Métodos Numéricos

Sistemas de equações não lineares

Teresa Monteiro

Departamento de Produção e Sistemas

Escola de Engenharia

Universidade do Minho

tm@dps.uminho.pt

Objetivos

Resolver sistemas de equações não lineares

Seja $F : \Omega \longrightarrow \Re^n$, em que Ω é um domínio em $\Re^n$. O objectivo é determinar as soluções do sistema de equações

$$F(x) = 0 \Leftrightarrow \begin{cases} f_1(x_1, \ldots, x_n) & = 0 \\ \quad\ldots & \ldots \\ f_n(x_1, \ldots, x_n) & = 0 \end{cases}$$

ou seja, o ponto $x^* = (x_1, \ldots, x_n)$ que verifica simultaneamente todas as n equações.

Exemplo em $\Re^2$:

$$\begin{cases} sen(x_1) + x_2^3 & = 0 \\ x_1 x_2 - e^{x1} & = 0 \end{cases}$$

Sumário

- Introdução
- Método de Newton
- Exercícios de aplicação

Introdução

Neste tipo de sistemas é muito difícil demonstrar a existência e unicidade de zeros de F.

Apenas em casos muito simples o sistema não linear admite uma solução analítica sendo que, os métodos para a sua resolução são de natureza iterativa devido ao carácter não linear das suas equações.

Surgem então dificuldades relacionadas com a **convergência** e sua **razão** e com a obtenção de uma **estimativa inicial** com algum significado.

O facto de se estar em $\Re^n$ também introduz dificuldades adicionais ao problema.

Existem vários métodos numéricos para a resolução de sistemas de equações não lineares, no entanto apenas será estudado o método de Newton.

Método de Newton

Algumas aulas atrás, apresentou-se o método de Newton para resolver uma equação não linear. Este consiste em aproximar a função $f(x)$ por um modelo linear local em x_k, sendo este modelo a recta tangente à função em x_k.

No caso dum sistema de n equações não lineares a ideia é a mesma - vai aproximar-se cada uma das funcões $f_i(x_{1,} \ldots, x_n)$ $i = 1, \ldots, n$, por um modelo linear em x_k

$$f_i(x) \approx f_i(x_k) + \nabla f_i(x_k)^T (x - x_k) \quad i = 1, \ldots, n$$

em que

$$\nabla f_i(x)^T = [\frac{\partial f_i(x)}{\partial x_1}, \ldots, \frac{\partial f_i(x)}{\partial x_n}]$$

é o transposto do vector gradiente de $f_i(x)$ (vector das n derivadas parciais).

Método de Newton

Considere-se a matriz do Jacobiano $J(x)$ (primeiras derivadas das funções $f_i(x)$)

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

Então o modelo linear para $F(x)$ centrado em x_k é

$$F(x) \approx L_k(x) = F(x_k) + J(x_k)(x - x_k)$$

sendo x_{k+1} o zero do modelo linear $L_k(x)$, *i.e.*,

$$L_k(x) = 0 \Leftrightarrow J(x_k)(x - x_k) = -F(x_k).$$

Método de Newton

Se denotarmos por $x - x_k$ por Δ_k, este vector é a solução do seguinte sistema linear:

$$J(x_k)\Delta_k = -\begin{bmatrix} f_1(x_k) \\ \vdots \\ f_n(x_k) \end{bmatrix} \quad \text{e} \quad x_{k+1} = x_k + \Delta_k$$

Este processo deve ser repetido, uma vez que a solução encontrada é solução de $L_k(x) = 0$, sendo uma aproximação da solução exacta de $F(x) = 0$.

Algoritmo básico

1. $k = 1$; fornecer aproximação inicial (vector x_1);
2. Resolver o sistema linear $J(x_k)\Delta_k = -f(x_k)$ para calcular o vector Δ_k (EGPP);
3. Actualizar o vector aproximação $x_{k+1} = x_k + \Delta_k$;
4. Verificar critério de paragem - se verificado $x^* \approx x_{k+1}$ [FIM] senão $k \leftarrow k+1$ e voltar para passo 2.

Método de Newton

Em termos computacionais, o método de Newton envolve em cada iteração k:

- o cálculo da matriz do Jacobiano em x_k
- a resolução dum sistema linear de dimensão n

Critério de Paragem (CP)

Critério de paragem

$$\frac{\|\Delta_k\|_2}{\|x_{k+1}\|_2} \leq \epsilon_1 \qquad \text{e} \qquad \|f_{k+1}\|_2 \leq \epsilon_2$$

A estimativa do erro relativo tem de ser suficientemente pequena e o valor da norma de f também.
(Nota: Δ, x e f são vectores em $\Re^n$).

