Lista de exercícios 2

Questão 1 (Burden et al. (2017), Exercícios 1.2(2)). Calcule os erros absoluto e relativo de cada aproximação $\tilde{p}$ para o valor exato p, dados a seguir.

(a) $p = e^{10}$ e $\tilde{p} = 22.000$
(b) $p = 10^{\pi}$ e $\tilde{p} = 1.400$
(c) $p = 8!$ e $\tilde{p} = 39.900$
(d) $p = 9!$ e $\tilde{p} = \sqrt{18\pi}\,(9/e)^9$

(a) $EA(\tilde{p}) = |22000 - e^{10}| \approx 26{,}466$

$ER(\tilde{p}) = \dfrac{EA(\tilde{p})}{|p|} \approx 1{,}20 \times 10^{-3}$

(b) $EA(\tilde{p}) = |1400 - 10^{\pi}| \approx 14{,}544$

$ER(\tilde{p}) \approx 1{,}050 \times 10^{-2}$

(c) $EA(\tilde{p}) = |39900 - 8!| \approx 420$

$ER(\tilde{p}) \approx 1{,}042 \times 10^{-2}$

(d) $EA(\tilde{p}) = |\sqrt{18\pi}\,(9/e)^9 - 9!| \approx 3343{,}127$

$ER(\tilde{p}) \approx 9{,}213 \times 10^{-3}$

Questão 2 (Burden et al. (2017), Exercícios 1.2(4)). Determine o maior intervalo no qual $\tilde{p}$ deve estar contido a fim de aproximar p com quatro dígitos significativos exatos para cada valor de p abaixo.

(a) π (b) e (c) $\sqrt{2}$ (d) $\sqrt[3]{7}$

EM CADA ITEM, DEVEMOS RESOLVER:

$$\frac{|\tilde{p}-p|}{|p|} \leq 5\times10^{-4} \Rightarrow |\tilde{p}-p| \leq |p|5\times10^{-4}$$

(a) $3{,}14002 \leq \tilde{p} \leq 3{,}14316$

(b) $2{,}71692 \leq \tilde{p} \leq 2{,}71964$

(c) $1{,}41351 \leq \tilde{p} \leq 1{,}41492$

(d) $1{,}91197 \leq \tilde{p} \leq 1{,}91389$

OBSERVAÇÃO. DO ITEM (a), TEMOS QUE:

$$|\tilde{p}-\pi| \leq 0{,}00157\ldots = 1{,}57\ldots\times10^{-3} \leq 5\times10^{-3}$$

O QUE NOS FARIA CONCLUIR, POR EXEMPLO, QUE $\tilde{p} = 3{,}14002$ POSSUI 3 DÍGITOS EXATOS À DIREITA DA VÍRGULA. NO ENTANTO, LEVANDO EM CONSIDERAÇÃO O ERRO RELATIVO, $\tilde{p}$ POSSUI 4 DÍGITOS SIGNIFICATIVOS EXATOS (À DIREITA DA VÍRGULA).

Questão 3. Para cada função a seguir, (i) aproxime o valor de $f(a)$ usando o polinômio de Taylor de primeira ordem para f definido em torno de x_0, (ii) calcule o erro relativo correspondente e (iii) obtenha um limitante para o erro de truncamento dessa aproximação para valores de x em um intervalo de tamanho unitário centrado em a.

(a) $f(x) = \ln x$; $a = 1{,}5$; $x_0 = 1$.
(b) $f(x) = \sqrt{x}$; $a = 9{,}5$; $x_0 = 9$.

(a)

(i)

$$P_1(x) = f(1) + f'(1)(x-1) = x-1$$

(ii)

$$P_1(1{,}5) = 0{,}5$$

$$ER(0{,}5) = \frac{|0{,}5 - \ln 1{,}5|}{|\ln 1{,}5|} \approx 2{,}3315 \times 10^{-1}$$

(iii) INTERVALO $\Rightarrow$ $[1{,}5 - 0{,}5 ; 1{,}5 + 0{,}5] = [1, 2]$

$$|P_1(x) - f(x)| = \frac{1}{2} \cdot |f''(\xi) \cdot (x - x_0)^2|$$

$$= \frac{1}{2} \cdot \frac{1}{\xi^2} \cdot (x-1)^2 \quad , \quad 1 \leq x \leq 2$$

$$\leq \frac{1}{2}$$

(b)

