UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
DEPARTAMENTO DE CIÊNCIAS NATURAIS, MATEMÁTICA E ESTATÍSTICA
DISCIPLINA: CÁLCULO NUMÉRICO
TURMA: 01
DOCENTE: MATHEUS DA SILVA MENEZES
DISCENTE: VICTOR BENOISTON JALES DE OLIVEIRA – 2016020720
ARTHUR WILLIAM PEREIRA CAVALCANTI – 2019021915
SAMUEL MARTINS BARBOSA - 2016020272

Trabalho referente à terceira unidade

MOSSORÓ – RN

JUNHO – 2021

Metodologia

Em todo o trabalho, será utilizado a linguagem python para resolver os problemas dados, com relação ao interpretador da linguagem, foi escolhido o *Pycharm*, na sua versão 2020.2, que roda o python 3.6. O trabalho conta com parte escrita (relatório presente), vídeo explicativo, e algoritmos desenvolvidos para resolução das questões.

Questão 01)

Os seguintes dados nos foram fornecidos:

Ano	1970	1980	1991	1996	2000	2010
Pop.	132.165	151.326	179.323	203.302	226.542	249.633

O método dos mínimos quadrados (MMQ) consiste em estabelecer a equação da reta,

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

Que melhor se ajuste aos pontos (no caso linear). Ou seja, determinaremos o a e o b. O erro é calculado pela distância do ponto à reta, e se dá por,

$$erro = |y_i - g(x_i)|$$

Logo, vamos procurar determinar g(x) com o menor erro possível. Portanto, a equação do erro total será,

$$Erro = \sum_{i=1}^{n} |y_i - g(x_i)|$$

Agora, teremos que calcular o erro mínimo, para isso, fazemos uso das derivadas parciais, só que, a função módulo, graficamente falando, tem um bico, logo, não é diferenciável nesse ponto. Logo, no lugar do módulo, utilizamos a mesma função, só que ao quadrado. Portanto, teremos,

$$\sum_{i=1}^{n} (y_i - a_{xi} - b)^2$$

Após dedução algébrica, teremos os seguintes resultados:

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

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

Utilizando o MMQ, temos a seguinte aproximação **linear**, e estimações:

```
         ESTATÍSTICA
------------------------------
Somatório de x = 11947
Somatório de y = 1142291
Somatório X.Y = 2277659745
Somatório X² = 23789497
Valor de a = 3079.199
Valor de b = -5940815.793
Valor de y médio = 190381.833
Valor de R² = 0.9733
------------------------------
A equação da reta é: g(x) = 3079.199X  -5940815.793
------------------------------
Estimativa população em 2003 = 226,819.017
Estimativa população em 2020 = 279,165.393
```

Ajuste exponencial

Nesse ajuste, temos a função,

$$g(x) = b.e^{ax}$$

Se aplicarmos ln aos dois lados, acabaremos com a seguinte função,

$$\ln(y) = ax + \ln(b).$$

Para utilizarmos a função anterior, precisaremos ao invés de b, aplicar o ln ao valor de b, ou seja, onde tínhamos b, teremos ln(b). Logo, ao acharmos b, vamos aplicar o exponencial a esse valor, daí, esse será nosso valor de b. Aplicando aos nossos valores, teremos:

```
         ESTATÍSTICA
------------------------------
Somatório de x = 11947
Somatório de y = 72.79682267069998
Somatório X.Y = 144967.86249867207
Somatório X² = 23789497
Valor de a = 0.017
Valor de b = 5.826918518456969e-10
Valor de y médio = 12.133
Valor de R² = 0.9986
------------------------------
A equação é: g(x) = 0.017X + 5.826918518456969e-10
------------------------------
Estimativa população em 2003 = 226,675.405
Estimativa população em 2020 = 301,462.547
```

Ajuste logarítmico

Temos a função,

$$g(x) = a.\ln(x) + b$$

Podemos perceber que é parecida com a aproximação linear, porém, ao invés de x, teremos ln(x), o resto, se repete. Portanto, basta mudarmos os valores de x, pelo seu logaritmo. Aplicando aos nossos valores, temos:

```
         ESTATÍSTICA
------------------------------
Somatório de x = 45.57872616695142
Somatório de y = 1142291
Somatório X.Y = 8678953.369540565
Somatório X² = 346.2369731099011
Valor de a = 6124117.307
Valor de b = -46331195.788
Valor de y médio = 190381.833
Valor de R² = 0.9728
------------------------------
A equação é: g(x) = 6124117.307ln(X)  -46331195.788
------------------------------
Estimativa população em 2003 = 226,801.803
Estimativa população em 2020 = 278,559.503
```

