EQUAÇÕES NÃO LINEARES E OPTIMIZAÇÃO UNIDIMENSIONAL

Joaquim J. Judice e Luís M. Fernandes

1 Introdução

A determinação de um mínimo de uma função não linear de uma variável tem uma grande importância em Análise Matemática. Nesse problema, dado um intervalo I e uma função

$$f : I \to R$$

procura-se encontrar um $\bar{x} \in \mathbb{R}^n$ tal que $f(\bar{x}) \leq f(x)$ para todo $x \in I$. Se f é uma função continuamente diferenciável em I, então uma condição necessária para a existência de $\bar{x}$ é $f'(\bar{x}) = 0$, onde $f'(x)$ é a derivada de f em x.

Se f é uma função não linear e não quadrática, então

$$f'(x) = 0 \tag{1}$$

representa uma equação não linear de uma variável. Deste modo o problema da determinação de um mínimo para uma função de uma variável está ligado à resolução de uma equação não linear.

Neste trabalho abordaremos a resolução de equações não lineares e de problemas de optimização não linear do ponto de vista numérico. Iremos assumir que as funções f são sempre pelo menos uma vez continuamente diferenciáveis e estudaremos apenas os métodos considerados mais eficientes para a resolução desses problemas. Esses processos são para ser utilizados em computadores, sendo por isso desenvolvidos para a obtenção de uma solução aproximada. Para um problema de optimização esses algoritmos procuram obter um mínimo local, para o que tentam resolver um sistema (1), usando direcções que impedem a obtenção de máximos locais.

O facto das funções f serem continuamente diferenciáveis nos seus domínios não significa que as expressões das suas derivadas sejam sempre simples. Em muitos casos essas funções são de tal modo complexas que a determinação das expressões das suas derivadas se torna bastante difícil. É então preferível obter valores aproximados para essas derivadas. Esse assunto será tratado neste trabalho.

Os processos de resolução que iremos considerar são iterativos, isto é, geram uma sucessão de números reais $(x_n)_{n\in\mathbb{N}}$, cujo limite é uma solução do problema em causa. A velocidade (ou taxa) de convergência para a solução é determinante no tempo de execução do algoritmo. Além disso a precisão que se pretende para a solução em causa é também bastante importante no funcionamento do algoritmo. Com efeito se pretendermos uma solução com precisão pequena, então o algoritmo deve demorar menos tempo do que se pretendermos uma solução com precisão elevada. Iremos ver que a precisão da solução está dependente do critério de paragem que estabelecemos para o algoritmo.

É importante notar nesta altura que a matemática numérica é um pouco diferente da análise matemática que temos vindo a estudar. Com efeito, apesar de estudarmos algoritmos que são considerados eficientes para a resolução dos problemas em causa, nada impede que não possa surgir um problema para o qual os algoritmos sejam incapazes de o resolver. Nesse aspecto, este tipo de matemática está mais próxima das ciências experimentais, como por exemplo a medicina, onde nem sempre um remédio muito recomendado para uma doença tem o efeito curativo desejado.

2 Taxas de Convergência de Sucessões

É perfeitamente conhecida a definição de limite de uma sucessão $(x_k)_{k\in\mathbb{N}}$ de números reais. Convém recordar que se $\bar{x} = \lim_k x_k$, então a diferença $|x_k - \bar{x}|$ vai-se tornando cada vez mais próxima de zero à medida que k vai aumentando. Assim podemos escrever

$$\lim_k x_k = \bar{x} \Leftrightarrow \lim_k |x - \bar{x}| = 0$$

A taxa ou rapidez de convergência de uma sucessão diz respeito à maneira mais ou menos rápida como esta diferença $|x_k - \bar{x}|$ se vai aproximando de zero.

Definição 1 *Diz-se que uma sucessão $(x_k)_{k\in\mathbb{N}}$ tem convergência linear se existe uma constante $c \in [0, 1[$ tal que*

$$|x_{k+1} - \bar{x}| \quad \leq \quad c\,|x_k - \bar{x}| \tag{2}$$

para todo $k \geq \bar{k}$, com $\bar{k} \in \mathbb{N}$.

A constante c tem um papel muito importante na velocidade de convergência. É sabido que se $c > 1$, então

$$\lim_k \frac{1}{c^k} = 0$$

Além disso

$$\frac{1}{c^{k+1}} = \frac{1}{c}\frac{1}{c^k}$$

e portanto a sucessão $(\frac{1}{c^k})_{k\in\mathbb{N}}$ tem convergência linear com constante $\frac{1}{c}$. Facilmente se conclui que quanto maior for a constante c, menor é o valor de $\frac{1}{c}$ e mais rápida é a convergência. Assim por exemplo se $c = 2$, tem-se a seguinte sucessão

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256}, \frac{1}{512}, \frac{1}{1024}, \frac{1}{2048}, \frac{1}{4096}, \frac{1}{8192}, \frac{1}{16384}, \ldots$$

Por outro lado se $c = 10$, obtém-se a sucessão

$$\frac{1}{10}, \frac{1}{100}, \frac{1}{1000}, \frac{1}{10000}, \frac{1}{100000}, \ldots$$

Portanto o quinto termo da última sucessão já é bastante menor que o termo de ordem 14 da outra sucessão.

Este exemplo mostra que a convergência linear para um algoritmo é desejável apenas quando a constante c é muito pequena.

Definição 2 *Diz-se que uma sucessão* $(x_k)_{k\in\mathbb{N}}$ *tem convergência superlinear se existe uma sucessão* $(c_k)_{k\in\mathbb{N}}$ *de termos não negativos e convergente para zero tal que*

$$|x_{k+1} - \bar{x}| \quad \leq \quad c_k \; |x_k - \bar{x}| \tag{3}$$

para todo $k \geq \bar{k}$*, com* $\bar{k} \in \mathbb{N}$.

A convergência superlinear é assim uma convergência linear em que as constantes se vão tornando muito pequenas à medida que k vai aumentando. Assim se uma sucessão tem convergência superlinear, então converge rapidamente para o seu limite na parte final.

Assim, por exemplo, consideremos a sucessão $(\frac{1}{k!})_{k\in\mathbb{N}}$. Como

$$\lim_k \frac{1}{k!} = 0, \quad \frac{1}{(k+1)!} = \frac{1}{k+1}\frac{1}{k!} \;\text{ e }\; \lim_k \frac{1}{k+1} = 0$$

então a convergência da sucessão é superlinear. Facilmente se verifica que a sucessão converge rapidamente para zero após uns primeiros termos relativamente elevados:

$$1, \frac{1}{2}, \frac{1}{6}, \frac{1}{24}, \frac{1}{120}, \frac{1}{720}, \frac{1}{5040}, \frac{1}{40320}, \frac{1}{362880}, \ldots$$

É de notar que uma sucessão com convergência linear pode convergir mais rapidamente para o seu limite do que uma outra sucessão com convergência superlinear. No entanto, o conhecimento de que um algoritmo possui convergência superlinear é uma garantia para o utilizador desse processo, pois sabe que pelo menos o método converge rapidamente na sua parte final. A convergência linear é mais perigosa nesse sentido, pois em geral não é conhecido o valor da constante c e portanto não há qualquer garantia que o processo seja lento ou rápido.