(i)

$$P_1(x) = f(9) + f'(9)(x-9) = 3 + \frac{1}{6}(x-9)$$

(ii)

$$P_1(9{,}5) = 3 + \frac{1}{12} = \frac{37}{12} \approx 3{,}0833$$

$$ER(9{,}5) = \frac{\left|37/12 - \sqrt{9{,}5}\right|}{\left|\sqrt{9{,}5}\right|} \approx 3{,}65 \times 10^{-4}$$

(iii) INTERVALO $\Rightarrow$ $[9{,}5-0{,}5 ; 9{,}5+0{,}5] = [9,10]$

$$|P_1(x) - f(x)| = \frac{1}{2} \cdot |f''(\xi) \cdot (x-x_0)^2|$$

$$= \frac{1}{2} \cdot \frac{1}{4\,\xi^{3/2}} \cdot (x-9)^2 \quad , \quad 9 \leq x \leq 10$$

$$\leq \frac{1}{8} \cdot \frac{1}{3^3} \approx 4{,}63 \times 10^{-3}$$

Questão 4 (Burden et al. (2017), Exercícios 1.1(18))**.** Sejam $f(x) = (1-x)^{-1}$, com $x < 1$, e $P_n(x)$ seu polinômio de Taylor de ordem n definido em torno de $x_0 < 1$.

(a) Determine a expressão de $P_n(x)$ para $x_0 = 0$.
(b) Determine o valor de n necessário para que $P_n(x)$ aproxime $f(x)$ com precisão de 10^{-6} para $x \in \left[0, \frac{1}{2}\right]$.

(a) $f(x) = 0!(1-x)^{-1}$

$f'(x) = 1!(1-x)^{-2}$

$f''(x) = 2!(1-x)^{-3}$

$f'''(x) = 3!(1-x)^{-4}$

$f^{(4)}(x) = 4!(1-x)^{-5}$

$\vdots$

$f^{(n)}(x) = n!(1-x)^{-(n+1)}$

$f^{(n+1)}(x) = (n+1)!(1-x)^{-(n+2)}$

SUBSTITUINDO $x = x_0 = 0$:

$f(x) = 0!$

$f'(x) = 1!$

$f''(x) = 2!$

$f'''(x) = 3!$

$f^{(4)}(x) = 4!$

$\vdots$

$f^{(n)}(x) = n!$

Assim,

$$P_n(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)(x-x_0)^2}{2!} + \frac{f'''(x_0)(x-x_0)^3}{3!} + \cdots + \frac{f^{(n)}(x_0)(x-x_0)^n}{n!}$$

$$= 0! + 1!x + \frac{\cancel{2!}}{\cancel{2!}}x^2 + \cdots + \frac{\cancel{n!}}{\cancel{n!}}x^n$$

$$= 1 + x + x^2 + \cdots + x^n$$

$$= \sum_{j=0}^{n} x^j$$

(b) OBSERVE QUE $P_n(x)$ É UMA PROGRESSÃO GEOMÉTRICA DE RAZÃO $q = x$ E ELEMENTO INICIAL $a_1 = 1$. ENTÃO:

$$P_n(x) = \frac{a_1(1-q^{n+1})}{1-q} = \frac{1-x^{n+1}}{1-x}$$

ASSIM, O ERRO DE TRUNCAMENTO FICARÁ:

$$R_n(x) = P_n(x) - f(x) = \frac{1-x^{n+1}}{1-x} - \frac{1}{1-x} = -\frac{x^{n+1}}{1-x}$$

IREMOS CALCULAR O MÁXIMO DA EXPRESSÃO ACIMA, AO INVÉS DE UTILIZAR A FÓRMULA DE LAGRANGE.

PARA $0 \leqslant x \leqslant \frac{1}{2}$, SABEMOS QUE:

$$|R_n(x)| = \left| -\frac{x^{n+1}}{1-x} \right| = \frac{x^{n+1}}{1-x} \qquad (*)$$

DERIVANDO $(*)$:

$$\left(\frac{x^{n+1}}{1-x} \right)' = \frac{x^n(-nx+n+1)}{(1-x)^2}$$

PARA $0 \leqslant x \leqslant 1/2$:

$$\frac{\overbrace{x^n}^{\geqslant 0} \cdot \overbrace{(-nx+n+1)}^{\leqslant n+1,\ \geqslant \frac{n}{2}+1}}{\underbrace{(1-x)^2}_{>0}} \geqslant 0 \Rightarrow \text{O MÁXIMO DE } \frac{x^{n+1}}{1-x} \text{ OCORRE}$$

É $x = 1/2$.