Condições de convergência

Se

- f é um vector de funções continuamente diferenciáveis,
- x^* é tal que $f(x^*) = 0$,
- $J(x^*)$ é não singular ($\exists\, (J(x^*))^{-1}$) e $(J(x^*))^{-1}$ é limitado ($\left\|(J(x^*))^{-1}\right\| \leq \beta,\ \beta > 0$),
- $J(x)$ matriz Lipschitz contínua na vizinhança de x^*,
- $x^{(1)}$ (aproximação inicial) na vizinhança de x^* - convergência local,
- então o método iterativo de Newton converge e

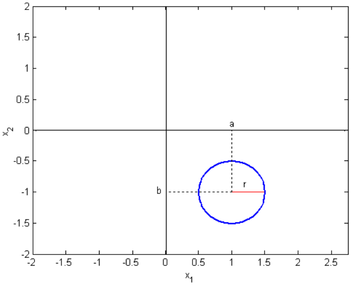
$$\frac{\left\|x^* - x^{(k+1)}\right\|}{\left\|x^* - x^{(k)}\right\|^p} \leq k \quad (k > 0,\ \ p = 2)$$

(convergência quadrática).

Exercício de aplicação 1

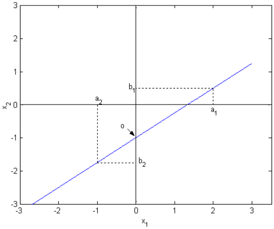
Usando o método de Newton, calcular um ponto de intersecção da circunferência com centro em $(0, 0)$ e raio $r = 3$ com a recta que passa pelos pontos $(0, 1)$ e $(1, 0)$.

Equação da circunferência:



centro: (a, b)
raio: r

Equação da recta:

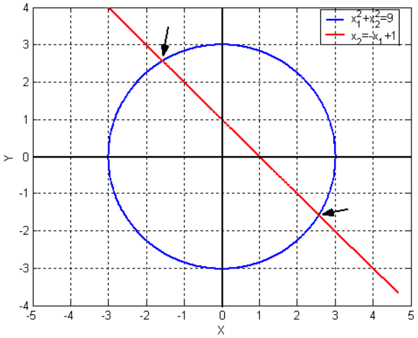


declive da recta: $m = \frac{b_1 - b_2}{a_1 - a_2}$

$$x_2 = m x_1 + o$$

Os pontos de intersecção são os pontos que verificam simultaneamente as duas equações, isto é, são os pontos que são solução do sistema:

$$\begin{cases} x_1^2 + x_2^2 = 9 \\ x_2 = -x_1 + 1 \end{cases}$$



$$\left\{ \begin{array}{rcl} x_1^2 + x_2^2 - 9 & = & 0 \\ x_2 + x_1 - 1 & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} f_1(x_1, x_2) & = & 0 \\ f_2(x_1, x_2) & = & 0 \end{array} \right.$$

Jacobiano das funções f_1 e f_2:

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 & 2x_2 \\ 1 & 1 \end{pmatrix}$$

- **1^a iteração**, k=1

$$x^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$f\begin{pmatrix} 0 \\ 0 \end{pmatrix} \equiv f(0,0) = \begin{pmatrix} -9 \\ -1 \end{pmatrix}$$

$$J\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

Matriz ampliada:

$$\left(\begin{array}{cc|c} 0 & 0 & 9 \\ 1 & 1 & 1 \end{array}\right)$$

A matriz é singular, o sistema é impossível.

Uma vez que não se pode continuar, deve escolher-se outra aproximação inicial. Seja

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{(1)} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

- **1^a iteração**, k=1

$$f\begin{pmatrix} 2 \\ 0 \end{pmatrix}^{(1)} = \begin{pmatrix} -5 \\ 1 \end{pmatrix}$$

$$J\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 1 & 1 \end{pmatrix}$$

Resolver (EGPP):

$$\begin{pmatrix} 4 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}^{(1)} = -\begin{pmatrix} -5 \\ 1 \end{pmatrix}$$

Matriz ampliada:

$$\left(\begin{array}{cc|c} 4 & 0 & 5 \\ 1 & 1 & -1 \end{array}\right)$$

Esta matriz, por acaso, é já triangular pelo que a resolução do sistema é imediata (substituição direta):

$$\Delta x_1 = \frac{5}{4} = 1.25, \quad \Delta x_2 = \frac{-1 - 1.25}{1} = -2.25$$

$$\left(\begin{array}{c} \Delta x_1 \\ \Delta x_2 \end{array}\right)^{(1)} = \left(\begin{array}{c} 1.25 \\ -2.25 \end{array}\right)$$