Definição 3 *Uma sucessão* $(x_k)_{k\in\mathbb{N}}$ *tem convergência de ordem* p *se existe* $c \geq 0$ *tal que*

$$|x_{k+1} - \bar{x}| \quad \leq \quad c \; |x_k - \bar{x}|^p \tag{4}$$

para todo $k \geq \bar{k}$*, com* $\bar{k} \in \mathbb{N}$*. Se* $p = 2$ *a convergência diz-se quadrática.*

Como exemplo de ilustração de uma sucessão nessas condições, consideremos $(\frac{1}{2^{2^k}})$. Então

$$\lim_k \frac{1}{2^{2^k}} = 0, \quad \frac{1}{2^{2^{k+1}}} = \frac{1}{2^{2^k 2}} = \frac{1}{(2^{2^k})^2}$$

e a sucessão tem convergência quadrática para zero. É fácil de ver que a sucessão converge muito rapidamente para zero, pois os seus termos são

$$\frac{1}{4}, \frac{1}{16}, \frac{1}{256}, \frac{1}{65536}, \ldots$$

É de notar que se a convergência é de ordem $p \geq 2$, então

$$|x_{k+1} - \bar{x}| \quad \leq \quad c \; |x_k - \bar{x}| \; |x_k - \bar{x}|^{p-1} \tag{5}$$

e portanto a sucessão possui convergência superlinear. Assim, a convergência de ordem $p \geq 2$ é ainda mais desejável do que a superlinear. Em geral, conseguir estabelecer que um algoritmo tem convergência quadrática já satisfaz plenamente os utilizadores dos algoritmos.

Como referimos anteriormente, o ponto inicial tem normalmente influência na convergência da sucessão dos pontos gerados por um algoritmo.

Definição 4 *Diz-se que um algoritmo tem convergência local se a convergência da sucessão dos pontos gerados pelo algoritmo é apenas assegurada para certos pontos iniciais. Se a convergência não depender do ponto inicial, então diz-se global.*

É importante notar que estes conceitos de convergência local e global não têm nada a ver com a taxa de convergência. É evidente que a convergência global é desejável, pois há a garantia do processo obter sempre uma solução para o problema em questão, independentemente do ponto inicial escolhido. No entanto, o número de iterações que o algoritmo necessita para obter essa solução depende evidentemente do ponto inicial.

3 Resolução de uma Equação Não Linear

Nesta secção iremos estudar o problema da resolução de uma equação da forma

$$f(x) = 0 \tag{6}$$

com f uma função continuamente diferenciável num intervalo $D \subseteq \mathbb{R}$ que contém pelo menos uma raiz da equação. Iremos começar por apresentar o método de Newton para a resolução desse problema, para depois discutir a sua convergência e implementação.

3.1 Método de Newton

Consideremos a equação

$$x^2 - 3 = 0$$

A função

$$\begin{aligned} f \quad : \quad & \mathbb{R} \to \mathbb{R} \\ & x \to x^2 - 3 \end{aligned}$$

é continuamente diferenciável em $\mathbb{R}$ e o seu gráfico é a parábola apresentada na figura a seguir.

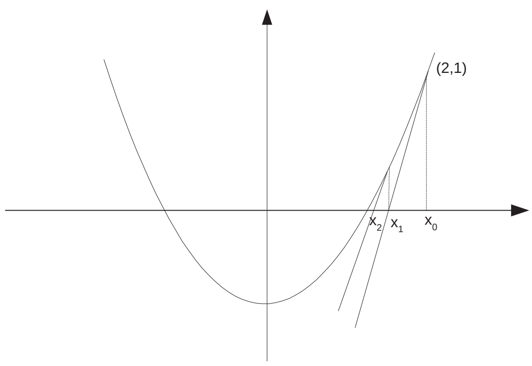

Figura 1: $y = x^2 - 3$

O método de Newton é um processo iterativo que gera uma sucessão de pontos $(x_k)_{k\in\mathbb{N}}$ tal que

$$\lim_k x_k = \bar{x}$$

é uma raiz da equação. Dado um ponto $x_k \in \mathbb{R}$ o próximo elemento x_{k+1} da sucessão é a raiz da aproximação linear de f em x_k. Assim, x_{k+1} é a abcissa do ponto de intersecção da recta tangente à curva de equação $y = f(x)$ em $P_k(x_k, f(x_k))$ com o eixo $X'X$.

Para obter a fórmula matemática que permite calcular x_{k+1}, consideremos a aproximação linear de f em x_k

$$g(x) = f(x_k) + f'(x_k)(x - x_k)$$

Como se pretende a raiz dessa função, então x_{k+1} tem de satisfazer a equação

$$-f(x_k) = f'(x_k)(x_{k+1} - x_k)$$

Assim x_{k+1} é dado de modo unívoco por

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{7}$$

desde que $f'(x_k) \neq 0$.

O método de Newton consiste em ir calculando os elementos de uma sucessão $(x_k)_{k\in\mathbb{N}}$ a partir da formula (7) até se obter um elemento x_k que satisfaz (aproximadamente) a equação $f(x) = 0$. O algoritmo necessita evidentemente de um ponto inicial x_0, que, como veremos mais adiante, tem um papel muito importante na eficiência do processo.

Consideremos novamente a equação $x^2 - 3 = 0$ e seja $x_0 = 2$ o ponto inicial (ver figura 1). Então

$$\begin{aligned} f'(x) &= 2x \Rightarrow f'(x_0) = 4 \\ f(x_0) &= 4 - 3 = 1 \end{aligned}$$

e portanto

$$x_1 = 2 - \frac{1}{4} = \frac{7}{4}$$

Para obter o segundo elemento da sucessão tem-se

$$\begin{aligned} f(x_1) &= \left(\frac{7}{4}\right)^2 - 3 = \frac{49}{16} - 3 = 0.0625 \\ f'(x_1) &= 2\,\frac{7}{4} = \frac{7}{2} = 3.5 \end{aligned}$$

e

$$x_2 = \frac{7}{4} - \frac{0.0625}{3.5} \approx 1.7321429$$

O processo poderia agora continuar com x_2. É de notar que a raiz da equação é $\sqrt{3} \approx 1.7320508$ pelo que x_2 já tem três casas decimais correctas. Aliás a figura 1 mostra exactamente que o segundo elemento da sucessão já está bastante próximo da raiz da equação.

Este exemplo mostra uma característica fundamental do método de Newton que o tem tornado muito recomendado para a solução de equações não lineares e de outros problemas: a rapidez da sua convergência desde que o ponto inicial seja bem escolhido. Seguidamente iremos debruçar-nos sobre esse assunto.

3.2 Convergência do Método de Newton

Definição 5 *Seja D um intervalo e $g\colon D \to \mathbb{R}$. Diz-se que g é contínua à Lipschitz em D com constante $\gamma > 0$ se*

$$|g(y) - g(x)| \leq \gamma |y - x| \tag{8}$$

para quaisquer $x, y \in D$.

Normalmente escrevemos

$$g \in Lip_\gamma(D)$$

para representar uma função nessas condições. Em geral não é fácil verificar se uma dada função é contínua à Lipschitz ou não. Com efeito é necessário determinar uma constante $\gamma > 0$ que verifique a desigualdade (8) para todos $x, y \in D$ e isso não é muito fácil de fazer na prática. Seguidamente apresentamos uma condição suficiente para que isso aconteça.

Teorema 1 *Se g é uma função continuamente diferenciável num intervalo $D \subseteq \mathbb{R}$ e existe $\gamma > 0$ tal que*

$$|g'(x)| \leq \gamma \qquad \text{para todo} \qquad x \in D$$

então g é contínua à Lipschitz em D com constante γ.

Demonstração: Como g é continuamente diferenciável em D, então, pelo teorema de Lagrange, para quaisquer $x, y \in D$ existe um número real c pertencente ao interior do intervalo de extremos x e y tal que

$$g(y) - g(x) = g'(c)(x - y)$$

Portanto

$$|g(y) - g(x)| = |g'(c)||x - y| \leq \gamma |x - y|$$

e isso demonstra o teorema. □

Por este teorema, toda a função continuamente diferenciável num intervalo fechado D é contínua à Lipschitz nesse conjunto. Com efeito a função derivada é contínua em D e portanto é limitada nesse intervalo.

