MATEMÁTICA COMPUTACIONAL

Cap. 5. Resolução Numérica de Sistemas Não-Lineares

Filipe J. Romeiras

Departamento de Matemática
Instituto Superior Técnico

Novembro de 2007

Apontamentos das aulas da disciplina do mesmo nome do 2º ano do
Mestrado Integrado em Engenharia Física Tecnológica

5. RESOLUÇÃO NUMÉRICA DE SISTEMAS NÃO-LINEARES

• Vamos agora considerar métodos iterativos para resolver sistemas de equações em $\mathbb{R}^n$,

$$f(x) = 0 \quad \Leftrightarrow \quad x = g(x),$$

onde

$$x = [x_1 \;\; x_2 \;\; \dots \;\; x_n]^T \in \mathbb{R}^n$$
$$f : D \subset \mathbb{R}^n \to \mathbb{R}^n, \qquad f(x) = [f_1(x) \;\; f_2(x) \;\; \dots \;\; f_n(x)]^T,$$
$$g : D \subset \mathbb{R}^n \to \mathbb{R}^n, \qquad g(x) = [g_1(x) \;\; g_2(x) \;\; \dots \;\; g_n(x)]^T.$$

Definição. Sendo $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$, chama-se **matriz Jacobiana** de f à matriz de derivadas parciais, caso existam,

$$J_f(x) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix} (x).$$

Método do ponto fixo

Definição. Diz-se que z é um **ponto fixo** de uma função $g : \mathbb{R}^n \to \mathbb{R}^n$ se e só se $z = g(z)$.

Definição. O **método do ponto fixo** é o método iterativo a um passo da forma

$$x^{(m+1)} = g(x^{(m)}), \quad m \in \mathbb{N}, \qquad x^{(0)} = \xi_0 \in \mathbb{R}^n \;\; \text{(dado)}.$$

Definição. Uma função $g : D \subset \mathbb{R}^n \to \mathbb{R}^n$, diz-se uma **função de Lipschitz** ou **Lipschitziana** em D se existir um número real $L \geq 0$ tal que

$$\|g(x) - g(y)\|_V \leq L\|x - y\|_V, \qquad \forall x, y \in D.$$

onde V designa qualquer norma em $\mathbb{R}^n$. Se $L < 1$ a função diz-se **contractiva** ou uma **contracção** em D, com **constante de contractividade** L.

Teorema do ponto fixo (em $\mathbb{R}^n$). Seja $g : D \to \mathbb{R}^n$ uma função contractiva num conjunto fechado $D \subset \mathbb{R}^n$ com constante de contractividade L e tal que $g(D) \subset D$. Então:

(1) g tem um e um só ponto fixo z em D;

(2) a sucessão $\{x^{(m)}\}_{m \in \mathbb{N}}$, definida por $x^{(m+1)} = g(x^{(m)})$, converge para z, qualquer que seja $x^{(0)} \in D$;

(3) verificam-se as seguintes estimativas de erro:

(i) $\|e^{(m+1)}\|_V \leq L\|e^{(m)}\|_V$;

(ii) $\|e^{(m)}\|_V \leq L^m\|e^{(0)}\|_V$;

(iii) $\|e^{(m)}\|_V \leq \dfrac{1}{1-L}\|x^{(m+1)} - x^{(m)}\|_V$;

(iv) $\|e^{(m+1)}\|_V \leq \dfrac{L}{1-L}\|x^{(m+1)} - x^{(m)}\|_V$;

(v) $\|e^{(m)}\|_V \leq \dfrac{L^m}{1-L}\|x^{(1)} - x^{(0)}\|_V$;

onde $e^{(m)} = z - x^{(m)}$, $\forall m \in \mathbb{N}$ e V designa qualquer norma em $\mathbb{R}^n$.

Dem.: $(\cdots)$

Definição. Um conjunto $D \subset \mathbb{R}^n$ diz-se **convexo** se

$$tx + (1-t)y \in D, \qquad \forall t \in [0,1], \qquad \forall x, y \in D,$$

isto é, se $x, y \in D$ então todos os pontos do segmento de recta que une x e y também pertencem a D.

Proposição. Seja $g \in C^1(D)$ onde D é um conjunto convexo em $\mathbb{R}^n$. Então se

$$\sup_{x\in D} \|J_g(x)\|_M \leq L < 1,$$

a função g é contractiva em D segundo a norma M associada à norma vectorial V em $\mathbb{R}^n$ com constante de contractividade L.

Proposição. Seja D um conjunto fechado e convexo em $\mathbb{R}^n$ e $g \in C^1(D)$ uma função tal que:

(i) $g(D) \subset D$; (ii) $\sup_{x\in D} \|J_g(x)\|_M \leq L < 1$.

Então são válidas as conclusões do teorema do ponto fixo.

Proposição (da convergência local). Seja $g \in C^1(V_z)$, onde V_z é uma vizinhança de z, ponto fixo de g, tal que $r_\sigma(J_g(z)) < 1$. Então o método do ponto fixo converge para z desde que $x^{(0)}$ esteja suficientemente perto de z. Além disso as estimativas de erro (i)-(v) são válidas em alguma bola fechada contida em V_z e que contenha z.

Exemplo. Considere o sistema de equações algébricas não-lineares

$$f(x) = 0, \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad f(x) = \begin{bmatrix} 3x_1 + x_2^2 \\ x_1^2 + 3x_2 - 1 \end{bmatrix}. \tag{S}$$

(a) Mostre que o sistema (S) tem uma e uma só solução z no conjunto

$$D = \left[-\frac{1}{3}, 0\right] \times \left[0, \frac{1}{3}\right].$$

(b) Mostre que o sistema (S) tem uma e uma só solução $\tilde{z}$ no conjunto

$$\tilde{D} = [-4, -2] \times [-4, -2].$$

(c) Utilize o método do ponto fixo para obter um valor aproximado da raiz z com um erro absoluto inferior a 10^{-6}.