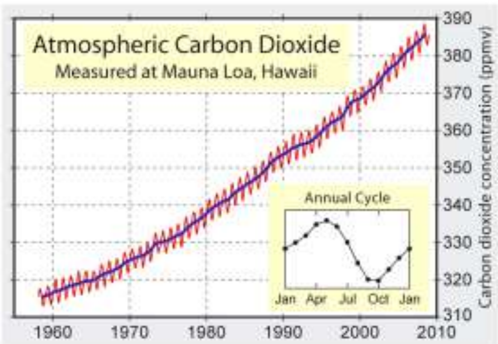
Parâmetro R²

O coeficiente de determinação (R^2) é um parâmetro utilizado para medir o nível de ajuste dos dados à reta, ou curva achada pelos métodos. Temos que, $0 \leq R^2 \leq 1$, quanto mais próximo de 1, maior o nível de ajuste dos dados. Se analisarmos o R^2 das aplicações anteriores, podemos chegar à conclusão de que o ajuste exponencial foi onde os pontos melhor se ajustaram, portanto, partindo dessa métrica, seria o ajuste mais apropriado para o nosso problema, o mais confiável.

Questão 02)

Na questão 02, temos os seguintes dados,

$x(ano.)$	1960	1975	1990	2000	2010
$f(x)(CO2)$	316	330	356	367	386



Ajuste linear

Como visto anteriormente, agora em posse da teoria e do algoritmo, basta substituir os valores de x e y, daí, teremos o seguinte resultado:

```
         ESTATÍSTICA
------------------------------
Somatório de x = 9916
Somatório de y = 1755
Somatório X.Y = 3482236
Somatório X² = 19666406
Valor de a = 1.729
Valor de b = -3077.934
Valor de y médio = 351.000
Valor de R² = 0.9375
------------------------------
A equação da reta é: g(x) = 1.729X  -3077.934
------------------------------
A estimativa da concentração em 2015 é de 405.982 ppm
A estimativa da concentração em 2020 é de 414.627 ppm
```

Ajuste exponencial

Como visto anteriormente, agora em posse da teoria e do algoritmo, basta substituir os valores de x e y, daí, teremos o seguinte resultado:

```
         ESTATÍSTICA
------------------------------
Somatório de x = 9916
Somatório de y = 29.291
Somatório X.Y = 58094.795
Somatório X² = 19666406
Valor de a = 0.005
Valor de b = 0.018018456231078655
Valor de y médio = 5.858
Valor de R² = 0.9993
------------------------------
A equação é: g(x) = 0.005X + 0.018018456231078655
------------------------------
A estimativa da concentração em 2015 é de 410.152 ppm
A estimativa da concentração em 2020 é de 420.491 ppm
```

Ajuste logarítmico

Como visto anteriormente, agora em posse da teoria e do algoritmo, basta substituir os valores de x e y, daí, teremos o seguinte resultado:

```
          ESTATÍSTICA
------------------------------
Somatório de x = 37.962
Somatório de y = 1755
Somatório X.Y = 13325.604
Somatório X² = 288.226
Valor de a = 3420.449
Valor de b = -25618.559
Valor de y médio = 351.000
Valor de R² = 0.9365
------------------------------
A equação é: g(x) = 3420.449ln(X)  -25618.559
------------------------------
A estimativa da concentração em 2015 é de 405.498 ppm
A estimativa da concentração em 2020 é de 413.974 ppm
```

Ajuste quadrático

Esse ajuste, leva em consideração a seguinte função,

$$P2(x) = cx^2 + bx + a$$

Portanto, encontraremos os valores de a, b e c, já vimos diversas técnicas de como fazer isso, e basta substituir os coeficientes na equação. Aplicando ao nosso resultado, teremos,

```
         ESTATÍSTICA
------------------------------
Somatório de x = 9935.000
Somatório de y = 1755.000
Somatório X.Y = 3489410.000
Somatório x² = 19742425.000
Somatório x³ = 39234470375.000
Somatório x^4 = 77977565970625
Somatório x²Y = 6938426050.000
Valor de R² = 0.9943
------------------------------
A equação é: g(x) = 0.007906843013247077X²  -29.9784358827749X + 28698.10106391284
------------------------------
A estimativa da concentração em 2015 é de 395.114 ppm
A estimativa da concentração em 2020 é de 404.743 ppm
```

Comparação de resultados

Método	Estimativa 2015 (ppm)	Estimativa 2020 (ppm)	R^2
Linear	405.982	414.627	0.9375
Exponencial	410.152	420.491	0.9993
Logaritmico	405.498	413.974	0.9365
Quadrático	395.114	404.743	0.9943

Portanto, se utilizarmos como métrica o coeficiente R^2, podemos concluir que a aproximação que mais se ajustou aos pontos, foi a exponencial, onde na teoria, teve um ajuste de 99.93% com os pontos.

