Método de Newton: Convergência Local, Unicidade de Solução e Aceleração da Convergência

Fernando Ricardo Moreira[1]

Resumo

Neste artigo é feita uma discussão sobre o método de Newton. Tratamos sobre a importância da escolha do ponto inicial na definição da seqüência de Newton. Damos alguns exemplos que mostram que a escolha do ponto inicial é fundamental para garantir a boa definição da seqüência e também para assegurar que ela convergirá. Observamos que o comportamento da derivada na raíz está relacionado com a taxa de convergência do método. Também é feita uma análise sobre a convergência do método, unicidade de solução para o caso local e damos uma estimativa do tamanho do raio da maior bola que contém uma única solução. Por último mostramos que a convergência do método pode ser acelerada.

Palavras-Chave: Convergência, Unicidade de Solução, Aceleração da Convergência.

Newton's method: Local Convergence, Uniqueness of Solution and Convergence Acceleration

Abstract

The article made a discussion about the method of Newton. We deal on the importance of choosing the starting point in defining the sequence of Newton. We give some examples that show that the choice of starting point is crucial to ensure the proper definition of the sequence and also to ensure that they converge. We noticed that the behavior of the secondary root is related to the rate of convergence of the method. It's also made an analysis on the convergence of the method, uniqueness of solution to the event site and give an estimate of the size of the radius of the largest ball that contains a single solution. Finally we show that the convergence of the method can be accelerated.

Keywords: Convergence, Uniqueness of Solution, Accelerating Convergence.

[1]Professor da Universidade Católica de Goiás

1 Sobre o Método de Newton

1.1 Histórico e Discussão preliminar sobre o Método de Newton

A idéia básica do Método de Newton é bastante simples, consiste na linearização de uma função derivável. Suponha $f : \mathbb{R} \to \mathbb{R}$ diferenciável e queremos resolver a equação $f(x) = 0$. Começando de um ponto inicial x_0 podemos construir uma aproximação linear de $f(x)$ em uma vizinhança de x_0:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

Note que pela equação anterior, temos

$$0 = f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

Portanto podemos assumir que

$$0 = f(x_0) + f'(x_0)(x - x_0).$$

Que tem solução, digamos

$$x_1 = x_0 - f'(x_0)^{-1} f(x_0).$$

Agora, se $f'(x_1) \neq 0$ repetimos o processo e encontramos

$$x_2 = x_1 - f'(x_1)^{-1} f(x_1).$$

Portanto, na k-ésima iteração, se $f'(x_k) \neq 0$, encontramos

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

A equação anterior nos motiva para a seguinte definição.

Definition 1. A seqüência do método de Newton para resolver $f(x) = 0$ onde f é uma função derivável é a seqüência dada pela seguinte regra de recorrência:

$$x_{k+1} = x_k - f'(x_k)^{-1} f(x_k) \quad k = 0, 1, 2, \cdots . \tag{1}$$

Remark 2. Para esta definição fazer sentido temos que garantir que $x_k \in I$, isto é , que x_k esteja no domínio de f e que $f'(x_k) \neq 0$ para todo $k = 0, 1, 2, \cdots$.

A idéia do método é exatamente esta, o uso da expressão linearizada de uma função derivável em lugar da própria função, visto que a expressão em questão é uma boa aproximação local da função (Para entender melhor o que significa boa aproximação local de uma função veja [6]). O método foi proposto por Isaac Newton em 1669 para encontrar raízes de funções polinomiais. Pouco tempo depois, em 1690 J. Raphson extendeu o método para funções reais quaisquer. Por isso é muito comum, na literatura, o método ser chamado método de Newton-Raphson. A consolidação do método está ligada a famosos matemáticos como L. A. Cauchy, J. Fourier entre outros. Em 1818, Fourier provou que o método convergia quadraticamente desde que o ponto inicial fosse tomado em uma vizinhança da solução procurada, enquanto Cauchy (1829-1847) mostrou que o método se extende naturalmente para funções de várias variáveis e usou o método para provar a existência de raízes de algumas equações. Em 1916, os matemáticos Fine e Bennet deram mais contribuições para o método. Fine em [1] provou a convergência para o caso n-dimensional sem a hipótese de existência de solução. Bennet em [2] extendeu o resultado para o caso de dimensão infinita. Mais recentemente, em 1948, L. V. Kantorovich em [5] provou a existência de solução e a convergência do método para operadores $\mathrm{T} : \mathcal{B}_1 \to \mathcal{B}_2$ onde $\mathcal{B}_1$ e $\mathcal{B}_2$ são espaços de Banach e T é um operador diferenciável qualquer. Os resultados de Bennet e Kantorovich merecem um destaque especial pois foram obtidos antes da descoberta dos fundamentos da Análise Funcional. Para mais informações sobre o desenvolvimento do método e outros resultados veja [7, 8, 9].

Sejam $I \subset \mathbb{R}$ um intervalo aberto e $f : I \to \mathbb{R}$ uma função de classe $C^1(I)$ e $C^0(\overline{I})$, onde $\overline{I}$ denota o fecho de I.

