Análise Numérica - 2009/2010
Tópicos de resolução do 1º Teste

1. Temos que $b_2 = a_2 + a_3 x$ e que $b_1 = a_1 + z b_2$. Assim, temos para o erro relativo

$$\begin{aligned}
\varepsilon_{b_2} &= \frac{a_2}{a_2 + a_3 z}\varepsilon_{a_2} + \frac{a_3 z}{a_2 + a_3 z}\varepsilon_{a_3} + \frac{z a_3}{a_2 + a_3 z}\varepsilon_z \\
&= \frac{a_2}{a_2 + a3_z}\varepsilon_{a_2} + \frac{a_3 z}{a_2 + a_3 z}(\varepsilon_{a_3} + \varepsilon_z) \qquad (1) \\
\varepsilon_{b_1} &= \frac{a_1}{a_1 + z b_2}\varepsilon_{a_1} + \frac{z b_2}{a_1 + z b_2}(\varepsilon_{b_2} + \varepsilon_z) \\
&= \frac{a_1}{a_1 + a_2 z + a_3 z^2}\varepsilon_{a_1} + \frac{a_2 z}{a_1 + a_2 z + a_3 z^2}\varepsilon_{a_2} + \frac{a_3 z^2}{a_1 + a_2 z + a_3 z^2}\varepsilon_{a_3} \\
&\quad + \frac{z(a_2 + 2a_3 z)}{a_1 + a_2 z + a_3 z^2}\varepsilon_z \qquad (2)
\end{aligned}$$

$$\begin{aligned}
e_{b_2} &= e_{a_2} + z e_{a_3} + a_3 e_z \qquad (3) \\
e_{b_1} &= e_{a_1} + z e_{b_2} + b_2 e_z \\
&= e_{a_1} + z e_{a_2} + z^2 e_{a_3} + (a_2 + 2 z a_3) e_z \qquad (4)
\end{aligned}$$

Se os coeficientes do polinómio não estiverem afectados de erro, i.e $e_{a_3} = e_{a2} = e_{a_1} = e_{a_0} = 0$ e $\varepsilon_{a_3} = \varepsilon_{a2} = \varepsilon_{a_1} = \varepsilon_{a_0} = 0$ teremos

$$\begin{aligned}
\varepsilon_{b_2} &= \frac{z a_3}{a_2 + a_3 z}\varepsilon_z \\
\varepsilon_{b_1} &= \frac{z(a_2 + 2a_3 z)}{a_1 + a_2 z + a_3 z^2}\varepsilon_z
\end{aligned}$$

$$\begin{aligned}
e_{b_2} &= a_3 e_z \\
e_{b_1} &= (a_2 + 2 z a_3) e_z
\end{aligned}$$

2. a) Decompondo a matriz de sistema como $A = L + D + U$ e tendo em conta que neste caso temos $L = 0$, o método de Jacobi escreve-se

$$\begin{aligned}
x^{(k+1)} &= D^{-1}(b - U x^{(k)}) \\
&= D^{-1} b - D^{-1} U x^{(k)} \\
&= c + G x^{(k)}
\end{aligned}$$

em que $c = D^{-1}b$ e $G = -D^{-1}U$. De acordo com a sugestão, como G é triangular superior, podemos garantir que $G^m = 0$ para $m \geq n$, o que implica que $\lim_{n\to\infty} G^n = 0$. Assim, sabemos que o método de Jacobi é convergente, qualquer que seja a aprox. inicial $x^{(0)} \in \mathbb{R}^n$. Mais concretamente,

$$
\begin{aligned}
x^{(1)} &= c + Gx^{(0)} \\
x^{(2)} &= c + Gx^{(1)} = c + Gc + G^2x^{(0)} \\
x^{(3)} &= c + Gx^{(2)} = c + Gc + G^2c + G^3x^{(0)} \\
&\vdots \\
x^{(m)} &= c + Gx^{(m-1)} = c + Gd + \cdots + G^{m-1}c + G^m x^{(0)} \\
&= (I + G + G^2 + \cdots G^m)c
\end{aligned}
$$

Em particular isto significa que, independentemente da aproximação inicial, a sucessão $\left(x^{(m)}\right)$ é constante para $m \geq n$, pelo que a solução exacta é obtida no máximo em n iterações.

b) No caso particular do sistema proposto temos

$$
c = D^{-1}b = (1/2, 1, 1)^T, \quad G = -D^{-1}U = \begin{pmatrix} 0 & -3/2 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}
$$

Assim, de acordo com alínea anterior teremos

$$
x = x^{(3)} = (I + G + G^2)c = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}
$$

c) Com base na alínea anterior temos que $\|G\|_\infty = 5/2$, $\|G\|_1 = 3$. Estes resultados não contradizem as conclusões anteriores, já que o facto de estas normas não serem inferiores a 1 não implica a divergência do método.