Cálculo da nova aproximação:

$$\left(\begin{array}{c} x_1 \\ x_2 \end{array}\right)^{(2)} = \left(\begin{array}{c} x_1 \\ x_2 \end{array}\right)^{(1)} + \left(\begin{array}{c} \Delta x_1 \\ \Delta x_2 \end{array}\right)^{(1)} =$$

$$= \left(\begin{array}{c} 2 \\ 0 \end{array}\right) + \left(\begin{array}{c} 1.25 \\ -2.25 \end{array}\right) = \left(\begin{array}{c} 3.25 \\ -2.25 \end{array}\right)$$

$$f(x^{(2)}) = \left(\begin{array}{c} 3.25^2 + (-2.25)^2 - 9 \\ 3.25 - 2.25 - 1 \end{array}\right) = \left(\begin{array}{c} 6.625 \\ 0 \end{array}\right)$$

$$\left\|\Delta x^{(1)}\right\|_2 = \left\| \begin{pmatrix} 1.25 \\ -2.25 \end{pmatrix} \right\|_2 = \sqrt{(1.25)^2 + (-2.25)^2} = 2.574$$

$$\left\|x^{(2)}\right\|_2 = \left\| \begin{pmatrix} 3.25 \\ -2.25 \end{pmatrix} \right\|_2 = \sqrt{(3.25)^2 + (-2.25)^2} = 3.953$$

$$\left\|f(x^{(2)})\right\|_2 = \left\| \begin{pmatrix} 6.625 \\ 0 \end{pmatrix} \right\|_2 = \sqrt{(6.625)^2 + (0)^2} = 6.625$$

Sejam $\varepsilon_1 = 0.1$ e $\varepsilon_2 = 0.5$

$$\frac{2.574}{3.953} = 0.651 \leq \varepsilon_1 \text{ Falso} \qquad 6.625 \leq \varepsilon_2 \text{ Falso}$$

Continuar o processo iterativo!

- **2^a iteração**, k=2

$$f(x^{(2)}) = \begin{pmatrix} 6.625 \\ 0 \end{pmatrix}$$

$$J(x^{(2)}) = J\begin{pmatrix} 3.25 \\ -2.25 \end{pmatrix} = \begin{pmatrix} 6.5 & -4.5 \\ 1 & 1 \end{pmatrix}$$

Resolver (EGPP):

$$\begin{pmatrix} 6.5 & -4.5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}^{(2)} = -\begin{pmatrix} 6.625 \\ 0 \end{pmatrix}$$

Matriz ampliada:

$$\left(\begin{array}{cc|c} 6.5 & -4.5 & -6.625 \\ 1 & 1 & 0 \end{array}\right), m_{2,1} = -0.154, \left(\begin{array}{cc|c} 6.5 & -4.5 & -6.625 \\ 0 & 1.693 & 1.020 \end{array}\right)$$

Substituição inversa:

$$\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}^{(2)} = \begin{pmatrix} -0.602 \\ 0.602 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{(3)} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{(2)} + \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}^{(2)} =$$

$$= \begin{pmatrix} 3.25 \\ -2.25 \end{pmatrix} + \begin{pmatrix} -0.602 \\ 0.602 \end{pmatrix} = \begin{pmatrix} 2.648 \\ -1.648 \end{pmatrix}$$

$$f(x^{(3)}) = \begin{pmatrix} 0.725 \\ 0 \end{pmatrix}$$

- Critério de paragem

$$\left\|f(x^{(3)})\right\|_2 = \left\|\begin{pmatrix} 0.725 \\ 0 \end{pmatrix}\right\|_2 = \sqrt{(0.725)^2 + (0)^2} = 0.725 \leq \varepsilon_2$$ Fals

Continuar o processo iterativo...

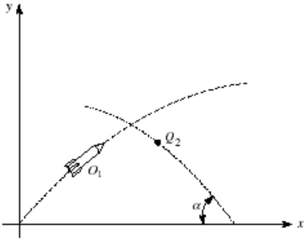
A posição de um determinado objecto O_1 no plano XY é descrita em função do tempo (t) pelas seguintes equações:

$$x_1(t) = t \qquad y_1(t) = 1 - e^{-t}$$

A posição de um segundo objecto O_2 é descrita pelas seguintes equações:

$$x_2(t) = 1 - t\cos(\alpha) \qquad y_2(t) = -0.1t^2 + tsen(\alpha)$$

em que α representa o ângulo, como mostra a figura



Exercício de aplicação 2

Determine os valores de t e α na posição em que os dois objectos colidem, *i.e.*, na posição em que se igualam as coordenadas x e y:

$$\begin{array}{rcl} t & = & 1 - t\cos(\alpha) \\ 1 - e^{-t} & = & -0.1t^2 + t sen(\alpha) \end{array}$$

Considere os valores iniciais $(t, \alpha)^{(1)} = (4.3, 2.4)$ e $\varepsilon_1 = \varepsilon_2 = 0.015$ ou no máximo duas iterações.