Além disso verifica-se a seguinte propriedade.

Teorema 2 *Sejam D um intervalo e f uma função continuamente diferenciável em D. Se $f' \in Lip_\gamma(D)$, então para quaisquer $a, b \in D$*

$$|f(b) - f(a) - f'(a)(b - a)| \leq \frac{\gamma (b - a)^2}{2}$$

Demonstração: Como f é continuamente diferenciável em D, então para quaisquer $a, b \in D$ tem-se

$$f(b) - f(a) = \int_a^b f'(x)\,dx$$

Portanto

$$\begin{aligned}
|f(b)-f(a)-f'(a)(b-a)| &= \left|\int_a^b [f'(x)-f'(a)]\,dx\right| \\
&\leq \left|\int_a^b |f'(x)-f'(a)|\,dx\right| \\
&\leq \left|\int_a^b \gamma|x-a|\,dx\right| = \left|\gamma\int_a^b |x-a|\,dx\right| \\
&= \frac{\gamma(b-a)^2}{2}
\end{aligned}$$

□

Como consequência deste teorema podemos estabelecer o seguinte resultado referente à taxa de convergência do método de Newton.

Teorema 3 *Sejam D um intervalo e f uma função continuamente diferenciável em D satisfazendo as seguintes condições:*

(i) $f' \in Lip_\gamma(D)$ *para certo* $\gamma > 0$.

(ii) A equação $f(x) = 0$ *tem uma solução* $\bar{x} \in D$ *e* $\rho = |f'(\bar{x})| > 0$.

Então existe η tal que para $x_0 \in D$ satisfazendo $|\bar{x} - x_0| < \eta$, a sucessão $(x_k)_{k\in\mathbb{N}}$ definida por

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}\ , \quad k = 0, 1, \ldots$$

é convergente para $\bar{x}$. Além disso

$$|x_{k+1} - \bar{x}| \quad \leq \quad \frac{\gamma}{\rho}|x_k - \bar{x}|^2 \tag{9}$$

Demonstração: Seja $\hat{\eta}$ o comprimento do maior intervalo de centro em $\bar{x}$ contido em D e

$$\eta = \min\{\hat{\eta}, \frac{\rho}{2\gamma}\} \tag{10}$$

Se $x_k \in D$ satisfaz $|\bar{x} - x_k| < \eta$, então

$$\frac{|f'(x_k) - f'(\bar{x})|}{|f'(\bar{x})|} \leq \frac{\gamma}{\rho}|x_k - \bar{x}| \leq \frac{\gamma}{\rho}\frac{\rho}{2\gamma} = \frac{1}{2}$$

e

$$|f'(\bar{x})| - |f'(x_k)| \leq |f'(\bar{x}) - f'(x_k)| \leq \frac{1}{2}|f'(\bar{x})|$$

Donde

$$|f'(x_k)| \geq \frac{1}{2}|f'(\bar{x})| = \frac{\rho}{2}$$

Por outro lado, como $f(\bar{x}) = 0$, tem-se

$$\begin{aligned}
|x_{k+1} - \bar{x}| &= \left|x_k - \bar{x} - \frac{f(x_k) - f(\bar{x})}{f'(x_k)}\right| \\
&= \frac{1}{|f'(x_k)|}|f(\bar{x}) - f(x_k) - f'(x_k)(\bar{x} - x_k)|
\end{aligned}$$

Pelo teorema 2 tem-se

$$|x_{k+1} - \bar{x}| \quad \leq \quad \frac{2}{\rho}\frac{\gamma}{2}|x_k - \bar{x}|^2$$

o que demonstra (9). Para provar o teorema, basta demonstrar que $x_k \in]\bar{x} - \eta, \bar{x} + \eta[\subseteq D$ para qualquer $k = 1, 2, \ldots$. Como por definição de η se tem

$$|x_0 - \bar{x}| < \eta \leq \frac{\rho}{2\gamma}$$

então por (9) vem

$$\begin{aligned} |x_1 - \bar{x}| &\leq \frac{\gamma}{\rho}|x_0 - \bar{x}||x_0 - \bar{x}| \\ &\leq \frac{\rho}{2\gamma}\frac{\gamma}{\rho}|x_0 - \bar{x}| = \frac{1}{2}|x_0 - \bar{x}| < \frac{1}{2}\eta < \eta \end{aligned}$$

Portanto $x_1 \in]\bar{x} - \eta, \bar{x} + \eta[\subseteq D$ e $x_k \in]\bar{x} - \eta, \bar{x} + \eta[$ para todo k, por indução. □

Este teorema mostra que a convergência do método de Newton é local e quadrática. Com efeito o ponto inicial x_0 tem de pertencer ao intervalo

$$]\bar{x} - \eta, \bar{x} + \eta[$$

Na prática este intervalo é impossível de determinar, pois a solução $\bar{x}$ da equação não é conhecida (se o fosse não era necessário utilizar o método de Newton). Além disso o valor de η pode ser calculado por (10), mas também aí há dificuldades por as constantes γ e ρ envolvidas nessa expressão serem de difícil cálculo.

As fórmulas (9) e (10) mostram que o valor de

$$\frac{\gamma}{\rho}$$

tem importância não só no comprimento do intervalo a que x_0 pode pertencer como na taxa de convergência do algoritmo. Com efeito quanto menor for essa quantidade, mais rápida será a convergência e em princípio maior será o comprimento do intervalo $]\bar{x} - \eta, \bar{x} + \eta[$ a que x_0 tem de pertencer. Isso significa que funções f para as quais essa quantidade seja muito pequena são as mais apropriadas para a resolução de equações $f(x) = 0$ com o método de Newton.

Exemplo: Para verificar a influência das constantes γ, ρ na taxa de convergência do método de Newton, consideremos as duas equações não lineares

$$f_1(x) = x^2 - 1 = 0$$
$$f_2(x) = x^4 - 1 = 0$$

cujos gráficos são apresentados na figura 2.

Então $\bar{x} = 1$ é raiz simples de ambas as equações, pois

$$f_1(\bar{x}) = f_2(\bar{x}) = 0$$

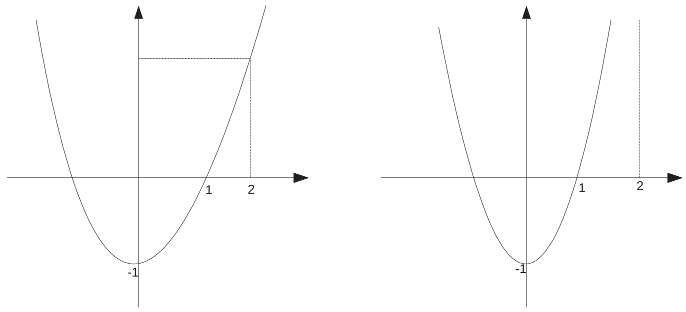

Figura 2: $f_1(x) = x^2 - 1$ $\qquad$ $f_2(x) = x^4 - 1$

e

$$f_1'(x) = 2x \;\Rightarrow\; f_1'(\bar{x}) = 2$$
$$f_2'(x) = 4x^3 \;\Rightarrow\; f_2'(\bar{x}) = 4$$

Seja $D = [0, 2]$. Como

$$f_1''(x) = 2$$
$$f_2''(x) = 12x^2$$

então as constantes γ_1 e γ_2 de Lipschitz para as funções f_1' e f_2' são respectivamente

$$\gamma_1 = 2, \quad \gamma_2 = \max_{x \in [0,2]} |f''(x)| = 48$$

Por outro lado podemos escrever $\rho_1 = |f_1'(\bar{x})| = 2$ e $\rho_2 = |f_2'(\bar{x})| = 4$ e portanto tem-se

$$\frac{\gamma_1}{\rho_1} = 1, \quad \frac{\gamma_2}{\rho_2} = 12$$

De acordo com o teorema 3, o método de Newton deve obter mais rapidamente a raiz da primeira equação. Efectuando as duas primeiras iterações do método de Newton em relação à primeira equação com $x_0 = 2$, tem-se