PORTANTO,

$$|R_n(x)| \leqslant \left. \frac{x^{n+1}}{1-x} \right|_{x=1/2} \overset{(1/2)^{n+1}/1/2}{=} \frac{1}{2^n}$$

RESOLVENDO:

$$\frac{1}{2^n} \leqslant 10^{-6}$$

OBTEMOS $n \geqslant 19{,}93\ldots \Rightarrow \boxed{n = 20}$

Questão 5. Um sistema de ponto flutuante é dado por $\beta = 2$, $t = 2$, $e_{\min} = -1$, $e_{\max} = 1$.

(a) Liste todos os números desse sistema.
(b) Converta os números (i) 1/3, (ii) 2/3, (iii) 0,9 et (iv) 9,6 para esse sistema.

FORMA GERAL: $\pm(d_0, d_1)_2 \times 2^e$

(a) NORMAIS:

$\pm(1,0)_2 \times 2^{-1}$ $\quad \pm(1,0)_2 \times 2^{0}$ $\quad \pm(1,0)_2 \times 2^{1}$

$\pm(1,1)_2 \times 2^{-1}$ $\quad \pm(1,1)_2 \times 2^{0}$ $\quad \pm(1,1)_2 \times 2^{1}$

SUB-NORMAIS: $\pm(0,1)_2 \times 2^{-1}$

ZEROS: $\pm(0,0)_2 \times 2^{-1}$

(b)

(i)

$\frac{1}{3} \times 2 = \frac{2}{3} < 1 \Rightarrow 0 \times 2^0$

$\frac{2}{3} \times 2 = \frac{4}{3} \geq 1 \Rightarrow 1 \times 2^1$

-1

$\frac{1}{3} \times 2 = \frac{2}{3}$

$\frac{1}{3} = (0,010101\overline{01})_2$

$= (1,01\overline{01})_2 \times 2^{-2} \times \frac{2^{-1}}{2^{-1}}$

EXPOENTE INDISPONÍVEL

$= (0,1|01\overline{01})_2 \times 2^{-1}$

$\approx +(0,1)_2 \times 2^{-1}$

LOGO, $fl(1/3) = +(0,1)_2 \times 2^{-1}$.

(ii)

$$\frac{2}{3}\times 2 = \frac{4}{3} \geq 1 \Rightarrow 1\times 2^{-1}$$

$$\frac{1}{3}\times 2 = \frac{2}{3} < 1 \Rightarrow 0\times 2^{0}$$

$$\frac{2}{3}\times 2 = \frac{4}{3}$$

$$\frac{2}{3} = (0{,}101010\overline{10})_2$$

$$= (1{,}0|1010\overline{10})_2 \times 2^{-1}$$

$$\approx +(1{,}1)_2 \times 2^{-1}$$

LOGO, $fl(2/3) = +(1{,}1)_2 \times 2^{-1}$.

(iii)

$$0{,}9\times 2 = \underline{1}{,}8 \geq 1$$

$$0{,}8\times 2 = \underline{1}{,}6 \geq 1$$

$$0{,}6\times 2 = \underline{1}{,}2 \geq 1$$

$$0{,}2\times 2 = \underline{0}{,}4 < 1$$

$$0{,}4\times 2 = \underline{0}{,}8 < 1$$

$$0{,}8\times 2 = 1{,}6$$

$$\vdots$$

COM ISSO,

$$(0{,}9)_{10} = (0{,}1\overline{1100})_2$$

$$= (1{,}1|\overline{1100})_2 \times 2^{-1}$$

$$\approx (1{,}1+0{,}1)_2 \times 2^{-1}$$

$$= (10{,}0)_2 \times 2^{-1}$$

$$= (1{,}0)_2 \times 2^{0}$$

LOGO, $fl(0{,}9) = +(1{,}0)_2 \times 2^{0}$.

(iv) O LIMITE P/ OVERFLOW É:

$$+(1{,}1)_2\times 2^{1} + \frac{1}{2}\cdot(0{,}1)_2\times 2^{1} = 3 + 0{,}5 = 3{,}5$$

SE $x \geq 3{,}5$, ENTÃO $fl(x) = +\infty$.

LOGO, $fl(9{,}6) = +\infty$.

