Matemática Computacional

MEBiol, MEBiom, MEFT, MEQ, MEM

1º Teste

Duração: 1 hora + 10 minutos   13/1/2021 – 11:30

Apresente todos os cálculos e justifique convenientemente todas as respostas.

1. Considere o sistema de ponto flutuante $\mathbb{F} = \mathrm{PF}(10, 6, -20, 20)$ com arredondamento simétrico.

(a)$_{[1.0]}$ Suponhamos que se pretende calcular a diferença $\sqrt{101} - \sqrt{100}$ no sistema $\mathbb{F}$. Com base na fórmula de propagação dos erros, quantos algarismos significativos terá o resultado? (Admita que o erro de arredondamento relativo cometido ao calcular a raíz é não superior à unidade de arredondamento de $\mathbb{F}$.)

(b)$_{[1.0]}$ Qual é o menor número inteiro positivo x, pertencente a $\mathbb{F}$, tal que $\mathrm{fl}(x+1) = \mathrm{fl}(x)$? Justifique.

2. Considere a sucessão definida pela fórmula

$$x_{n+1} = 2.5x_n(1 - x_n),\ n = 0, 1, \ldots$$

(a)$_{[1.5]}$ Com base no teorema do ponto fixo, justifique que esta sucessão é convergente, qualquer que seja $x_0 \in [0.35, 0.65]$.

Seja z o limite da sucessão considerada na alínea (a).

(b)$_{[1.0]}$ Determine z.

(c)$_{[1.5]}$ Sendo $x_0 = 0.65$, determine um majorante de $|z - x_{10}|$.

3.$_{[1.5]}$ Sejam $g \in C^1([a, b])$ e $0 < L < 1$ tais que $g([a, b]) \subset [a, b]$ e

$$-L \leq g'(x) < 0,\ \forall x \in [a, b].$$

Sendo $x_0 \in [a, b]$, mostre que a sucessão $x_{n+1} = g(x_n)$, $n = 0, 1, 2, \ldots$ satisfaz

$$|z - x_{n+1}| \leq \frac{L}{1+L}|x_{n+1} - x_n|, \quad \forall n \in \mathbb{N},$$

onde $z = \lim_{n\to+\infty} x_n$.

4. Considere o sistema linear

$$\begin{cases} ax_1 + x_2 + x_3 = 1 \\ x_1 + ax_2 = 1 \\ x_2 + ax_3 = 1 \end{cases}$$

onde $a \in \mathbb{R}$.

(a)$_{[1.0]}$ Indique condições sobre a que garantam que o método de Jacobi é convergente quando aplicado a este sistema.

(b)$_{[1.5]}$ Sendo $a = 2$ e admitindo que $\|x - x^{(0)}\|_1 \leq 0.1$, determine um majorante de $\|x - x^{(1)}\|_1$, onde $x^{(1)}$ é a primeira iterada do método de Gauss-Seidel e x é a solução exacta do sistema.

RESOLUÇÃO

1.

(a) A unidade de arredondamento neste sistema é $u = 0.5 \times 10^{-5}$.

Os números $x = 101$, $y = 100$ e $\sqrt{y}$ têm representação exata em $\mathbb{F}$. Sejam

$$z = \sqrt{x} - \sqrt{y}, \quad z_{\mathbb{F}} = \mathrm{fl}(\mathrm{fl}(\sqrt{x}) - \sqrt{y}) = 0.499000 \times 10^{-1} \qquad (t = -1).$$

Com base na fórmula de propagação dos erros, o erro relativo de $z_{\mathbb{F}}$ é dado por

$$\delta_{z_{\mathbb{F}}} = \frac{\sqrt{x}}{\sqrt{x} - \sqrt{y}} \delta_{arr_1} + \delta_{arr_2} = \frac{\sqrt{x}}{z} \delta_{arr_1} + \delta_{arr_2}$$

onde $\delta_{arr_1} = \delta_{\mathrm{fl}(\sqrt{x})}$ e, pelo enunciado, $|\delta_{arr_i}| \le u = 0.5 \times 10^{-5}$, $i = 1, 2$. Então o erro absoluto de $z_{\mathbb{F}}$ satisfaz

$$|e_{z_{\mathbb{F}}}| = z|\delta_{z_{\mathbb{F}}}| \le \sqrt{x}|\delta_{arr_1}| + z|\delta_{arr_2}| \le (\sqrt{x} + z)u = (2\sqrt{x} - \sqrt{y})u.$$