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 2 - \frac{3}{4} = \frac{5}{4} = 1.25$$
$$x_2 = \frac{5}{4} - \frac{\frac{9}{16}}{\frac{10}{4}} = \frac{5}{4} - \frac{\frac{9}{10}}{4} = 1 + \frac{1}{40} = 1.025$$

o que indica já uma grande proximidade em relação à raiz da equação na segunda iteração do processo. Em relação à equação $f_2(x) = 0$ e também para $x_0 = 1$, vem

$$x_1 = 2 - \frac{15}{32} = \frac{49}{32} = 1.53125$$
$$x_2 = \frac{64}{15} - \frac{(\frac{49}{32})^4 - 1}{4(\frac{49}{32})^3} = 1.218$$

e portanto a convergência é mais lenta neste caso. $\square$

É importante acrescentar que na prática é difícil de calcular os valores de γ e ρ e assim determinar o intervalo onde o método tem convergência garantida. Notar que o teorema anterior é apenas uma condição suficiente para que a convergência do método seja assegurada. Por exemplo no caso das equações $f_i(x) = 0, i = 1, 2$ do exemplo anterior $\hat{\eta} = 1$ e portanto

$$\eta_1 = \min\{1, \frac{1}{2}\} = \frac{1}{2}$$
$$\eta_2 = \min\{1, \frac{1}{24}\} = \frac{1}{24}$$

No entanto o método converge em ambos os casos para a raiz $\bar{x} = 1$ quando o ponto inicial $x_0 = 2$ não pertence aos intervalos $]\bar{x} - \eta_1, \bar{x} + \eta_1[$ e $]\bar{x} - \eta_2, \bar{x} + \eta_2[$.

Da discussão anterior chega-se à conclusão que a escolha do ponto inicial para o método de Newton é o grande problema desse processo. Para atestar das dificuldades que uma má escolha pode acarretar, consideremos a equação

$$\operatorname{arctg} x = 0$$

É possível provar que existe um valor α pertencente ao intervalo $]1.39, 1.40[$ tal que se $x_0 = \alpha$ então $x_1 = -\alpha$, $x_2 = \alpha$, $x_3 = -\alpha$ e assim sucessivamente, ou seja, o algoritmo é incapaz de obter a raiz $x = 0$ da equação (ver figura 3).

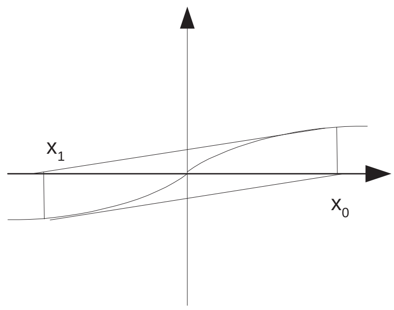

Figura 3: $y = \operatorname{arctg} x \qquad x_1 = -\alpha,\ x_0 = \alpha, \qquad \alpha \in]1.39, 1.40[$

É portanto necessário modificar o método de Newton de modo a torná-lo global, isto é, capaz de obter uma raiz da equação independentemente do ponto inicial escolhido. Numa próxima secção iremos estudar uma forma de resolver esse problema.

A hipótese (ii) do teorema anterior assume que a derivada da função f não se anula na raiz $\bar{x}$ da equação para a qual o método converge. Assim $f'(\bar{x}) \neq 0$ e a raiz $\bar{x}$ é simples. Suponhamos agora que $f'(\bar{x}) = 0$. Consideremos o caso mais simples em que os elementos da sucessão $(x_k)_{k \in \mathbb{N}}$ usados pelo algoritmo satisfazem $|f'(x_k)| \geq \rho/2$. Então a demonstração da desigualdade (9) mantém-se válida. Contudo como f é continuamente diferenciável e $f'(\bar{x}) = 0$, é natural que ρ tenha de ser uma quantidade muito pequena, o que implica uma redução de $|x_k - \bar{x}|$ muito mais lenta. Essa discussão permite entender que o método de

Newton tem uma rapidez muito menor quando a raiz $\bar{x}$ da equação $f(x) = 0$ não é simples. Seguidamente apresentamos a aplicação do método de Newton à resolução de duas equações $x^2 - 1 = 0$ e $x^2 - 2x + 1 = 0$.

$x^2 - 1 = 0$	$x^2 - 2x + 1 = 0$
2	2
1.25	1.5
1.025	1.25
1.00030487804	1.125
1.0000000464	1.0625
1.	1.03125

Notar que 1 é raiz simples de $x^2 - 1 = 0$, isto é, $f'(1) \neq 0$ e que se verifica uma taxa de convergência quadrática. Na segunda equação $x = 1$ é raiz dupla, pois satisfaz $f'(1) = 0$. É visível uma taxa de convergência muito mais lenta. Com efeito começando com o mesmo ponto inicial, o algoritmo demorou 6 iterações para obter a raiz $x = 1$ da equação $x^2 - 1 = 0$, enquanto que no caso da segunda equação apenas as décimas estão correctas na sexta iteração do algoritmo.

3.3 Critério de Paragem

Na teoria os algoritmos deveriam terminar quando se encontrasse um ponto $\bar{x}$ tal que $f(\bar{x}) = 0$. O facto da precisão dos computadores ser finita torna impossível a determinação de um tal ponto. Em vez disso, os algoritmos terminam quando se obtém um elemento x_k tal que

$$|f(x_k)| < TOL \tag{11}$$

com TOL uma tolerância para zero. O valor dessa tolerância depende da precisão da máquina ϵ_M do computador e da precisão que é pretendida para a solução da equação $f(x) = 0$. Na prática TOL é normalmente igual a $\sqrt{\epsilon_M}$.

Os algoritmos discutidos nesta secção são iterativos, isto é, geram uma sucessão de pontos $(x_k)_{k\in\mathbb{N}}$ cujo limite é uma raiz da equação $f(x) = 0$. É evidente que à medida que nos aproximamos do limite da sucessão, os termos da sucessão tendem a ficar muito próximos uns dos outros, isto é, o erro relativo

$$\frac{|x_k - x_{k-1}|}{|x_k|}$$

tende a tornar-se muito pequeno. Por isso é também boa política incorporar um critério do tipo

$$\frac{|x_k - x_{k-1}|}{max\{|x_k|, 1\}} < TOL \tag{12}$$

com TOL uma tolerância para zero (normalmente $TOL = \sqrt{\epsilon_M}$). O denominador da expressão (12) tem como objectivo a ultrapassagem de problemas que podem ocorrer quando $|x_k|$ for muito pequeno. É frequente o critério (12) verificar-se com $|f(x)|$ pequeno mas sem satisfazer o critério (11). Nessas condições o algoritmo deve terminar, pois deve ser difícil conseguir melhorar a precisão da raiz da equação. Por outro lado se o critério (11) é satisfeito sem a verificação da condição (11), é boa política baixar o valor da tolerância TOL e deixar o algoritmo continuar por mais umas iterações, pois é muito possível que se melhore a precisão da raiz da equação que o algoritmo obtém.