Questão 6 (Cheney e Kincaid (2008), Problemas 2.1)**.** Sabe-se que a operação aritmética de ponto flutuante $a \oplus (b \oplus c)$ pode ter o resultado diferente de $(a \oplus b) \oplus c$, isto é, a operação de adição não é associativa. Exemplifique isso por meio de um exemplo.

CONSIDERE O SISTEMA:

$$\beta = 10, \quad t = 2, \quad e_{mín} = 0, \quad e_{máx} = 1$$

TOMANDO

$$a = b = +(9,9)_{10} \times 10^1$$

$$c = -(9,9)_{10} \times 10^1$$

TEMEMOS:

$$(a \oplus b) \oplus c = +\infty$$

$$a \oplus (b \oplus c) = a \oplus 0 = +(9,9)_{10} \times 10^1$$

Questão 7. Realize as operações a seguir considerando um formato decimal normalizado com precisão de três dígitos, com o intervalo de expoentes válidos ilimitado. Calcule o erro absoluto e o erro relativo tomando o valor exato constituído de pelo menos cinco algarismos de precisão.

(a) $133 \oplus 0{,}921$
(b) $133 \ominus 0{,}499$
(c) $(121 \ominus 0{,}327) \ominus 119$
(d) $(121 \ominus 119) \ominus 0{,}327$

(a) $fl(133) = 1{,}33 \times 10^{2}$, $fl(0{,}921) = 9{,}21 \times 10^{-1}$

$133 \oplus 0{,}921 = fl(1{,}33 \times 10^{2} + 9{,}21 \times 10^{-1})$

$= fl(1{,}33|921 \times 10^{2}) = 1{,}34 \times 10^{2}$

VALOR EXATO $= 1{,}3392 \times 10^{2}$

EA $= 8 \times 10^{-2}$

ER $= 5{,}974 \times 10^{-4}$

(b) $133 \ominus 0{,}499 = fl(1{,}33 \times 10^{2} - 4{,}99 \times 10^{-1})$

$= fl(1{,}32|501 \times 10^{2}) = 1{,}33 \times 10^{2}$

VALOR EXATO $= 1{,}325 \times 10^{2}$

EA $= 5 \times 10^{-1}$

ER $= 3{,}7736 \times 10^{-3}$

(c) $(121 \ominus 0{,}327) \ominus 119 = 2{,}00 \times 10^{0}$

VALOR EXATO $= 1{,}6730$

$EA = 0{,}327$

$ER = 1{,}9546 \times 10^{-1}$

(d) $(121 \ominus 119) \ominus 0{,}327 = 1{,}67 \times 10^{0}$

EXATO $= 1{,}673 \times 10^{0}$

$EA = 3 \times 10^{-3}$

$ER = 1{,}7932 \times 10^{-3}$

Questão 8. O sistema de precisão dupla do IEEE é um sistema numérico binário caracterizado por ter uma precisão de 53 bits, $e_{\min} = -1022$ e $e_{\max} = 1023$. Neste contexto, é verdade que $1/3 \oplus 2/3$ é diferente de 1 quando usamos:

(a) arredondamento para o mais próximo?
(b) arredondamento em direção ao zero?

DA QUESTÃO 5, CONCLUÍMOS QUE:

$$fl(1/3) = +(1{,}0101\ldots01)_2 \times 2^{-2}$$

$$fl(2/3) = +(1{,}0101\ldots01)_2 \times 2^{-1}$$

COM ISSO,

$$fl(1/3) + fl(2/3) = +(\overbrace{1{,}111\ldots11}^{53\text{ BITS}}\,|\,1)_2 \times 2^{-1}$$

fl

MAIS PRÓXIMO

$$+(1{,}111\ldots11)_2 \times 2^{-1} \quad \downarrow +1$$

$$\shortparallel$$

$$+(1{,}000\ldots00)_2 \times 2^{0}$$

$$\equiv 1$$

DIREÇÃO AO ZERO

$$+(1{,}111\ldots11)_2 \times 2^{-1}$$

$$\equiv \sum_{i=1}^{53} 2^{-i} = \frac{2^{-1}\cdot(1-2^{-53})}{1-2^{-1}}$$

$$\approx 0{,}999$$

Questão 9 (Adaptado de REAMAT/UFRGS, Exemplo 2.7.2; Ascher e Greif (2011), Exercícios 2.5(15-16))**.**

(a) Calcule as raízes da equação:

$$x^2 + 300x - 0{,}014 = 0$$

aplicando a fórmula de Bhaskara, considerando um sistema decimal com precisão de seis dígitos.