Agora há que procurar uma majoração da forma

$$|e_{z_{\mathbb{F}}}| \le 0.5 \times 10^{t-k} = 0.5 \times 10^{-1-k}$$

com $k \in \mathbb{N}$ o maior possível, que será o número de algarismos significativos obtido com base neste raciocínio. De

$$|e_{z_{\mathbb{F}}}| \le (2\sqrt{x} - \sqrt{y})u \approx 10.0998 \times 0.5 \times 10^{-5} \approx 0.0000505 \le 0.5 \times 10^{-1-2} = 0.5 \times 10^{t-2},$$

conclui-se que podemos garantir que o resultado tem, pelo menos, dois algarismos significativos. Aqui usámos o valor de $z_{\mathbb{F}}$, mas podíamos ter usado

$$z = 0.49875621... \times 10^{-1}.$$

Nota: Se não fosse permitido usar o valor de z acima calculado (num sistema com mais precisão do que o sistema $\mathbb{F}$), mas pretendêssemos estimar de forma mais rigorosa o valor de $2\sqrt{x} - \sqrt{y}$:

$$\sqrt{x} = \sqrt{y} + (x - y)\frac{1}{2\sqrt{\xi}}, \quad y < \xi < x$$

donde, para $y = 100$, $x = 101$ e $100 < \xi < 101$, se obtém

$$\sqrt{101} = 10 + \frac{1}{2\sqrt{\xi}} < 10 + \frac{1}{20} = 10.05,$$

$$\sqrt{101} - 10 < 0.5 \times 10^{-1}, \quad 2\sqrt{101} - \sqrt{100} = 2(\sqrt{101} - 10) + 10 < 10^{-1} + 10 = 10.1.$$

(b) Se $x = 10^6 = 1000000$ então é óbvio que $x \in \mathbb{F}$ e $\mathrm{fl}(x) = 0.100000 \times 10^7$. Nesse caso, $x + 1 = 1000001$ e

$$\mathrm{fl}(x + 1) = \mathrm{fl}(0.1000001 \times 10^7) = 0.100000 \times 10^7 = \mathrm{fl}(x).$$

Por outro lado, se x é inteiro e $0 < x < 1000000$, então $\mathrm{fl}(x + 1) \ne \mathrm{fl}(x)$, porque tanto x como $x + 1$ têm representação exacta no sistema $\mathbb{F}$.

Logo, a resposta é $x = 1000000$.

2.

(a) Verifiquemos as condições do teorema do ponto fixo.

- Sendo $g(x) := 2.5x(1-x)$, g é continuamente diferenciável em $I := [0.35, 0.65]$. Temos $g(0.35) = g(0.65) = 0.56875 \in I$.

  Por outro lado, $g'(x) = 2.5 - 5x$ anula-se em $x = 0.5$. Temos $g(0.5) = 0.625 \in I$, que é o valor máximo de g em I. Por conseguinte, podemos garantir que

  $$g(I) \subset I.$$

- Já vimos que $g'(x) = 2.5 - 5x$, logo g' é decrescente em I e o seu módulo atinge o máximo num dos extremos de I. Mais precisamente,

  $$\max_{x\in I} |g'(x)| = |g'(0.35)| = |g'(0.65)| = 0.75 = L.$$

  Logo $L < 1$, satisfazendo a outra condição do teorema do ponto fixo.

Estando satisfeitas ambas as condições do teorema do ponto fixo, está garantido que a sucessão considerada, com qualquer $x_0 \in I$, converge para o único ponto fixo da função g, localizado no intervalo I.

