AT1 - Metodos Numericos

March 21, 2024

3. A constante de Euler é definida como

$$e = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x$$

0.1 Resposta

[ ]:
```
def binario16_para_decimal(vetor, i_expoente=2, i_mantissa=6):
    sinal = -1 if vetor[0] == 1 else 1 # Guadra o sinal do número
    sinal_expoente = -1 if vetor[1] == 1 else 1 # Guarda o sinal do expoente

    expoente = 0
    for i in range(i_expoente, i_mantissa):
        expoente += vetor[i] * 2 ** (i_mantissa - 1 - i)  # Calculo iterativo␣
 ↪do expoente de n bits
    expoente *= sinal_expoente # Expoente recebe o sinal após conversão

    mantissa = 0
    for i in range(i_mantissa, len(vetor)):
        mantissa += vetor[i] * 2 ** (i_mantissa - 1 - i) # Calculo iterativo da␣
 ↪mantissa de n bits

    return sinal * mantissa * (2 ** expoente)

vetor = [0,1,0,0,1,1,1,0,1,0,1,1,0,0,0,1] ## Vetor dado
print("a)Número decimal representado é:", binario16_para_decimal(vetor))
```

```
a)Número decimal representado é: 0.0841064453125
```

[ ]:
```
vetor_inferior = [0,1,0,0,1,1,1,0,1,0,1,1,0,0,0,0]
print("b)\tO número de máquina imediatamente inferior é: ", vetor_inferior)
print("\tNúmero decimal imediatamente inferior é: ",␣
 ↪binario16_para_decimal(vetor_inferior))
```

```
b)      O número de máquina imediatamente inferior é:  [0, 1, 0, 0, 1, 1, 1, 0,
```

```
1, 0, 1, 1, 0, 0, 0, 0]
        Número decimal imediatamente inferior é:  0.083984375
```

c) Qual´e o numero de maquina imediato superior? Informe seu valor decimal.

[ ]:
```python
vetor_superior = [0,1,0,0,1,1,1,0,1,0,1,1,0,0,1,0]
print("O número de máquina imediatamente superior é: ", vetor_superior)
print("Número imediatamente superior é: ",␣
 ↪binario16_para_decimal(vetor_superior))
```

```
O número de máquina imediatamente superior é:  [0, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1,
1, 0, 0, 1, 0]
Número imediatamente superior é:  0.084228515625
```

d) Qual o intervalo de números reais que podem ser representados na notacão de 16 bits?

O maior número decimal possível será quando todos os bits da mantissa forem 1 e o expoente for o máximo possível, considerando que o sinal do expoente seja positivo. O menor número decimal possível será quando todos os bits da mantissa forem 1 e o expoente for o máximo possível, considerando que o sinal do expoente seja positivo e o sinal do número negativo.

[ ]:
```python
vetor_inicio_faixa = [1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1]
vetor_final_faixa = [0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1]
print("O menor número real possível de ser representado é: ",␣
 ↪int(binario16_para_decimal(vetor_inicio_faixa)))
print("O maior número real possível de ser representado é: ",␣
 ↪int(binario16_para_decimal(vetor_final_faixa)))
```

```
O menor número real possível de ser representado é:  -32736
O maior número real possível de ser representado é:  32736
```

e) Utilize os resultados anteriores para explicar porque apenas um subconjunto dos números reais pode ser representado pelo computador.

Limitação da mantissa: Com 10 bits disponíveis para a mantissa, podemos representar até 2^1 = 1024 diferentes combinações binárias. Isso nos permite representar a parte fracionária dos números com até 10 bits de precisão. No entanto, essa precisão é limitada e números que exigem mais de 10 bits para representar sua parte fracionária precisarão ser arredondados ou truncados, resultando em uma perda de precisão.

Limitação do expoente com sinal: Com 4 bits disponíveis para o expoente (incluindo 1 bit para o sinal do expoente), podemos representar números muito grandes ou muito pequenos. O bit de sinal do expoente nos permite representar números positivos e negativos. Com 4 bits, podemos representar números de aproximadamente -8 a 7.

Representação finita: Como todos os números de ponto flutuante são representados em uma quantidade finita de bits, há um limite para a precisão e a faixa de valores que podemos representar. Com um total de 16 bits (10 para a mantissa, 1 para o sinal, 1 para o sinal do expoente e 4 para o expoente), estamos limitados a representar números de ponto flutuante em uma faixa específica e com uma precisão limitada.

Assim, mesmo com um formato de ponto flutuante de 16 bits, apenas um subconjunto finito e discreto dos números reais pode ser representado com precisão, devido às limitações no tamanho da mantissa e do expoente, bem como à representação finita dos números em um computador.