Nota: os cálculos devem ser feitos em radianos.

$$\left\{\begin{array}{l} t-1+tcos(\alpha)=0 \\ 1-e^{-t}+0.1t^2-tsen(\alpha)=0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} f_1(t,\alpha)=0 \\ f_2(t,\alpha)=0 \end{array}\right.$$

$$J=\begin{pmatrix} \frac{\partial f_1}{\partial t} & \frac{\partial f_1}{\partial \alpha} \\ \frac{\partial f_2}{\partial t} & \frac{\partial f_2}{\partial \alpha} \end{pmatrix} = \begin{pmatrix} 1+cos(\alpha) & -tsen(\alpha) \\ e^{-t}+0.2t-sen(\alpha) & -tcos(\alpha) \end{pmatrix}$$

Aproximação inicial:

$$\begin{pmatrix} t \\ \alpha \end{pmatrix}^{(1)} = \begin{pmatrix} 4.3 \\ 2.4 \end{pmatrix}$$

1^a iteração, $k = 1$:

$$J(4.3, 2.4) = \begin{pmatrix} 0.262606 & -2.90449 \\ 0.198105 & 3.170793 \end{pmatrix}$$

$$\begin{pmatrix} f_1(4.3, 2.4) \\ f_2(4.3, 2.4) \end{pmatrix} = \begin{pmatrix} 0.129207 \\ -0.069060 \end{pmatrix}$$

Resolvendo por EGPP

$$\begin{pmatrix} 0.262606 & -2.90449 \\ 0.198105 & 3.170793 \end{pmatrix} \Delta^{(1)} = - \begin{pmatrix} 0.129207 \\ -0.069060 \end{pmatrix}$$

$$\Leftrightarrow \left(\begin{array}{cc|c} 0.262606 & -2.90449 & -0.129207 \\ 0.198105 & 3.170793 & 0.069060 \end{array}\right)$$

$$\Delta^{(1)} = \begin{pmatrix} -0.148505 \\ 0.031058 \end{pmatrix}$$

$$\begin{pmatrix} t \\ \alpha \end{pmatrix}^{(2)} = \begin{pmatrix} t \\ \alpha \end{pmatrix}^{(1)} + \begin{pmatrix} -0.148505 \\ 0.031058 \end{pmatrix} = \begin{pmatrix} 4.1515 \\ 2.431058 \end{pmatrix}$$

Critério de paragem:

$$\frac{\left\|\Delta^{(1)}\right\|_2}{\left\|\begin{pmatrix} t \\ \alpha \end{pmatrix}^{(2)}\right\|_2} = \frac{0.15172}{4.81092} = 0.03154 \leq 0.015 \quad \text{(falso)}$$

2^a iteração, $k = 2$:

$$J(4.1515, 2.431058) = \begin{pmatrix} 0.241987 & -2.70777 \\ 0.193802 & 3.146892 \end{pmatrix}$$

$$\begin{pmatrix} f_1(4.1515, 2.431058) \\ f_2(4.1515, 2.431058) \end{pmatrix} = \begin{pmatrix} 0.004608 \\ -1.6367 \times 10^{-5} \end{pmatrix}$$

Resolver por EGPP:

$$\begin{pmatrix} 0.241987 & -2.70777 & | - 0.004608 \\ 0.193802 & 3.146892 & |1.6367 \times 10^{-5} \end{pmatrix}$$

$$\Delta^{(2)} = \begin{pmatrix} -0.01124 \\ 0.0006973 \end{pmatrix}$$

$$\begin{pmatrix} t \\ \alpha \end{pmatrix}^{(3)} = \begin{pmatrix} t \\ \alpha \end{pmatrix}^{(2)} + \begin{pmatrix} -0.01124 \\ 0.0006973 \end{pmatrix} = \begin{pmatrix} 4.14026 \\ 2.43176 \end{pmatrix}$$

Critério de paragem:

$$\frac{\|\Delta^{(2)}\|_2}{\left\| \begin{pmatrix} t \\ \alpha \end{pmatrix}^{(3)} \right\|_2} = \frac{0.01126}{4.80158} = 0.000235 \leq \varepsilon_1 \qquad \text{(verdade)}$$

$$\left\| \begin{pmatrix} f_1(4.14026, 2.43176) \\ f_2(4.14026, 2.43176) \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} 6.6 \times 10^{-6} \\ 5.0 \times 10^{-6} \end{pmatrix} \right\|_2 = 8.3 \times 10^{-6} \leq \varepsilon_2$$

(verdade)

As duas condições do critério de paragem são verificadas logo:

$$\left(\begin{array}{c} t \\ \alpha \end{array}\right)^{(*)} \approx \left(\begin{array}{c} t \\ \alpha \end{array}\right)^{(3)} = \left(\begin{array}{c} 4.14026 \\ 2.43176 \end{array}\right)$$