Questão 04

No nosso quarto problema, é dado a seguinte integral,

$$\int_1^4 e^x dx$$

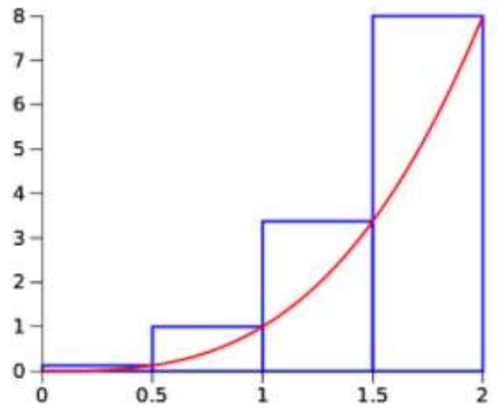
A partir dela, faremos aproximações usando os seguintes métodos.

- Soma de Riemann à direita;
- Soma de Riemann à esquerda;
- Método dos Trapézios;
- Método de 1/3 de Simpson.

Em todos os métodos, serão utilizados 1, 4, 10 e 100 subintervalos.

Soma de Riemann à direita

A soma de Riemann à direita, ou mão direita, é feita a partir do final do intervalo do retângulo, como o exemplo abaixo sugere, temos R[0; 0.5], utilizaremos o 0.5.



Soma de Riemann à Direita

Onde o início do retângulo se dá pela direita,

É trivial perceber que quanto mais repartições, o erro será menor em todos os métodos.

Além disso, temos que a soma é dada por,

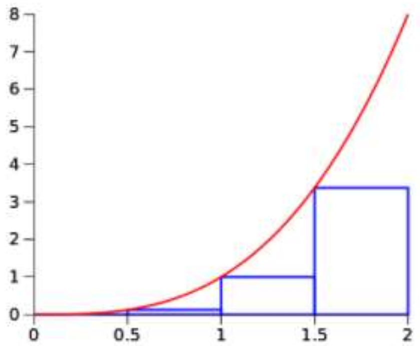
$$\sum_{i=1}^{n} f(ci).\Delta xi$$

Portanto, aplicando ao nosso caso, teremos o seguinte resultado,

```
Para 1 subintervalo:
------------------------------
Soma de Riemann = 0.000
Integral = 51.880
Erro = 51.880
Porcentagem do erro = 100.000%
------------------------------
Para 4 subintervalos:
------------------------------
Soma de Riemann = 32.796
Erro = 19.084
Porcentagem do erro = 36.786%
------------------------------
Para 10 subintervalos:
------------------------------
Soma de Riemann = 43.671
Erro = 8.209
Porcentagem do erro = 15.823%
------------------------------
Para 100 subintervalos:
------------------------------
Soma de Riemann = 51.024
Erro = 0.856
Porcentagem do erro = 1.650%
```

Soma de Riemann à esquerda

A soma de Riemann à esquerda, ou mão esquerda, é feita a partir do início do intervalo do retângulo, como o exemplo abaixo sugere, temos R[0; 0.5], utilizaremos o 0.



Soma de Riemann à Esquerda[1]

Além disso, temos que a soma é dada por,

$\sum_{i=1}^{n-1} f(ci).\Delta xi$

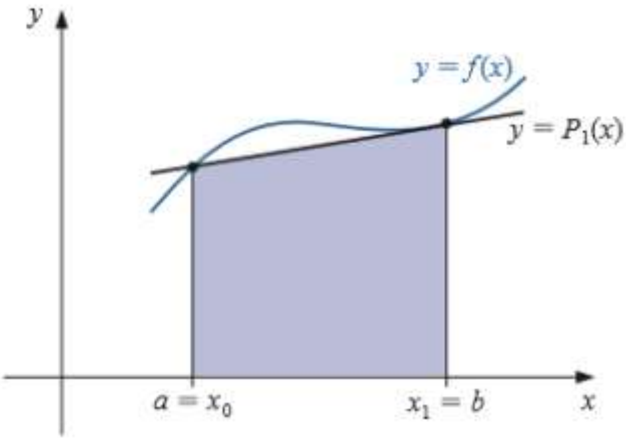
Portanto, aplicando ao nosso caso, teremos o seguinte resultado,

```
Para 1 subintervalo:
------------------------------
Soma de Riemann = 8.155
Integral = 51.880
Erro = 43.725
Porcentagem do erro = 84.281%
------------------------------
Para 4 subintervalos:
------------------------------
Soma de Riemann = 34.834
Erro = 17.046
Porcentagem do erro = 32.856%
------------------------------
Para 10 subintervalos:
------------------------------
Soma de Riemann = 44.486
Erro = 7.393
Porcentagem do erro = 14.251%
------------------------------
Para 100 subintervalos:
------------------------------
Soma de Riemann = 51.106
Erro = 0.774
Porcentagem do erro = 1.493%
```