3. A constante de Euler é definida como

$$e = \lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x$$

3. A constante de Euler é definida como

$$e = \lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x$$

[ ]:
```python
import math

def calcular_raiz_i(a, b, c):
    delta = round(b ** 2, 4) - round(round(4 * a, 4) * c, 4)

    return round((- b + round(math.sqrt(delta), 4)) / round(2 * a, 4), 4)

raiz_i = calcular_raiz_i(1, 62.10, 1)
print("Raiz é: ", raiz_i)
```

```
Raiz é:  -0.0161
```

3. A constante de Euler é definida como

$$e = \lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x$$

[ ]:
```python
import math

def calcular_raiz_i(a, b, c):
    delta = round(b ** 2, 4) - round(round(4 * a, 4) * c, 4)

    return round((- b - round(math.sqrt(delta), 4)) / round(2 * a, 4), 4)

raiz_i = calcular_raiz_i(1, 62.10, 1)
print("Raiz é: ", raiz_i)
```

Raiz é: -62.0839

3. A constante de Euler é definida como

$$e = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x$$

[ ]:
```python
import math

def calcular_raiz_iii(a, b, c):
    delta = round(b ** 2, 4) - round(round(4 * a,4) * c,4)

    return  round(round(-2 * c, 4) / (b + round(math.sqrt(delta),4)), 4)

print("Raiz é: ", calcular_raiz_iii(1, 62.10, 1))
```

Raiz é: -0.0161

3. A constante de Euler é definida como

$$e = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x$$

[ ]:
```python
import math

def calcular_raiz_iii(a, b, c):
    delta = round(b ** 2, 4) - round(round(4 * a,4) * c,4)

    return  round(round(-2 * c, 4) / (b - round(math.sqrt(delta),4)), 4)

print("Raiz é: ", calcular_raiz_iii(1, 62.10, 1))
```

Raiz é: -62.1118

a) Calcule os erros absolutos para cada estimativa.

$$E_T = \left|p - p^*\right|$$

Valores reais = X1: -0.0161072374, X2: -62.0838927626

i - Ea = | -0.0161072374 - (-0.0161) | = 0.0000072374

ii - Ea = | -62.0838927626 - (-62.0839) | = 0.0000072374

iii - Ea = | -0.0161072374 - (-0.0161) | = 0.0000072374

iv - Ea = | -62.0838927626 - (-62.1118) | = 0.0279072374

b) Calcule os erros relativos para cada estimativa.

3. A constante de Euler é definida como

$$e = \lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x$$

Valores reais = X1: -0.0161072374, X2: -62.0838927626

i - Er = | -0.0161072374 - (-0.0161) | / 0.0161072374 = 0.000449326

ii - Er = | -62.0838927626 - (-62.0839) | / 62.0838927626 = 0.0000001166

iii - Er = | -0.0161072374 - (-0.0161) | / 0.0161072374 = 0.000449326

iv - Er = | -62.0838927626 - (-62.1118) | / 62.0838927626 = 0.0004495085

c) Considerando os resultados de (a) e (b) qual definição de erro´e mais apropriada? Justifique.

Dados os resultados encontrados o erro relativo é levemente preferível por expressas o erro em relação ao tamanho do valor verdadeiro, quanto maior o erro em relação ao valor real, podemos dizer que pior é a estimativa. Porém erro absoluto pode ser útil quando estamos mais interessado na magnitude pura do erro, independentemente da maquinitude do valor verdadeiro. Em conclusão percebo que é útil considerar ambos os tipos de erro para obter uma compreensão completa do desempenho ou precisão de um método ou estimativa.

d) Se compararmos os resultados obtidos pelas expressoes (i) e (ii), com os das expressões (iii) e (iv). Porque os resultados s˜ao diferentes? Qual das alternativas devemos utilizar?

As alternativas i e ii pra resulução de raízes de uma equação de 2° grau, são preferíveis porque garantem uma precisão maior no cálculo das raízes, minimizando tanto os erros relativos quanto os absolutos em comparação com as alternativas iii e iv.