(d) Mostre que o sistema (S) tem apenas duas soluções, z e $\tilde{z}$, em $\mathbb{R}^2$.

Método de Newton generalizado

Definição. O **método de Newton** é o método iterativo a um passo da forma

$$x^{(m+1)} = x^{(m)} - [J_f(x^{(m)})]^{-1} \cdot f(x^{(m)}), \quad m \in \mathbb{N}, \qquad x^{(0)} = \xi_0 \text{ (dado)}.$$

Na prática resolve-se em cada iteração o sistema linear

$$J_f(x^{(m)}) \cdot \Delta x^{(m)} = -f(x^{(m)}),$$

e define-se a iterada seguinte por

$$x^{(m+1)} = x^{(m)} + \Delta x^{(m)}.$$

Teorema de Kantorovich (condições suficientes de convergência).

Seja $D \subset \mathbb{R}^n$ um conjunto aberto e convexo e $f \in C^1(D)$. Suponhamos que para alguma norma em $\mathbb{R}^n$ e algum $x^{(0)} \in D$ são verificadas as seguintes condições:

(i) $\det(J_f(x)) \neq 0, \ \forall x \in D; \quad \exists M_1 > 0 : \|[J_f(x)]^{-1}\| \leq \dfrac{1}{M_1}, \ \forall x \in D;$

(ii) $\exists M_2 > 0 : \|J_f(x) - J_f(y)\| \leq M_2 \|x - y\|, \ \forall x, y \in D$;

(iii) sendo $\varepsilon_0 = 2\|x^{(1)} - x^{(0)}\|$ e $K = \dfrac{M_2}{2M_1}$, verifica-se a desigualdade $2K\varepsilon_0 < 1$;

(iv) $\bar{B}(x^{(0)}, \varepsilon_0) \subset D$.

Então:

(1) f tem um único zero $z \in \bar{B}(x^{(0)}, \varepsilon_0)$;

(2) a sucessão de Newton com condição inicial $x^{(0)}$ é bem definida, permanece em $\bar{B}(x^{(0)}, \varepsilon_0)$ e converge para z.

(3) verifica-se a estimativa de erro a priori

$$\|z - x^{(m)}\| \leq \frac{1}{K} (K\varepsilon_0)^{2^m}.$$

Nota. Se $f \in C^2(D)$ a condição (ii) pode ser substituída pela condição:

(ii)' $\exists M_2 > 0 : \|H_{f_i}(x)\| \leq M_2, \quad \forall x \in D, \ \forall i \in \{1, \ldots, n\}$,

onde f_i designa uma componente de f e H_{f_i} é a matriz Hessiana de f_i, com componentes

$$(H_{f_i})_{jk} = \frac{\partial^2 f_i}{\partial x_j \partial x_k}.$$

Notas.

- ⋄ A condição (i) implica a existência de inversa para a matriz Jacobiana (equivalente no caso real a $f'(x) \neq 0$), e serve ao mesmo tempo para definir M_1 (que corresponde no caso real a $\min |f'(x)|$).
- ⋄ A condição (ii) implica a limitação dos valores da matriz Hessiana (caso $f \in C^2$) e define M_2 (que corresponde no caso real a $\max |f''(x)|$).
- ⋄ A condição (iii) permite garantir que as iteradas vão permanecer na bola $\bar{B}(x^{(0)}, \varepsilon_0)$ (e corresponde no caso real à condição $\left|\frac{f(a)}{f'(a)}\right| \leq b - a, \quad \left|\frac{f(b)}{f'(b)}\right| \leq b - a$).
- ⋄ A condição (iv) é óbvia e podemos mesmo considerar $\bar{D} = \bar{B}(x^{(0)}, \varepsilon_0)$.

Proposição (da convergência local). Seja $f \in C^2(V_z)$, onde V_z é uma vizinhança de z, zero de f, tal que $\det(J_f(z)) \neq 0$. Então o método de Newton converge para z desde que $x^{(0)}$ esteja suficientemente perto de z.

Proposição (da ordem de convergência). Seja $f \in C^2(V_z)$, onde V_z é uma vizinhança de z, zero de f, tal que $\det(J_f(z)) \neq 0$. Então o método de Newton, quando converge para z, tem convergência pelo menos quadrática, isto é, existe $K > 0$ tal que

$$\|z - x^{(m+1)}\| \leq K\|z - x^{(m)}|^2, \qquad m \in \mathbb{N},$$

ou

$$\|z - x^{(m)}\| \leq \frac{1}{K} \left(K\|z - x^{(0)}|\right)^{2^m}, \qquad m \in \mathbb{N}.$$

Exemplo. Considere o sistema de equações algébricas não-lineares

$$f(x) = 0, \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad f(x) = \begin{bmatrix} 3x_1 + x_2^2 \\ x_1^2 + 3x_2 - 1 \end{bmatrix}. \tag{S}$$

para o qual se verificou que tinha apenas duas raízes em $\mathbb{R}^2$, z e $\tilde{z}$.

(a) Determine o valor aproximado da raiz z usando três iteradas do método de Newton generalizado com aproximação inicial $x = [0 \ \ 0]^T$.

(b) Obtenha uma estimativa do erro da solução aproximada obtida na alínea (a) (usando a norma do máximo).