3.
a) Escrevendo o sistema na forma $F(\vec{x}) = \vec{0}$ temos

$$
F(x, y, z) = \begin{pmatrix} -x^2 + y^2 + z^2 + 1 \\ 2xy + 2y^2 + z^2 + 3z - 3 \\ 2xz - yz + 2 \end{pmatrix},
$$

$$
J_F(x, y, z) = \begin{pmatrix} -2x & 2y & 2z \\ 2y & 2x + 4y & 2z + 3 \\ 2z & -z & 2x - y \end{pmatrix}.
$$

Designando $\mathbf{x} = (x, y, z)$, o método de Newton Generalizado consiste em, partindo de uma aproximação inicial $\mathbf{x}^{(0)}$, e enquanto se verificar a condição $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| > \varepsilon$ fazer :

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - J_F^{-1}(\mathbf{x}^{(k)})F(\mathbf{x}^{(k)}).$$

b) Para evitar calcular a matriz inversa é usual substituir o algoritmo apresentado na alínea a) por outro equivalente em que, para cada $k > 0$

i. Calcula-se $\mathbf{y}$, solução de $J_F(\mathbf{x}^{(k)})\mathbf{y} = -F(\mathbf{x}^{(k)})$.

ii. Faz-se $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{y}$.

Neste caso particular, para calcular $\mathbf{x}^{(1)}$ começamos por resolver o sistema linear

$$J_F(\mathbf{x}^{(0)})\mathbf{y} = -F(\mathbf{x}^{(0)}) \Leftrightarrow \begin{pmatrix} -1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}$$

A resolução este sistema linear pelo método de Crout consiste em fatorizar a matriz de sistema na forma A = L U, em que L é uma matriz triangular inferior e U é uma matriz triangular superior com diagonal unitária, procedendo-se depois à resolução dos sistemas triangulares por substituição ascendente ou descendente.

$$J_F(\mathbf{x}^{(0)}) = LU = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$L\mathbf{w} = -F(\mathbf{x}^{(0)}) \Leftrightarrow \mathbf{w} = (1, 1/3, -4)^T$$

$$U\mathbf{y} = \mathbf{w} \Leftrightarrow \mathbf{y} = (16/3, 13/3, -4)^T$$

Assim,

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{y} = (35/6, 29/6, -4)^T.$$

4.

a) Comecemos por notar que a equação pode ser escrita de modo equivalente na forma $x = g(x)$, com $g(x) = \cos(x)/2$. Vejamos que estão verificadas as condições de aplicação do teorema do ponto fixo à função g no intervalo $I = [0, \pi/4]$:

i. A função g é contínua em I

ii. Como g é decrescente em I, temos que $g(I) = [g(\pi/4), g(0)] = [\sqrt{2}/4, 1/2] \subset I$, pelo que $g(I) \subseteq I$.

iii. $|g'(x)| = |-\sin(x)/2| \leq \sqrt{2}/4 < 1$, pelo que g é contrativa em I.

Assim, o teorema do ponto fixo garante que a equação tem uma e uma só raíz em I e que que a sucessão $x_{k+1} = \cos(x_k)/2$ converge para essa raíz, qualquer que seja $x_0 \in I$. Além disso, como $x = 0, \pi/4$ claramente não são soluções, podemos afirmar que $\alpha \in]0, \pi/4[$.

b) Da alínea a) já sabemos que o método iterativo referido converge para α, qualquer que seja $x_0 \in [o, \pi/4]$. Trata-se apenas de verificar que a convergência é linear. As estimativas de erro do método do ponto fixo garantem uma convergência pelo menos linear. Para a ordem de convergência ser superior a 1, teríamos que ter $g'(\alpha) = 0$, o que não é possível já que $g'(\alpha) = -\sin(\alpha)/2$ e esta quantidade não se anula para nenhum $\alpha \in]0, \pi/4[$.

c) Devemos neste caso utilizar uma estimativa *a priori* , por exemplo

$$\begin{aligned}
|\alpha - x_{16}| &\leq L^{16}|\alpha - x_0| \\
&\leq \left(\sqrt{2}/4\right)^{16} \pi/8 \\
&\approx 2.341 \times 10^{-8}.
\end{aligned}$$