

1 - (5) Use o **Método da Bissecção** para aproximar a raiz cúbica real de um número real r . No cálculo **proponha tabela a ser preenchida e aproxime essa raiz com um erro absoluto menor que 0,01**.

Sabendo que o Método das tangentes (ou de Newton-Raphson) para resolver equações às substitui por uma função iterativa na forma

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

diga como faria por esse método para calcular uma das raiz cubicas de r (usando a mesma equação anterior, isto é: $f(x) = x^3 - r = 0$.) **Resolva PELOS DOIS METODOS e diga quantas iterações foram preciso em cada caso.** Considere a mesma precisão e use r como o número formado pelos **últimos 2 dígitos de sua de matricula**.

Solução geral:

Para o aluno acima: $f(x) = x^3 - 12 = 0$, resolvendo-se sabe-se que uma solução será $(12)^{1/3} = 2,2894284851066637356160844238794.....$

Solução M. da Bissecção :

Considerando usar a tabela abaixo para a Bissecção:

Ite- ra- ção	a	b	F(a)	F(b)	x_{medio}	$F(x_{medio})$	Con- ver- ge?	Subs- titui a ou b ?
0	2,2	2,3	-1,352	0,167	2,25	-0,6094	não	a
1	2,25	2,3	-0,6094	0,167	2,275	-0,2255	não	a
2	2,275	2,3	-0,2255	0,167	2,288	-0,0303	não	a
3	2,2875	2,3	-0,0303	0,167	2,29375	0,068081787109375	nao	b
4	2,2875	2,29375	-0,0303	0,06808	2,290625	0,018824371337891	nao	b
5	2,2875	2, 290625	0,0303	0,0188243713	2,2890625	-0,00575399398804	sim	paro

Iteração 0: com os valores de intervalos iniciais: $a=2,2$ e $b=2,3$ tem-se $F(a)=f(2,2)=[(2,2)^3 - 12] = -1,352$, $F(b)=f(2,3)=[(2,3)^3 - 12] = 0,167$, $x_{medio} = 2,25$, $f(x_{medio}) = -0,609375$, $tolerância = 0,01 \Rightarrow$ ou $|f(x_{medio})| = 0,609375 > 0,01$ não converge! $F(a)$ tem mesmo sinal de $F(a)$ - será colocado no lugar de a

Iteração 1: com os valores de intervalos : $a=2,25$ e $b=2,3$ tem-se $x_{medio} = 2,275$ e $f(x_{medio}) = f(2,275) = (2,275)^3 - 12 = -0,225453125$ e $|f(2,275)| = 0,2254.... > 0,01$. não converge! $F(x_{medio})$ tem mesmo sinal de $F(a)$ - será colocado no lugar de a

Iteração 2: com os valores de intervalos : $a=2,275$ e $b=2,3$ tem-se $x_{medio} = 2,2875$ e $f(x_{medio}) = f(2,2875) = (2,2875)^3 - 12 = -0,030298828125$ e $|f(2,2875)| = 0,030298828125.... > 0,01$. não converge! $F(x_{medio})$ tem mesmo sinal de $F(a)$ - será colocado no lugar de a

Iteração 3: com os valores de intervalos : $a=2,2875$ e $b=2,3$ tem-se $x_{\text{medio}} = 2,29375$ e $f(x_{\text{medio}}) = f(2,29375) = (2,29375)^3 - 12 = 0,068$ e $|f(2,29375)| = 0,068... > 0,01$ ainda não converge! e $F(x_{\text{medio}})$ tem mesmo sinal de $F(b)$ - será colocado no lugar de b

Iteração 4: com os valores de intervalos : $a=2,2875$ e $b=2,29375$ tem-se $x_{\text{medio}} = 2,290625$ e $f(x_{\text{medio}}) = f(2,290625) = (2,290625)^3 - 12 = 0,0188243713$ e $|f(2,290625)| = 0,018... > 0,01$. não converge! mas $F(x_{\text{medio}})$ tem mesmo sinal de $F(b)$ - será colocado no lugar de b

Iteração 5: com os valores de intervalos : $a=2,2875$ e $b=2,290625$ tem-se $x_{\text{medio}} = 2,2890625$ e $f(x_{\text{medio}}) = f(2,2890625) = (2,2890625)^3 - 12 = -0,006$ e $|f(2,2890625)| = -0,006... < 0,01$ converge! e pode-se para o processo

Resposta , além da tabela preenchida :

Pelo M. da Bissecção foram preciso 5 iteração , considerando o intervalo inicial escolhido a iteração zero!

