar do tempo. Assim, podemos acompanhar a evolução de variáveis em decorrência da variação do tempo. Aprendemos a utilizar regressão linear e regressão não linear no estudo de séries temporais. Aprendemos a calcular também médias móveis e a representá-las graficamente.

REFERÊNCIAS

CASTANHERIA, N. P. **Matemática aplicada**. 3. ed. Curitiba: Ibpex, 2010.

DEMANA, F. D. et al. **Pré-cálculo**. 2. ed São Paulo, Pearson, 2013.

FLEMING, D. M.; GONÇALVES, M. B.; **Cálculo A**: função de uma variável. 2. ed. São Paulo: Pearson, 2007.