3. A constante de Euler é definida como

$$e = \lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x$$

a) Utilize a definição anterior para calcular estimativas de e para x = $\{10\ , 10^1\ , 10^2\ , 10^3\ , . . . , 10^{32}\}$.

```
[ ]: def lim_const_euler(x):
    euler_estimado = 1

    if x == 0: ## Caso Base
        vetor_resultados_euler.append(euler_estimado)

    for i in range(1, x + 1): ## Caso Geral
        euler_estimado += 1 / math.factorial(i)
    vetor_resultados_euler.append(euler_estimado)

vetor_resultados_euler = []
for x in range(0,33):
    lim_const_euler(10 ** x)
    print(f"O número de Euler com x = 10^{x}", "é: 0",␣
 ↪vetor_resultados_euler[x])
```

```
O número de Euler com x = 10^0 é: 0 2.0
O número de Euler com x = 10^1 é: 0 2.7182818011463845
O número de Euler com x = 10^2 é: 0 2.7182818284590455
O número de Euler com x = 10^3 é: 0 2.7182818284590455
O número de Euler com x = 10^4 é: 0 2.7182818284590455
```

```
---------------------------------------------------------------------------
KeyboardInterrupt                         Traceback (most recent call last)
/home/andrei/Documents/GIT/MESTRADO/Métodos Numéricos/AT1 - Metodos Numericos.
 ↪ipynb Cell 31 line 1
     <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
 ↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
 ↪ipynb#X42sZmlsZQ%3D%3D?line=10'>11</a> vetor_resultados_euler = []
     <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
 ↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
 ↪ipynb#X42sZmlsZQ%3D%3D?line=11'>12</a> for x in range(0,33):
---> <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
 ↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
 ↪ipynb#X42sZmlsZQ%3D%3D?line=12'>13</a>     lim_const_euler(10 ** x)
     <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
 ↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
 ↪ipynb#X42sZmlsZQ%3D%3D?line=13'>14</a>     print(f"O número de Euler com x =␣
 ↪10^{x}", "é: 0", vetor_resultados_euler[x])

/home/andrei/Documents/GIT/MESTRADO/Métodos Numéricos/AT1 - Metodos Numericos.
 ↪ipynb Cell 31 line 8
      <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
 ↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
 ↪ipynb#X42sZmlsZQ%3D%3D?line=4'>5</a>     vetor_resultados_euler.
 ↪append(euler_estimado)
      <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
 ↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
 ↪ipynb#X42sZmlsZQ%3D%3D?line=6'>7</a> for i in range(1, x + 1): ## Caso Geral
```

```
----> <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
↪ipynb#X42sZmlsZQ%3D%3D?line=7'>8</a>     euler_estimado += 1 / math.
↪factorial(i)
      <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
↪ipynb#X42sZmlsZQ%3D%3D?line=8'>9</a> vetor_resultados_euler.
↪append(euler_estimado)

KeyboardInterrupt:
```

b) Para cada estimativa calcule o desvio relativo considerando como valor exato e = 2,718281828459.

[ ]:
```
for x in range(0,32):
    euler_real = 2.718281828459
    desvio = euler_real - vetor_resultados_euler[x]
    desvio_absoluto = abs(desvio)
    desvio_relativo = desvio_absoluto / euler_real
    print(f"O desvio relativo do número de Euler Estimado com x = 10^{x}", "é:
↪", desvio_relativo)
```

```
O desvio relativo do número de Euler Estimado com x = 10^0 é:
0.26424111765710306
O desvio relativo do número de Euler Estimado com x = 10^1 é:
1.0047749646339605e-08
O desvio relativo do número de Euler Estimado com x = 10^2 é:
1.6827242906040848e-14
O desvio relativo do número de Euler Estimado com x = 10^3 é:
1.6827242906040848e-14
O desvio relativo do número de Euler Estimado com x = 10^4 é:
1.6827242906040848e-14
```

```
---------------------------------------------------------------------------
IndexError                                Traceback (most recent call last)
/home/andrei/Documents/GIT/MESTRADO/Métodos Numéricos/AT1 - Metodos Numericos.
↪ipynb Cell 33 line 3
      <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
↪ipynb#X44sZmlsZQ%3D%3D?line=0'>1</a> for x in range(0,32):
      <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
↪ipynb#X44sZmlsZQ%3D%3D?line=1'>2</a>     euler_real = 2.718281828459
----> <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
↪ipynb#X44sZmlsZQ%3D%3D?line=2'>3</a>     desvio = euler_real -␣
↪vetor_resultados_euler[x]
      <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
↪ipynb#X44sZmlsZQ%3D%3D?line=3'>4</a>     desvio_absoluto = abs(desvio)
```

```
    <a href='vscode-notebook-cell:/home/andrei/Documents/GIT/MESTRADO/
↪M%C3%A9todos%20Num%C3%A9ricos/AT1%20-%20Metodos%20Numericos.
↪ipynb#X44sZmlsZQ%3D%3D?line=4'>5</a>      desvio_relativo = desvio_absoluto /␣
↪euler_real

IndexError: list index out of range
```

c) Qual o comportamento do erro com o aumento de x? Comente os resultados.

Quanto maior o valor de x, mais termos da série precisam ser calculados para obter uma precisão desejada. Se o valor de x for muito grande, a série de Taylor pode se tornar divergente ou o cálculo pode se tornar numericamente instável devido a problemas de precisão finita em representar números muito grandes ou muito pequenos em ponto flutuante.

Portanto, enquanto aumentar x em valores razoáveis geralmente resultará em um erro de aproximação decrescente (desde que a série seja calculada com precisão suficiente), aumentar x além de um certo ponto pode introduzir erros significativos devido a limitações numéricas. Em muitos casos, é necessário usar técnicas de cálculo mais avançadas para lidar com valores extremos de x.