3.4 Aproximação das Derivadas

Da descrição do método de Newton, conclui-se que em cada iteração é necessário calcular a derivada de f no ponto x_k correspondente. Se $f(x)$ tem uma expressão muito complexa, é difícil determinar a expressão da derivada $f'(x)$. Nesse caso é preferível aproximar $f'(x_k)$ por um valor suficientemente próximo. Da definição de derivada de f em x_k tem-se

$$f'(x_k) = \lim_{h_k \to 0} \frac{f(x_k + h_k) - f(x_k)}{h_k}$$

e portanto

$$f'(x_k) \approx \frac{f(x_k + h_k) - f(x_k)}{h_k} = a_k \tag{13}$$

quando h_k é muito pequeno. A determinação do valor de h_k é bastante importante, pois tem de se ter alguma garantia que a_k seja um valor bastante próximo de $f'(x_k)$. O próximo teorema fornece alguma ajuda para a escolha desse valor de h_k.

Teorema 4 *Seja D um intervalo e f uma função continuamente diferenciável em D. Se $f' \in Lip_\gamma(D)$, então*

$$|a_k - f'(x_k)| \leq \frac{\gamma |h_k|}{2}$$

Demonstração: Do teorema 2, tem-se

$$|f(x_k + h_k) - f(x_k) - h_k f'(x_k)| \leq \frac{\gamma |h_k|^2}{2}$$

Portanto

$$\left| \frac{f(x_k + h_k) - f(x_k)}{h_k} - f'(x_k) \right| \leq \frac{\gamma |h_k|}{2}$$

e isso demonstra o resultado. □

Este teorema mostra que se se pretender que a_k tenha t casas decimais correctas em relação à derivada $f'(x_k)$, então é conveniente que h_k seja da ordem 10^{-t} ou inferior. Por outro lado se h_k é um valor demasiado pequeno, então

$$x_k + h_k \approx x_k$$

e a_k vem igual a zero, com todos os problemas daí inerentes. Uma boa escolha para h_k que tem sido muito recomendada na prática é dada por

$$h_k = \sqrt{\epsilon_M} \max\{1, |x_k|\}. \tag{14}$$

Este processo de calcular numericamente a derivada de uma função é conhecido em Análise Numérica por diferenças finitas progressivas. Existem outras regras semelhantes, tais como as diferenças finitas regressivas e as diferenças centrais, que não serão discutidas neste curso.

3.5 Método de Newton com Derivadas Aproximadas e Método da Secante

O Método de Newton com Derivadas Aproximadas tem a forma

$$x_{k+1} = x_k - \frac{f(x_k)}{a_k} \tag{15}$$

onde a_k é dado por (13).

Para estudar a taxa da convergência da sucessão dos pontos gerados pelo método de Newton, seja $\bar{x}$ a solução da equação $f(x) = 0$ tal que $|f'(\bar{x})| = \rho > 0$. Então tem-se

$$\begin{aligned} x_{k+1} - \bar{x} &= x_k - \bar{x} - \frac{f(x_k)}{a_k} \\ &= \frac{1}{a_k}[f(\bar{x}) - f(x_k) - a_k(\bar{x} - x_k)] \end{aligned}$$

pois $f(\bar{x}) = 0$. Portanto

$$|x_{k+1} - \bar{x}| = \frac{1}{|a_k|}|f(\bar{x}) - f(x_k) - f'(x_k)(\bar{x} - x_k) + [f'(x_k) - a_k](\bar{x} - x_k)|$$

e pelo teorema 2 vem

$$|x_{k+1} - \bar{x}| \leq \frac{1}{|a_k|}(\frac{\gamma}{2}|\bar{x} - x_k|^2 + |f'(x_k) - a_k||\bar{x} - x_k|)$$

Como $|f'(\bar{x})| = \rho > 0$ então, usando o teorema 4, tem-se

$$|x_{k+1} - \bar{x}| \leq \frac{\gamma}{2|a_k|}(|\bar{x} - x_k| + |h_k|)|\bar{x} - x_k|$$

Mas

$$|f'(x_k)| - |a_k| \leq |f'(x_k) - a_k|$$

Portanto por (10) e pelo teorema 3, vem

$$|a_k| \geq |f'(x_k)| - |f'(x_k) - a_k| \geq \frac{\rho}{2} - \frac{\gamma|h_k|}{2} = \frac{1}{2}(\rho - \gamma|h_k|) > 0$$

para h_k muito pequeno. Então

$$|x_{k+1} - \bar{x}| \leq \frac{\gamma}{\rho - \gamma|h_k|}(|\bar{x} - x_k| + |h_k|)|\bar{x} - x_k| \tag{16}$$

Assim, se as condições do teorema 3 são verificadas, o método de Newton com derivadas aproximadas tem convergência superlinear se a sucessão $(h_k)_{k\in\mathbb{N}}$ tender para zero. Com efeito

$$\lim_k |x_k - \bar{x}| + |h_k| = 0$$

Podemos assim concluir que o método de Newton tem uma convergência um pouco mais lenta se as derivadas são aproximadas. Além disso o uso das derivadas aproximadas a_k obriga a um cálculo adicional do valor da função f no ponto $x_k + h_k$. Isso pode ser bastante custoso se a expressão da função $f(x)$ é muito complexa. Nesses casos o método da secante pode ter alguma utilidade.

O Método da Secante consiste em fazer em cada iteração k,

$$x_{k+1} = x_k - \frac{f(x_k)}{a_k}$$

com

$$a_k = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

onde x_0 e x_{-1} são dois pontos dados. Assim a_k é o declive da recta secante ao gráfico da função f nos pontos de abcissas x_{k-1} e x_k. Essa propriedade dá o nome ao método.

Este processo pode ser visto como um método de Newton com derivadas aproximadas em que $|h_k| = |x_k - x_{k-1}|$. Mas $(h_k)_{k\in\mathbb{N}}$ tende para zero, e por isso o algoritmo tem convergência superlinear quando converge para uma raiz simples da equação. Como $f(x_{k-1})$ tinha sido calculado na iteração $(k-1)$, então há apenas a necessidade de calcular um valor da função em cada iteração. Esta vantagem do método da secante em relação ao método de Newton tem a contrapartida de uma mais lenta velocidade de convergência. A seguir apresentamos os resultados da aplicação do método de Newton, método de Newton com derivadas aproximadas e método da secante para a determinação de uma raiz da equação $x^2 - 1 = 0$. Em todos os processos $x_0 = 2$ é o ponto inicial.

Método de Newton	Método de Newton com derivadas aproximadas	Método da Secante
2.	2.	2.
1.25	1.25000002664	1.25000002645
1.025	1.025000017905	1.076923084491
1.000304878	1.0003048001120	1.0082644643823
1.00000004646	1.00000004647	1.0003048781354
1.0		1.0000012544523
		1.0000000001912
		1.0

É de notar que no método de Newton com derivadas aproximadas o valor de h_k é dado por (14). Como $\epsilon_M = 10^{-16}$ no computador em causa, então

$$h_k = 10^{-8}|x_k|\,,\ k = 0, 1, \ldots$$

É interessante notar que em cada ponto gerado pelos métodos de Newton com ou sem derivadas aproximadas as 7 primeiras casas decimais dos pontos da mesma iteração coincidem. Isso é consequência do valor de h_k e também da simplicidade da expressão de $f(x) = x^2 - 1$. Além disso o método de Newton com derivadas aproximadas não consegue obter a raiz 1.0, devido ao facto de $(h_k)_{k\in\mathbb{N}}$ não tender para zero.

Como conclusão desta secção podemos afirmar que o método de Newton é um processo eficiente para a resolução de equações não lineares, mas pode não convergir para uma raiz da equação ou mesmo divergir na prática. Se a expressão $f(x)$ é simples, então a forma pura (7) deve ser usada, isto é, a derivada $f'(x_k)$ deve ser calculada em cada iteração k. Se a expressão $f(x)$ é bastante complexa, então derivadas aproximadas ou o método da secante têm utilidade.