Critério de correção M. da Bissecção:

a) Esse método tem valor 3

b) o aluno deve verificar a convergência na $f(x)$, a função deve ser zerada, neste caso se $|f(x)| < 0,01$. O Aluno **perde 0,5** se não parou na nesta convergência. **Para antes -0,5**

c) parar ou fazer verificação apenas em relação a valores em x , não é sempre verdade, pois existem funções e mesmo polinômios mal condicionados. Aluno **perde 0,5** se apenas verificou em x

d) era pedido uma proposta de tabela. Aluno **perde 0,5** se não fez isso.

e) deveriam ser feito algumas iterações para mostrar que sabe mesmo fazer. **Apenas explicar=** o aluno **perder 0,5**

f) usando $[f(a)+f(b)]/2$ e não $[(a)+(b)]/2$ para as iterações **perde 2 pontos**, isso é o própria definição de Bissecção

g) errou nas substituição dos limites, ouno motivo que levaria asubstituir, **perder 0,5** cada vez que isso foi feito.

Solução pelo Método das tangentes (ou de Newton-Raphson):

$$f(x) = x^3 - 12 \Rightarrow f'(x) = 3x^2 \Rightarrow x_{i+1} = x_i - [x_i^3 - 12] / [3x_i^2]$$

$$\text{desenvolvendo : } x_{i+1} = \{3x_i^3 - [x_i^3 - 12]\} / [3x_i^2]$$

$$x_{i+1} = [2x_i^3 + 12] / [3x_i^2]$$

chutando $x_0=2$

$$x_1 = [2(2)^3 + 12] / [3(2)^2] = (16 + 12) / 12 = 28 / 12 = 2,333....$$

$$f(x_1) = x_1^3 - 12 = 0,703703703703 > 0.01 \text{ não atingiu a tolerância}$$

$$x_2 = [2(2,3333)^3 + 12] / [3(2,3333)^2] = 37,407407407407407407... / 16,3333333333... = 2,2902494....$$

$$f(x_2) = x_2^3 - 12 = 0,01291306295198596978 > 0.01 \text{ não atingiu a tolerância}$$

$$x_3 = [2(2,2902494)^3 + 12] / [3(2,2902494)^2] = [24,025826125903971939568 + 12] / 15,73572694260108 = 2,2894287793194881488109799245087$$

$$f(2,2894287793194881488109799245087) = (2,2894287793194881488109799245087)^3 - 12 = 4,6263349604076027036595791248173 \times 10^{-6} < \text{tolerância} \Rightarrow \text{convergiu!!!}$$

Pelo M. das tangentes foram preciso 3 iterações , considerando o valor inicial a iteração zero!

Critério de correção M. das tangentes:

- a) Esse método tem **valor 2** . Não foi feito **(- 2)**
- b) como não é possível divisão por zero, esse chute inicial não pode ocorrer. **(Zero neste item se foi feito)**
- c) o aluno montou a equação de iteração errada: **menos 1,5**
- d) verificou a convergência na diferença entre os x e não em $f(x)$ - **menos 0,5 cada vez que for feito**
- e) não deixa claro onde verificou a convergência -> **menos 0,5 cada vez que for feito**
- f) o aluno não montou a equação de iteração mas apenas responde : **menos 2**

2 – **(2)** Considere **3 números decimais** formados, cada um **agrupando 3 dígitos de sua matrícula**. Por exemplo, se seu número fosse: 319 005 112 . Esses seriam na base decimal $(319)_{10}$, $(005)_{10}$ e $(112)_{10}$

Represente esses números na forma normalizada de pontos flutuantes, na base decimal. Suponha que a máquina que você está usando considera apenas **3 algarismos** significativos na **mantissa** nas operações, com **truncamento** dos demais e com **expoentes de -2 a 2**: **F(10, 3, -2, 2)**. Com essas considerações **faça as operações** :

a) $(a+10b)-10c$ (*i.e.* some o primeiro número com o segundo multiplicado por 10 e depois de ter o resultado , subtraia do resultado o terceiro numero multiplicado por 10)

b) $a+10(b-c)$ (*i.e.* inicie pela subtração dos dois últimos números, depois a multiplicação do resultado por 10 e finalmente esse resultado for somado com o primeiro número).

Considere todas as restrições que devem ser usadas nesta notação F(10, 3, -2, 2) e as regras de soma e subtração dos números de ponto flutuantes. Verifique se o resultado será o mesmo , *i.e.* se o resultado da troca de ordem na operação de fato (como matematicamente deveria ser) não altera o resultado. **(1,0 -> 0,5 por cada operação feita corretamente na ordem pedida na questão)**

Explique como você pode transformar o resultado em binários. Finalmente como são esses **números representados em binário** ? **(1,0 -> 0,5 pela explicação e 0,5 pela transformação correta)**

OBS.