(b) Sabendo que a sucessão converge para um dos pontos fixos de g, comecemos por determinar esses pontos fixos:

$$g(z) = z \iff 2.5z(1-z) = z \iff z = 0 \vee z = 0.6.$$

Visto que $z = 0 \notin I$, então a sucessão converge para $z = 0.6$.

(c) Podemos basear-nos na estimativa do erro a priori

$$|z - x_{10}| \leq L^{10} \frac{|x_1 - x_0|}{1 - L}.$$

Dado que $x_1 = g(0.65) = 0.56875$ e que $L = 0.75$, temos

$$|z - x_{10}| \leq 0.75^{10} \frac{0.65 - 0.56875}{1 - 0.75} = 0.0183019.$$

Em alternativa, usando a estimativa

$$|z - x_{10}| \leq L^{10} |z - x_0|$$

e atendendo a que $|z - x_0| \leq |0.65 - 0.35| = 0.3$, obtemos

$$|z - x_{10}| \leq 0.75^{10} 0.3 = 0.0168941.$$

Também se poderia responder, tendo em conta que na alínea anterior se calculou $z = 0.6$. Nesse caso, com $x_0 = 0.65$, obtém-se

$$|z - x_{10}| \leq 0.75^{10} 0.05 = 0.002816.$$

3. Pelo enunciado do problema sabemos que g satisfaz as condições do teorema do ponto fixo, pelo que $x_n \to z$ e verifica-se

$$|z - x_{n+1}| \leq L|z - x_n|. \tag{1}$$

Por outro lado, sabemos que $g'(x) < 0$, para todo $x \in [a,b]$, pelo que a sucessão das iteradas é alternada. Nestas condições verifica-se $x_{n+1} < z < x_n$ ou $x_{n+1} < z < x_{n+1}$, para todo $n \in \mathbb{N}$. Em qualquer nos casos, temos

$$|x_n - x_{n+1}| = |z - x_n| + |z - x_{n+1}|$$

e portanto,

$$|z - x_n| = |x_n - x_{n+1}| - |z - x_{n+1}|. \tag{2}$$

Substituindo (2) em (1), concluímos que

$$|z - x_{n+1}| \leq L|x_n - x_{n+1}| - L|z - x_{n+1}|$$

donde resulta a estimativa pretendida

$$|z - x_{n+1}| \leq \frac{L}{1+L}|x_n - x_{n+1}|.$$

4.

(a) A matiz do sistema considerado tem a forma

$$A = \begin{bmatrix} a & 1 & 1 \\ 1 & a & 0 \\ 0 & 1 & a \end{bmatrix}.$$

Para que o método de Jacobi seja convergente basta que a matriz A tenha a diagonal estritamente dominante, por linhas ou por colunas. Para ter a diagonal estritamente dominante por linhas, basta que, em cada linha, o módulo do elemento da diagonal seja superior à soma dos módulos dos restantes elementos, ou seja, que se verifique

$$\begin{cases} |a| > 1 + 1 = 2, \\ |a| > 1. \end{cases}$$

Portanto, o método de Jacobi é convergente se tivermos $|a| > 2$.

(b) Para obter o majorante pretendido deve usar-se a estimativa do erro a posteriori:

$$\|x - x^{(1)}\|_1 \leq \|C_{GS}\|_1 \|x - x^{(0)}\|_1. \tag{3}$$

Para o sistema considerado, com $a = 2$, a matriz de iteração do método de Gauss-Seidel tem a forma

$$C_{GS} = -(L_A + D_A)^{-1} U_A = \begin{bmatrix} 0 & -1/2 & -1/2 \\ 0 & 1/4 & 1/4 \\ 0 & -1/8 & -1/8 \end{bmatrix}.$$

Daqui, $\|C_{GS}\|_1 = 1/2 + 1/4 + 1/8 = 7/8 = 0.875$. Finalmente, substituindo em (3), obtém-se

$$\|x - x^{(1)}\|_1 \leq 0.875 \times 0.1 = 0.0875.$$

Note-se que a matriz C_{GS} pode ser obtida sem calcular $(L_A + D_A)^{-1}$, a partir das fórmulas computacionais do método.