4 Optimização Unidimensional

4.1 Método de Newton Global

Considere o seguinte problema

$$\begin{array}{rl} \text{Minimizar} & f(x) \\ \text{sujeito a} & x \in D \end{array}$$

com f uma função duas vezes continuamente diferenciável num intervalo D. Como a determinação do mínimo global da função f é em geral muito difícil, limitar-nos-emos a estudar o problema da determinação de um mínimo local de f. Notar que o problema

$$\begin{array}{rl} \text{Maximizar} & f(x) \\ \text{sujeito a} & x \in D \end{array}$$

se reduz a minimizar $g(x) = -f(x)$ em D.

Se $\bar{x} \in int(D)$ é mínimo local de f, então $\bar{x}$ satisfaz a equação

$$f'(x) = 0 \tag{17}$$

Contudo a verificação de tal equação não garante um mínimo local, podendo a raiz dessa equação ser um máximo local ou um ponto de sela. Por isso os processos para a determinação de um mínimo local de uma função procuram obter uma raiz da equação (17) de modo a que o valor da função diminua em cada iteração.

Para a determinação de uma raiz da equação (17) é conveniente usar o método de Newton, que neste caso terá a forma

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}, \quad k = 0, 1, \dots \tag{18}$$

desde que $f''(x_k) \neq 0$. É fácil de ver que nessas condições x_{k+1} é o ponto que satisfaz $m'(x_k) = 0$, onde $m(x_k)$ é a aproximação quadrática de f em x_k, ou seja,

$$m(x_k) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2$$

cujo gráfico é uma parábola. As figuras a seguir ilustram a determinação do ponto x_k nos casos em que x_k pertence a uma parte convexa ou côncava.

Tal como no caso das equações não lineares, o método de Newton tem convergência local e quadrática desde que convirja para um ponto $\bar{x}$ satisfazendo $f''(\bar{x}) \neq 0$. Dada a natureza local da sua convergência, é importante desenvolver uma técnica com convergência global que incorpore o método de Newton. Para descrever uma iteração k desse método de Newton global, seja x_k tal que $f''(x_k) \neq 0$ o ponto corrente e calcule-se x_{k+1} a partir da fórmula (18). Dois casos podem acontecer e são discutidos a seguir.

(i) Se $f(x_{k+1}) < f(x_k)$, então o ponto x_{k+1} deve ser aceite, pois houve um decréscimo no valor da função que se pretende minimizar.

(ii) Se $f(x_{k+1}) \geq f(x_k)$, então é necessário encontrar um novo ponto z_{k+1} tal que $f(z_{k+1}) < f(x_k)$. Esse ponto pode estar no interior do intervalo I_k que une os pontos x_k e x_{k+1} ou no intervalo J_k que tem como ponto extremo x_k e satisfaz $J_k \bigcap I_k = \{x_k\}$.

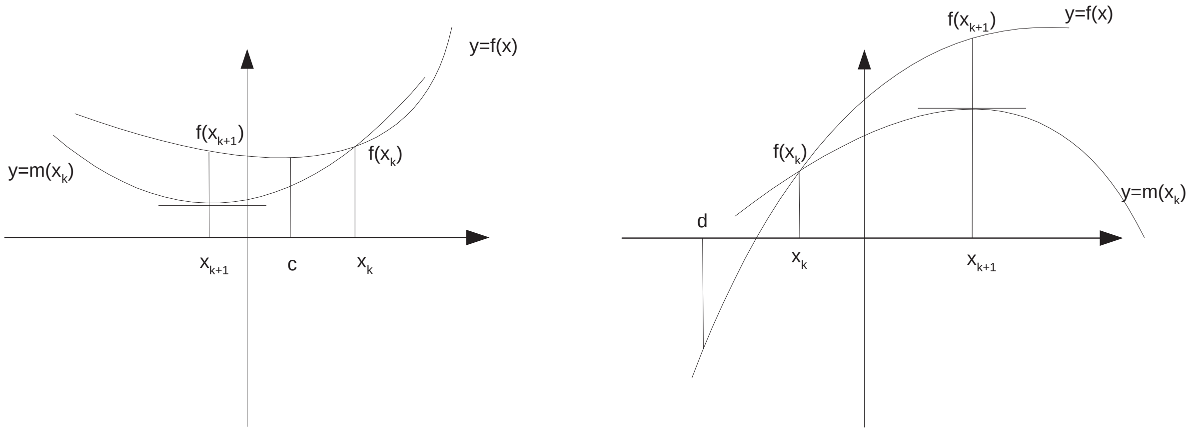

Figura 4: $c = x_k + \frac{1}{2}(x_{k+1} - x_k)$ $\qquad$ $d = x_k - \frac{1}{2}(x_{k+1} - x_k)$

Em qualquer dos casos o algoritmo determina o novo ponto x_{k+1} ou z_{k+1} com menor valor da função f e uma nova iteração é repetida. Seguidamente iremos mostrar que um tal ponto z_{k+1} referido em (ii) existe sempre. Para isso é necessário estabelecer o seguinte resultado.

Teorema 5 *Se f é uma função continuamente diferenciável num intervalo D, $\bar{x}, \bar{y} \in D$, I é o intervalo de extremos $\bar{x}$ e $\bar{y}$ e J é o intervalo em que $\bar{x}$ é um dos extremos e satisfaz $I \bigcap J = \{\bar{x}\}$, então*

(i) Se $f'(\bar{x})(\bar{y} - \bar{x}) < 0$, existe $\bar{z} \in I$ tal que $f(\bar{z}) < f(\bar{x})$.

(ii) Se $f'(\bar{x})(\bar{y} - \bar{x}) > 0$, existe $\bar{z} \in J$ tal que $f(\bar{z}) < f(\bar{x})$.

Demonstração: (i) Se $\bar{x} < \bar{y}$, então $f'(\bar{x}) < 0$. Como f é continuamente diferenciável em D, existe $\bar{z} \in]\bar{x}, \bar{y}[$ tal que $f'(x) < 0$ para todo $x \in [\bar{x}, \bar{z}]$. Pelo teorema de Lagrange tem-se

$$f(\bar{z}) - f(\bar{x}) = f'(c)(\bar{z} - \bar{x})$$

com $c \in]\bar{x}, \bar{z}[$. Mas $f'(c) < 0$ e portanto,

$$f(\bar{z}) < f(\bar{x}).$$

Se $\bar{x} > \bar{y}$, então $f'(\bar{x}) > 0$ e existe $\bar{z} \in]\bar{y}, \bar{x}[$ tal que $f'(x) > 0$ para todo $x \in [\bar{z}, \bar{x}]$. Mas, pelo teorema de Lagrange, existe $c \in]\bar{z}, \bar{x}[$ tal que

$$f(\bar{z}) - f(\bar{x}) = f'(c)(\bar{z} - \bar{x}) < 0$$

e o resultado está demonstrado também neste caso.

(ii) A demonstração é semelhante à anterior. $\square$

Este resultado permite mostrar a existência na iteração k do algoritmo de Newton global de um ponto z_{k+1} tal que $f(z_{k+1}) < f(x_k)$. Basta para isso considerar $x_k = \bar{x}$, $x_{k+1} = \bar{y}$ e $z_{k+1} = \bar{z}$ nesse teorema. Além disso tem-se

(i) $f'(x_k)(x_{k+1} - x_k) < 0 \Rightarrow z_{k+1} \in int(I_k)$

(ii) $f'(x_k)(x_{k+1} - x_k) > 0 \Rightarrow z_{k+1} \in int(J_k)$

onde I_k e J_k são os intervalos definidos anteriormente no ponto (ii) da descrição do algoritmo.