Estamos considerando aqui a definição de PONTO FLUTUANTE NORMALIZADO apresentado na página 26 do livro texto distribuído aos alunos em pdf. Esse livro está no cite do curso e corresponde a referência [1]

F. G. Muniz Cunha, J. K. de Sousa Castro, [Calculo Numerico](#), Ministerio da Educacao, U A B, 2010

Também foi incluído no material que se encontra do <Sala de Aula> e usado nos pdf da professora e nos videos recomendados, gravados pelos professores do IC durante a pandemia. Mais especificamente ao vídeo 5 do Tema 1: [Representando numeros de ponto flutuante](#)

Solução:

Para o texto não ficar muito sobrecarregado todos os números estão na base dez mesmo em o subscrito 10 entre parênteses, será indicado quando passar a base 2 , pelo 2 como subscrito depois do parênteses.

Primeiro vamos usar a notação normalizada $F(10, 3, -2, 2)$ nos decimais:

$(319)_{10} = 0,319 \times 1000 = 0,319 \times 10^3$ mas aplicando a restrição do expoente apenas poder estar entre -2 e 2 $\Rightarrow 0,319 \times 10^2$ esse número teve perda na representação $F(10, 3, -2, 2)$, ficando 10 vezes menor.

$(005)_{10} = 0,5 \times 10$ esse numero não tem perda na representação $F(10, 3, -2, 2)$

$(112)_{10} = 0,112 \times 1000 = 0,112 \times 10^3$ mas aplicando a restrição do expoente apenas poder estar entre -2 e 2

$\Rightarrow 0,11 \times 10^2$ Logo esse número teve perda na representação $F(10, 3, -2, 2)$, ficando 10 vezes menor.

Fazendo agora as Operações nas ordens pedidas:

$$a) (a+10b)-10c = (0,319 \times 10^2 + 10 \times 0,5 \times 10) - 10 \times 0,11 \times 10^2$$

$$\text{calculando o parênteses: } \Rightarrow 0,319 \times 10^2 + 10 \times 0,5 \times 10 = 0,319 \times 10^2 + 0,5 \times 10^2 = 0,81 \times 10^2$$

Nesta parte , nesta operação, não houve perda adicional.

Continuando, isto é, diminuindo esse resultado do último numero $0,11 \times 10^2$:

$$0,81 \times 10^2 - 0,11 \times 10^2 = 0,70 \times 10^2$$

Resultado de a): devido aos limites -2 e 2 do expoente na ordem a) o resultado fica: $0,7 \times 10^2$

b) trocando a ordem : $a+10(b-c) =$

$$0,319 \times 10^2 + 10 (0,5 \times 10 - 0,11 \times 10^2) = 0,319 \times 10^2 + 10 (0,05 \times 10^2 - 0,11 \times 10^2) =$$

$$0,319 \times 10^2 + 10 (- 0,06 \times 10^2) = 0,319 \times 10^2 - (0,6 \times 10^2) = 0,319 \times 10^2 - 0,600 \times 10^2 = - 0,281 \times 10^2$$

devido aos limites -2 e 2 do expoente nesta ordem o resultado fica: $- 0,281 \times 10^2$

Resposta ao pedido explicitamente na questão: o resultado **não foi o mesmo** , i.e. a troca de ordem na operação alterou o resultado.

A verdade matemática não foi preservada .

Explicando como transformar um valor da base 10 em binários= base 2.

A **parte antes da virgula** (inteira) de um número representado na base decimal pode ser transformada na base binária por divisões sucessivas até ser possível. Considera-se o valor final e os restos, descritos na ordem do menos significativo para o mais significativo (ou na ordem inversa da divisão, pois o resto associado a cada vez mais divisão estará em uma potência de 2 maior e na primeira divisão já se descobre se o número é par ou ímpar, o que em binário será representado pelo valor menos significativo).

Os resultados foram: $0,69 \times 10^2 = 69 = (1000101)_2$

pois: $69 / 2 = 34$ com resto **1** (será o menos significativo em binário, um numero ímpar)

$$34 / 2 = 17 \text{ com resto } \mathbf{0}$$

$$17 / 2 = 8 \text{ com resto } \mathbf{1}$$

$$8 / 2 = 4 \text{ com resto } \mathbf{0}$$

$$4 / 2 = 2 \text{ com resto } \mathbf{0}$$

$$2 / 2 = \mathbf{1} \text{ com resto } \mathbf{0}$$

esse valor final será o mais significativo em binário. seguido dos restos ate a primeira divisão

$$(1000101)_2 = 69 = (69)_{10}$$

Agora o segundo resultado :

$-0,381 \times 10^2 = -38,1$ o sinal pode ser representado apenas pelo primeiro bit estar ligado.