Método dos trapézios

No método dos trapézios, teremos o seguinte somatório,

$$At = \frac{h}{2}\left[y_0 + \sum_{i=1}^{n-1} y_i + \ y_n\right]$$



Portanto, aplicando aos nossos valores, temos o seguinte resultado,

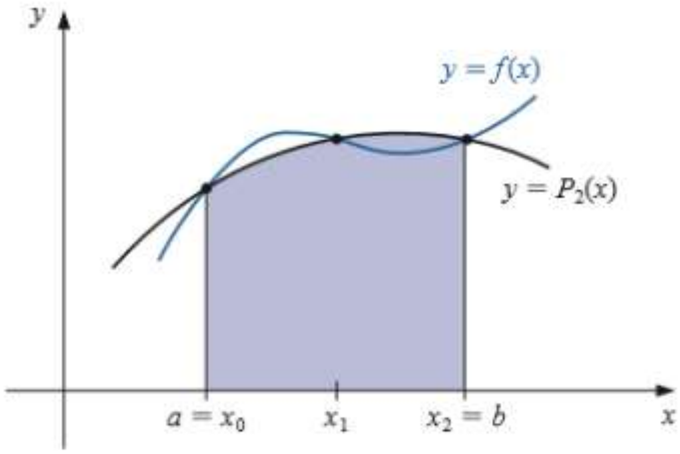
```
Para 1 subintervalo:
------------------------------
Soma dos trapézios = 85.975
Integral = 51.880
Erro = 34.095
Porcentagem do erro = 65.719%
------------------------------
Para 4 subintervalos:
------------------------------
Soma dos trapézios = 54.289
Erro = 2.409
Porcentagem do erro = 4.644%
------------------------------
Para 10 subintervalos:
------------------------------
Soma dos trapézios = 52.268
Erro = 0.389
Porcentagem do erro = 0.749%
------------------------------
Para 100 subintervalos:
------------------------------
Soma dos trapézios = 51.884
Erro = 0.004
Porcentagem do erro = 0.007%
```

Método de 1/3 de Simpson

Já no método de 1/3 de Simpson, teremos a seguinte aproximação,

$$I_2 = \frac{h}{3}[y_0 + 4y_1 + y_2]$$

Observação: É necessário ter um número par de pontos.



Portanto, aplicando aos nossos valores, temos o seguinte resultado,

```
1/3 de Simpson = 57.316
Integral = 51.880
Erro = 5.437
Porcentagem do erro = 10.479%
------------------------------
Para 4 subintervalos:
------------------------------
1/3 de Simpson = 51.965
Erro = 0.085
Porcentagem do erro = 0.165%
------------------------------
Para 10 subintervalos:
------------------------------
1/3 de Simpson = 51.882
Erro = 0.002
Porcentagem do erro = 0.004%
------------------------------
Para 100 subintervalos:
------------------------------
1/3 de Simpson = 51.880
Erro = 0.000
Porcentagem do erro = 0.000%
```

Comparação

Comparação				
Método	Subintervalos	Integral	Resultado	Erro (%)
Riemann direita	1	51.880	0	100.000
Riemann direita	4	51.880	32.796	36.786
Riemann direita	10	51.880	43.671	15.823
Riemann direita	100	51.880	51.024	1.650
Riemann esquerda	1	51.880	8.155	84.281
Riemann esquerda	4	51.880	34.834	32.856
Riemann esquerda	10	51.880	44.486	14.251
Riemann esquerda	100	51.880	51.106	1.493
Trapézios	1	51.880	85.975	65.719
Trapézios	4	51.880	54.289	4.644
Trapézios	10	51.880	52.268	0.749
Trapézios	100	51.880	51.884	0.007
1/3 Simpson	1	51.880	57.316	10.479
1/3 Simpson	4	51.880	51.965	0.165
1/3 Simpson	10	51.880	51.882	0.004
1/3 Simpson	100	51.880	51.880	0.000

Utilizando como métrica o erro, podemos concluir que o método mais confiável foi o 1/3 de Simpson. É possível concluir que, indiferente do método, quanto mais subintervalos forem fornecidos, menor será o erro, consequentemente, maior será a confiança.

Referências bibliográficas

[1] SymPy Development Team. SymPy documentation. 2021. Disponível em: < https://docs.sympy.org/latest/modules/integrals/integrals.html>;

[2] Numpy Development Team. NumPy documentation. 2008-2020. Disponível em: < https://numpy.org/doc/> ;

[3] NYKAMP, D.Q. “Calculating the area under a curve using Riemann sums.”. From Math Insight. Disponível em: < https://mathinsight.org/calculating_area_under_curve_riemann_sums>