O método de Newton Global determina esse ponto z_{k+1} por tentativas da forma

$$z = x_k \pm \frac{1}{2^p}(x_{k+1} - x_k), \quad p = 0, 1, \ldots$$

até a obtenção de um inteiro não negativo $\bar{p}$ tal que

$$z = x_k \pm \frac{1}{2^{\bar{p}}}(x_{k+1} - x_k)$$

satisfaz $f(z) < f(x_k)$. É de notar que essa pesquisa é iniciada com $p = 0$, o que significa que x_{k+1} é o primeiro ponto a ser investigado. Os passos do algoritmo de Newton Global são apresentados a seguir.

Algoritmo de Newton Global

Seja dado $x^0 \in D$.

Para k = 0,1,2,...

Se $f''(x_k) \neq 0$, calcule

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

Se $f'(x_k) \approx 0$ termine com x_k (em geral x_k é um mínimo local de f em D).

Se $f(x_{k+1}) \geq f(x_k)$ faça

Para p = 1,2,...

Calcule

$$z = \begin{cases} x_k + \frac{1}{2^p}(x_{k+1} - x_k), & \text{se } f'(x_k)(x_{k+1} - x_k) < 0 \\ x_k - \frac{1}{2^p}(x_{k+1} - x_k), & \text{se } f'(x_k)(x_{k+1} - x_k) > 0 \end{cases}$$

Se $f(z) < f(x_k)$ faça $x_{k+1} = z$.

$\square$

A próxima figura ilustra o primeiro caso do cálculo de z com $p = 1$.

A convergência do método de Newton Global para um ponto $\bar{x}$ satisfazendo $f'(\bar{x})$ está assegurada, pois em cada iteração o valor da função f decresce estritamente. Além disso o método não pode convergir para um máximo local e em geral obtém um mínimo local $\bar{x}$. Em relação à taxa de convergência, suponhamos que o ponto $\bar{x}$ obtido pelo método satisfaz a

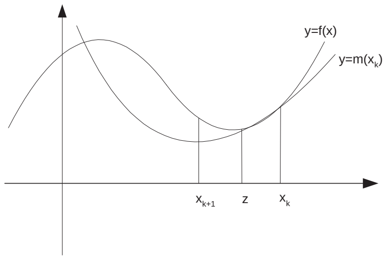

Figura 5: $z = x_k + \frac{1}{2}(x_{k+1} - x_k)$

condição suficiente de optimalidade de 2ª ordem, isto é, $f''(\bar{x}) > 0$. Então, de acordo com o teorema 3, existe um intervalo de centro em $\bar{x}$ tal que o método de Newton simples é usado até ao fim, isto é, $f(x_{k+1}) < f(x_k)$ em todas as iterações até final. Portanto o método de Newton global mantém a taxa de convergência quadrática quando converge para um ponto $\bar{x}$ satisfazendo a condição suficiente de optimalidade de 2ª ordem.

Apesar da convergência do método de Newton estar assegurada (desde que $f''(x_k) \neq 0$ em cada iteração k) para qualquer ponto inicial x_0, a escolha desse ponto tem grande importância na determinação do mínimo local. Com efeito é de esperar que diferentes escolhas de pontos iniciais conduzam a mínimos locais distintos. Poderá ainda acontecer que para uma determinada escolha do ponto inicial o método de Newton vá gerando sucessivos pontos com valores da função estritamente decrescentes sem verificação da condição de optimalidade. Nesse caso, o algoritmo deve parar com indicação de que a função é ilimitada inferiormente e por isso não é possível obter um mínimo local. O próximo exemplo ilustra estes dois casos.

Exemplo: Considere-se o problema da minimização da função

$$f(x) = x^3 - 3x$$

em $\mathbb{R}$. A representação gráfica dessa função é apresentada na figura a seguir.

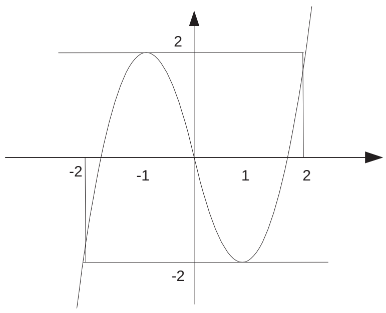

Figura 6: $y = x^3 - 3x$

É fácil de ver que f só tem um mínimo local em $x = 1$ e que não tem mínimo global. Se considerarmos o ponto inicial $x_0 = 2$ então tem-se

$$f(x_0) = 2^3 - 3.2 = 2$$
$$f'(x) = 3(x^2 - 1) \Rightarrow f'(x_0) = 9$$
$$f''(x) = 6x \Rightarrow f''(x_0) = 12$$

e

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} = 2 - \frac{9}{12} = \frac{5}{4}$$

Como

$$f(x_1) = f(\frac{5}{4}) = \left(\frac{5}{4}\right)^3 - 3\frac{5}{4} = \frac{5}{4}\left(\frac{25}{16} - 3\right) < 0$$

então $f(x_1) < f(x_0)$ e o ponto x_1 é aceite. Na segunda iteração vem

$$\begin{aligned} x_2 &= x_1 - \frac{f'(x_1)}{f''(x_1)} = \frac{5}{4} - \frac{3[(\frac{5}{4})^2 - 1]}{6\frac{5}{4}} = \frac{5}{4} - \frac{3\frac{9}{16}}{\frac{3.5}{2}} \\ &= \frac{5}{4} - \frac{9}{40} = \frac{41}{40} = 1.025 \end{aligned}$$

É fácil de ver que $f(x_2) < f(x_1)$, pelo que o ponto x_2 é aceite. É de notar que x_2 já tem uma casa decimal correcta em relação ao mínimo local ($\bar{x} = 1$) para o qual o método converge neste caso. Como a convergência é quadrática, então o algoritmo obteria o mínimo local rapidamente.

Consideremos agora o ponto inicial $x_0 = -2$. Então

$$f(x_0) = (-2)^3 - 3(-2) = -2$$
$$f'(x_0) = 9, \quad f''(x_0) = -12$$

e

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} = -2 - \frac{9}{-12} = -\frac{5}{4}$$

Como

$$f(x_1) = \left(-\frac{5}{4}\right)^3 - 3(-\frac{5}{4}) = -\frac{5}{4}\left(\frac{25}{16} - 3\right) > 0$$

então $f(x_1) > f(x_0)$ e o ponto x_1 não pode ser aceite. Além disso

$$f''(x_0) = -12$$
$$f'(x_0)(x_1 - x_0) = 9(-\frac{5}{4} - (-2)) > 0$$

Portanto considera-se o ponto

$$x_p = x_0 - \frac{1}{2}(x_1 - x_0) = -2 - \frac{1}{2}(-\frac{5}{4} + 2) = -\frac{19}{8}$$

e facilmente se conclui da figura 6 que

$$f(-\frac{19}{8}) < f(-2).$$

Donde $x_1 = -\frac{19}{8}$ é o novo ponto gerado pelo algoritmo, que prosseguiria até dar uma indicação que a função é ilimitada inferiormente, não sendo possível obter um mínimo local.

Para terminar esta secção iremos discutir o Critério de Paragem do processo e a aproximação das derivadas.

4.2 Critério de Paragem

Como se pretende resolver a equação $f'(x) = 0$, então o primeiro Critério de Paragem tem a forma

$$|f'(x_k)| < TOL$$

com TOL uma tolerância para zero. Por outro lado, como o método gera uma sucessão de pontos, então também deve ser usado o critério

$$\frac{|x_k - x_{k-1}|}{\max\{|x_k|, 1\}} < TOL$$

A tolerância TOL tem normalmente um valor igual a $\sqrt{\epsilon_M}$.

4.3 Método de Newton com Derivadas Aproximadas e Método da Secante

Se assumirmos que a derivada $f'(x)$ se pode calcular explicitamente, então três casos podem acontecer e são discutidos a seguir.