Assim considerando o restante do numero $= > 38 + 0,1$

parte inteira: 38 de novo se transforma fazendo divisões sucessivas por 2:

$$38 / 2 = 19 \text{ com resto } \mathbf{0}$$

$$19 / 2 = 9 \text{ com resto } \mathbf{1}$$

$$9 / 2 = 4 \text{ com resto } \mathbf{1}$$

$$4 / 2 = 2 \text{ com resto } \mathbf{0}$$

$$2 / 2 = \mathbf{1} \text{ com resto } \mathbf{1}$$

1 não divide por 2 para, e esse como sofre mais divisão será o mais significativo $-38 = -(110110)_2$

$$-0,38 \times 10^2 = -38 = (-110110)_2$$

A **parte depois da virgula** de um número representado na base decimal , pode ser transformada na base binária por multiplicações sucessivas por 2, enquanto o resultado tiver uma parte depois da virgula, ou fracionaria. Considera-se a parte inteira resultante do processo e a parte fracionaria de forma diferente. A parte inteira formará cada valor depois da virgula, iniciando do mais significativo para o de menor valor. A parte fracionaria deve ser usada para aumentar o número de aproximações , enquanto não resultar em um seqüência que se repita (periodicamente).

O valor **0,1** (foi inclusive avo do primeiro trabalho) representa uma valor que se repete infinitamente na base 2, pois resulta na tabela abaixo.

Multiplic.	Resultado	Parte inteira	Parte fracionária
0.1×2	0.2	0	0.2
0.2×2	0.4	0	0.4
0.4×2	0.8	0	0.8
0.8×2	1.6	1	0.6
0.6×2	1.2	1	0.2
0.2×2	0.4	0	0.4
0.4×2	0.8	0	0.8
0.8×2	1.6	1	0.6
0.6×2	1.2	1	0.2
0.2×2	0.4	0	0.4
0.4×2	0.8	0	0.8
0.8×2	1.6	1	0.6
0.6×2	1.2	1	0.2

como ainda tem parte fracionária deveria continuar mas note que depois do primeiro 0 a seqüência 0011 se repete 3 vezes e deve continuar assim infinitamente, isso indica que esse número não tem uma representação finita na base 2 ou seja

$$0.1_{10} = 0.000110011\underline{0011}_2$$

Finalmente somando as duas formas de conversão

$$\mathbf{-0,381 \times 10^2} = -38,1$$

parte, sinal + intera + fracionaria em binários

$$(38)_{10} + (0.1)_{10} = (-110110)_2 + (0.000110011\underline{0011})_2 = -(\mathbf{110110.000110011\underline{0011}})_2$$

Resposta os resultados obtidos na base 2 são: $(1000101)_2$ e $(\mathbf{110110.000110011\underline{0011}})_2$

Critério de correção:

- a) Não usando representação normalizada da referencia [1] do curso mencionada acima **perde 1,0**
- b) respostas corretas mas contas erradas **perde 0, 5 por conta errada**
- c) passou para binário um numero que não é uma solução sua **perde 0, 5 por numero <inventado>**

3 – (3) Imagine que **3 estagiários (E)** medindo o **raio externo (F)** de uma arruela e a **espessura (P)** tiveram os valores tabelados. Qual deve ser **a espessura e o raio** mais provável? **(ate 1)** Qual **deles fez melhor a tarefa** pedida **(ate 2)**?

Estag.	F1	F2	F3	F4	F5	P1	P2	P3	P4	P5
E1	11,51	11,16	11,31	11,42	11,35	0,01	0,03	0,02	0,04	0,01
E2	11,49	11,51	11,48	11,52	11,53	0,02	0,01	0,03	0,02	0,01
E3	11,53	11,55	11,51	11,54	11,52	0,02	0,03	0,04	0,02	0,01

Solução:

Observando todos os valores nota-se que tem uma medida, a F2, do E1 que está bem fora dos valores das demais.

Para não interferir na media essa deveria ser retirada do conjunto das medias de F, se fazendo 14 somas, sem esse valor. **(0,5 para soma com 14 valores e resposta em 2 casa decimais)**

Fora isso poder-se-ia considerar o valores mais prováveis como as medias aritméticas simples de todas as medidas.