(b) Sabendo que os valores exatos das raízes com seis dígitos são $\xi_2 = -3{,}00000 \times 10^2$ e $\xi_1 = 4{,}66667 \times 10^{-5}$, discuta o que pode ter ocorrido com o resultado do item (a).

(c) Proponha uma solução para contornar esta dificuldade, calculando os novos valores e seus respectivos erros relativos. [Dica: utilize as Equações 1.1, 1.2 e 1.3 do livro-texto.]

(d) Implemente a fórmula de Bhaskara e o método escolhido no item (c) em funções denominadas `bhaskara` e `proposto`, respectivamente. Essas funções devem ter como entrada números reais a, b e c em precisão dupla e retornar as raízes do polinômio quadrático $ax^2 + bx + c$. Avalie o desempenho das funções implementadas nos seguintes casos:

(i) $a = 1, b = -10^5, c = 1$.
(ii) $a = 6 \times 10^{30}$, $b = 5 \times 10^{30}$, $c = -4 \times 10^{30}$.
(iii) $a = 10^{-30}$, $b = -10^{30}$, $c = 10^{30}$.

(a) $\Delta = b \otimes b \ominus 4 \otimes a \otimes c = 9{,}00001 \times 10^4$

$\xi_1 = (-b \oplus \sqrt{\Delta}) \oslash (2 \otimes a) = +0{,}00000 \times 10^{e_{min}}$

$\xi_2 = (-b \ominus \sqrt{\Delta}) \oslash (2 \otimes a) = -3{,}00000 \times 10^2$

(b) COMO $b > 0$ E $-b \approx \sqrt{\Delta}$, A EXPRESSÃO PARA O CÁLCULO DE ξ TEVE UMA SUBTRAÇÃO DE NÚMEROS MUITO PRÓXIMOS, LEVANDO A UMA PERDA DE PRECISÃO.

(c) CONFORME DESCRITO NA REFERÊNCIA APONTADA, FAREMOS:

SE $b < 0$:

$$\xi_1 = (-b \oplus \sqrt{\Delta}) \oslash (2 \otimes a)$$

$$\xi_2 = (-2 \otimes c) \oslash (b \ominus \sqrt{\Delta})$$

SENÃO:

$$\xi_1 = (-2 \otimes c) \oslash (b \oplus \sqrt{\Delta})$$

$$\xi_2 = (-b \ominus \sqrt{\Delta}) \oslash (2 \otimes a)$$

P/ $a = 1$, $b = 300 \geqslant 0$ E $c = -0{,}014$, A RAIZ ξ_1 SERÁ:

$$\xi_1 = (-2 \otimes c) \oslash (b \oplus \sqrt{\Delta})$$

$$= -2{,}80000 \times 10^{-2} \oslash (3{,}00000 \times 10^{2} \oplus 3{,}00000 \times 10^{2})$$

$$= -4{,}66667 \times 10^{-5}$$

$$ER(\xi_1) \approx 8{,}69841 \times 10^{-7} \qquad ER(\xi_2) = 0$$

Questão 10 (Ascher e Greif (2011), Exercícios 3.6(2))**.** O polinômio $P(x) = (x-2)^9$ pode ser escrito em ao menos três formas distintas:

$$\begin{aligned}
P(x) &= (x-2)^9 \qquad (1)\\
&= x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 + 5376x^3 - 4608x^2 + 2304x - 512 \qquad (2)\\
&= -512 + x(2304 + x(-4608 + x(5376 + x(-4032 + x(2016 + \\
&\qquad x(-672 + x(144 + x(-18 + x)))))))). \qquad (3)
\end{aligned}$$

As duas primeiras expressões são conhecidas na álgebra como as formas *fatorada* e *expandida*. Já a última é a famosa *regra de Hörner*, bastante conhecida por quem lida com métodos numéricos.

(a) Implemente essas três formas de se calcular o valor de $P(x)$ em funções denominadas `fatorada`, `expandida` e `horner`. Elas devem ser avaliadas em 161 pontos regularmente espaçados no intervalo $[1{,}92; 2{,}08]$. Plote os resultados em três figuras separadas usando a `matplotlib`.

(b) Explique as diferenças observadas nos gráficos obtidos.