(i) Se a derivada de $f'(x)$ é relativamente simples, então $f''(x)$ deve ser calculada explicitamente.

(ii) Se a expressão da derivada $f''(x)$ é difícil de obter, mas $f'(x)$ não é muito complexa, então derivadas aproximadas devem ser usadas, isto é, faz-se

$$x_{k+1} = x_k - \frac{f'(x_k)}{a_k} \tag{19}$$

com

$$a_k = \frac{f'(x_k + h) - f'(x_k)}{h_k}$$

e h_k dado por (14).

(iii) Se a expressão de $f'(x)$ é muito complexa, então o método da secante pode ser usado com vantagem. Nesse processo usa-se a fórmula (19) com

$$a_k = \frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}$$

onde x_0 e x_{-1} são dados inicialmente.

Tal como no caso das equações não lineares, o método da secante tem convergência superlinear desde que a função satisfaça a condição de optimalidade de 2ª ordem $f''(\bar{x}) > 0$ no limite $\bar{x}$ da sucessão. O método da secante é mais lento do que o método de Newton com derivadas aproximadas, que por sua vez é relativamente mais lento e menos preciso que o método de Newton com derivadas exactas. O caso de $f'(x)$ não ser conhecida explicitamente não será discutido neste trabalho.

5 Método de Newton Global para Equações Não Lineares

Consideremos novamente uma equação não linear

$$f(x) = 0$$

e a função de mérito

$$g(x) = \frac{1}{2}[f(x)]^2$$

Então verificam-se as seguintes propriedades

Teorema 6 *(i) $\bar{x}$ é raiz da equação $f(x) = 0$ em $D \subseteq \mathbb{R}$ se e só se $\bar{x}$ é mínimo global da função de mérito g em D e $g(\bar{x}) = 0$.*

(ii) Se f é continuamente diferenciável num intervalo $D \subseteq \mathbb{R}$, então g é continuamente diferenciável em D e

$$g'(x) = f'(x)f(x)$$

para todo $x \in D$.

O método de Newton global discutido na secção anterior permite determinar um mínimo local (ou ponto de sela) $\bar{x}$ da função de mérito g em D. Como $\bar{x}$ satisfaz $g'(\bar{x}) = 0$ então, pela propriedade anterior

$$f(\bar{x}) = 0 \quad \text{ou} \quad f'(\bar{x}) = 0$$

Assim o método de Newton Global permite determinar uma solução da equação não linear $\bar{x}$, desde que essa raiz seja simples. O processo pode ainda ser modificado de forma a utilizar a Fórmula de Newton (7) para equações não lineares em vez da fórmula de Newton (18) para a minimização da função de mérito. Introduzindo essa modificação no algoritmo, podemos escrever os passos do método de Newton global na seguinte forma:

Método de Newton Global

Seja dado $x^0 \in \mathbb{R}$.

Para k = 0,1,2,...

(i) Se $f(x_k) \approx 0$, termine: x_k é raiz da equação $f(x) = 0$.

(ii) Se $f'(x_k) \approx 0$, termine: x_k é mínimo local (ou ponto de sela) da função g.

(iii) Calcule

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Se $g(x_{k+1}) \geq g(x_k)$ faça

Para p = 1,2, ...

Calcule

$$z = \begin{cases} x_k + \frac{1}{2^p}(x_{k+1} - x_k), & \text{se } g'(x_k)(x_{k+1} - x_k) < 0 \\ x_k - \frac{1}{2^p}(x_{k+1} - x_k), & \text{se } g'(x_k)(x_{k+1} - x_k) > 0 \end{cases}$$

Se $g(z) < g(x_k)$ faça $x_{k+1} = z$. □

É de notar que o método pode ser usado mesmo no caso em que o ponto x_{k+1} definido pela fórmula de Newton (7) não pertence ao domínio da função, bastando considerar $g(x_{k+1}) = \infty$ para o efeito.

Exemplo: Considere-se a equação

$$\log x = 0$$

que como se sabe tem uma única raiz $\bar{x} = 1$.

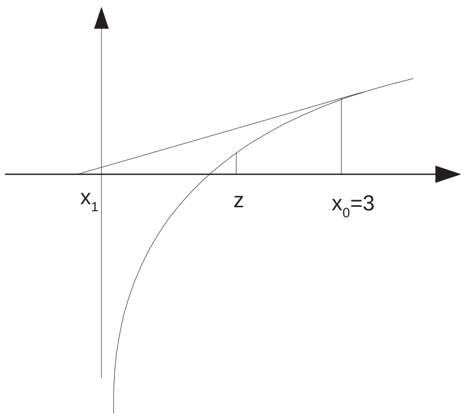

Figura 7: $z = x_0 + \frac{1}{2}(x_1 - x_0)$

Se considerarmos $x_0 = 3$, então tem-se

$$f(x_0) = \log 3, \quad f'(x_0) = \frac{1}{x_0} = \frac{1}{3}$$

e portanto

$$x_1 = 3 - \frac{\log 3}{\frac{1}{3}} = 3(1 - \log 3) < 0$$

Então $\log x_1$ não existe, ou seja, $g(x_1) = \infty$. Como $g'(x_0)(x_1 - x_0) < 0$ calcula-se

$$z = x_0 + \frac{1}{2}(x_1 - x_0) = \frac{3(1 - \log 3) + 3}{2} = \frac{3(2 - \log 3)}{2} > 0$$

Além disso

$$f(z) = \log[\frac{3}{2}(2 - \log 3)] = \log 3 + \log(2 - \log 3) - \log 2 > 0$$

Como a função logarítmica é estritamente crescente, então

$$log(2 - \log 3) - \log 2 < 0$$

e portanto

$$0 < f(z) < \log 3 = f(x_0) \Rightarrow g(z) < g(x_0)$$

Isso significa que o ponto

$$x_1 = z = \frac{3(2 - \log 3)}{2}$$

é aceite e o processo continuaria até à obtenção de um valor aproximado da raiz $\bar{x} = 1$.

Da descrição deste processo conclui-se que o método de Newton Global pode terminar com uma solução da equação não linear $f(x) = 0$ ou com um mínimo local (ou ponto de sela) da função de mérito que não seja solução da equação. Este último caso ocorre raramente na prática e pode ser ultrapassado reiniciando o método com um novo ponto inicial. A taxa de convergência do método de Newton Global mantém-se quadrática desde que o algoritmo convirja para um ponto $\bar{x}$ que seja uma raiz simples da equação não linear. Por isso o método de Newton Global é universalmente usado na prática quando o cálculo das derivadas da função f não é demasiado complexo. Como vimos anteriormente, a versão do algoritmo com derivadas aproximadas ou o método da secante são preferíveis no caso em que as derivadas são muito difíceis de calcular.

Referencias

- R. Burden, J. Douglas Faires e A. Reynolds, Numerical Analysis, Prindle, Weber and Schmidt, Massachusetts, 1981.
- J. Dennis e R. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM, Philadelphia, 1996.
- E. Fernandes, Computação Numérica, Universidade do Minho, 1998.
- J. A. Ferreira e M. F. Patrício, Análise Numérica, Departamento de Matemática, Universidade de Coimbra, 1998.
- J. Júdice, Funções Reais de Uma Variável Real, Departamento de Matemática, Universidade de Coimbra, 1988.
- J. Mathews e K. Fink, Numerical Methods using Matlab, Prentice Hall, New Jersey, 1999.
- Luís Nunes Vicente, Apontamentos de Matemática Numérica II, Departamento de Matemática, Universidade de Coimbra, 2006.
- H. Pina, Métodos Numéricos, Mc Graw-Hill, Lisboa, 1995.