Se isso **não for** observado tem-se como raio=**11,46** e a **espessura e 0,02** como mais prováveis (em responder formalmente + **0,1**)

Depois, devem ser calculados 6 valores de erros, 2 para cada E. Depois transformados em 2 tipos de erros relativos para cada um dos 3 E. Os valores das medidas F e P apenas podem ser combinados se estiverem como relativos pois tem ordens de grandeza diferentes e se referem a outros tipos de valores. Esses erros relativos de cada E devem ser comparados. O funcionário que tiver o menor erro relativo **é que terá feito melhor a tabela pedida**

Assim: Considerando que não se sabe os tamanhos corretos, e que todas as medidas deste tipo tem a mesma possibilidade de estarem corretas, o valor mais provável seria a média simples de todos eles, ou seja $F = 171,93/15 = 11,462$ **(0,4)**. Idem para as espessuras sendo $P = 0,32/15 = 0,021333$ **(0,4)**.

Obs. 1- Reparem que não faz sentido se ter/usar mais **valores de casa decimais** do que a precisão das medidas. Os valores adicionais são apenas resultado da conta e não transmitem a realidade das medições.

Assim se deve calcular o erro relativo de cada E em relação a essa F e a essa P separadamente .

para as F:

$$E1: |11,51-11,46| + |11,16-11,46| + |11,31-11,46| + |11,42-11,46| + |11,35-11,46| = 0,65$$

$$E2: |11,49-11,46| + |11,51-11,46| + |11,48-11,46| + |11,52-11,46| + |11,53-11,46| = 0,23$$

$$E3: |11,53-11,46| + |11,55-11,46| + |11,51-11,46| + |11,54-11,46| + |11,52-11,46| = 0,35$$

para as P:

$$E1: |0,01 - 0,02| + |0,03 - 0,02| + |0,02 - 0,02| + |0,04 - 0,02| + |0,01 - 0,02| = 0,05$$

$$E2: |0,02 - 0,02| + |0,01 - 0,02| + |0,03 - 0,02| + |0,02 - 0,02| + |0,01 - 0,02| = 0,03$$

$$E3: |0,02 - 0,02| + |0,03 - 0,02| + |0,04 - 0,02| + |0,02 - 0,02| + |0,01 - 0,02| = 0,04$$

Relativizando, se tem para as F:= $E_i / 11,46$

erros	relativo
0,65	0,056719
0,23	0,02007
0,35	0,030541

E para as P: $E_i / 0,02$.

Repare que como as espessuras são muito pequenas ao se transformarem em erros relativos, elas acabam sendo muito maiores que os valores relativos das F .

Isso porque na verdade por serem muito pequenas, qualquer erro no limiar de avaliação do equipamento usado já produz algo muito grande.

erro	relativo	somando relativos
0,05	2,5	2,556719
0,03	1,5	1,52007
0,04	2	2,030541

de modo que o E que fez menor erro relativo foi (o que errou menos em ambas as tarefas pedidas) :**E2**

O mais correto aqui seria na verdade fazer análises em cima de 3 critérios diferentes:

- 1- Pela erro feito na medida F2 por E1, já se nota que esse estagiário é mais <distraído que os outros>;
- 2- Considerando os erros em separado das medidas F e P se nota que m maior erro em ambas foi de E1 e o menor de E2 (**melhor E2**)
- 3- Considerando os erros relativos de P e F somados, E podem ser somados, também se tem que **E1 é o pior e E2 o melhor.**

Mas todas indicam na verdade o mesmo pior e melhor Estagiário.

Critério de correção:

- a) cada resposta correta faz o aluno receber **0,4** pontos nesta questão;
- b) **desconto de 0,1** se foram usadas mais casas decimais na resposta: **0,3** pontos nesta questão se usado mais casa decimais.
- c) **recebe 0,1** pela resposta correta até o final e **0,1** por comentar que consideração fez para chegar a ela.
- d) se por algum motivo separou por medidas Fi ou Pi , pode ter no Maximo 0,8. nesta questão, pois não entendeu como se calcular o valor.
- e) Na avaliação dos erros não podem ser usados valores médios separados por estagiário pois isso faria se ter uma tendência na medida, usando apenas dados dele para o avaliar. Se fez isso tem o valor **0** no restante da questão
- f) Cada valor de erro correto = **0,3** => $6 \times 0,3 = 1,8$. Transformados em 2 relativos corretos para cada E: **0,2=> 0,2+1,8=2,0 (a resposta do melhor estagiário relacionado a menor erro relativo fica com ponto na parte anterior da questão) .**
- g) cada resposta ou valor errado desconta-se **0,1**