Primeiro suponha que a seqüência definida pela equação (1) anterior está bem definida e que converge para um valor $x_* \in \overline{I}$. Se além disso, f' é limitada numa vizinhança de x_*, então $f(x_*) = 0$. De fato, basta notar que a equação que define o método de Newton é equivalente à equação

$$f(x_k) - f'(x_k)(x_{k+1} - x_k) = 0\,, \qquad k = 0, 1, 2, \ldots.$$

Agora, fazendo $k \to \infty$ na equação anterior e observando que f é contínua em $\overline{I}$, f' é limitada e $\lim x_k = x_*$, temos que $f(x_*) = 0$.

Na discussão anterior, supomos que a seqüência de Newton está bem definida, isto é, $x_k \in U$ e $f'(x_k)$ é não singular para todo $k \in \mathbb{N}$. Esta é uma das questões mais delicadas de se resolver. De fato, pode acontecer que algum ponto da seqüência de Newton esteja fora do domínio de definição da função f, ou ainda, que a derivada não esteja definida em algum ponto da seqüência.

2 Convergência do Método - Caso local

Nesta seção consideramos um operador $F : U \to \mathcal{B}_2$ Fréchet derivável sobre int(U), onde $U \subset \mathcal{B}_1$ e $\mathcal{B}_1, \mathcal{B}_2$ são espaços de Banach. Admitimos que a equação $F(x) = 0$ tem solução. Mostramos que, sob certas condições, a seqüência gerada pelo método de Newton para resolver $F(x) = 0$ está bem definida, converge para uma solução com taxa de convergência quadrática, desde que o ponto inicial x_0 seja tomado numa vizinhança apropriada da solução. A taxa quadrática de convergência é garantida através de uma hipótese sobre o comportamento de F' no ponto solução x_*.

Agora, vamos enunciar e demonstrar um teorema sobre a convergência do método de Newton para o caso local.

Theorem 3. *Seja $F : U \to \mathcal{B}_2$ um operador Fréchet derivável sobre o aberto convexo $U \subset \mathcal{B}_1$, onde $\mathcal{B}_1$ e $\mathcal{B}_2$ são espaços de Banach. Suponha que existam $L > 0$, $p \in (0, 1]$ e $x_* \in U$ tais que*

1. $F(x_*) = 0$ e $F'(x_*)$ é não singular ;
2. $B(x_*, 1/(2L\beta)^{1/p}) \subset U$, onde $\beta = \|F'(x_*)^{-1}\|$;

3. $F' \in \mathsf{Hol}_L^p\left(B(x_*, 1/(2L\beta)^{1/p})\right)$.

Theorem 4. *Então a seqüência $\{x_k\}$ gerada pelo método de Newton para resolver $F(x) = 0$, com ponto inicial $x_0 \in B(x_*, 1/(2L\beta)^{1/p})$*

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad k = 0, 1, \dots, \tag{2}$$

está bem definida, contida na bola $B(x_*, 1/(2L\beta)^{1/p})$ e converge para x_*. Além disso, a seqüência $\{x_k\}$ tem ordem de convergência $p+1$ para x_*, isto é,

$$\|x_{k+1} - x_*\| \leqslant \frac{2L\beta}{p+1}\|x_k - x_*\|^{p+1}, \qquad k = 0, 1, 2, \dots . \tag{3}$$

Para facilitar a demonstração do Teorema 3 precisaremos de dois resultados auxiliares. Primeiro o seguinte lema.

Lemma 5. *Seja $F : U \to \mathcal{B}_2$ um operador Fréchet derivável sobre o aberto convexo $U \subset \mathcal{B}_1$, onde $\mathcal{B}_1$ e $\mathcal{B}_2$ são espaços de Banach. Suponha que existam $L > 0$ e $x_* \in U$ que satisfaçam as mesmas condições do Teorema 3 . Então, para todo $x \in B(x_*, 1/(2L\beta)^{1/p})$ a derivada $F'(x)$ é não singular e vale a seguinte estimativa*

$$\left\|F'(x)^{-1}\right\| \leq 2\left\|F'(x_*)^{-1}\right\| = 2\beta. \tag{4}$$

Demonstração. Tome $x \in B(x_*, 1/(2L\beta)^{1/p})$. Por hipótese, $F' \in Hol_L^p(B(x_*, 1/(2L\beta)^{1/p}))$, assim após algumas manipulações algébricas, obtemos

$$\begin{aligned}\|F'(x_*)^{-1}\left(F'(x) - F'(x_*)\right)\| \quad &\leq \left\|F'(x_*)^{-1}\right\| \|F'(x) - F'(x_*)\| \\ &\leqslant \beta\, L\|x - x_*\|^p < \tfrac{1}{2}.\end{aligned}$$

Portanto temos que $F'(x)$ é inversível. Ainda, pelo mesmo teorema, temos que é valida a seguinte estimativa

$$\left\|F'(x)^{-1}\right\| \leq \frac{\left\|F'(x_*)^{-1}\right\|}{1 - \|F'(x_*)^{-1}(F'(x) - F'(x_*))\|} \leq 2\left\|F'(x_*)^{-1}\right\| = 2\beta,$$

pois, $\|F'(x_*)^{-1}\left(F'(x) - F'(x_*)\right)\| < 1/2$. Assim, completamos a demonstração. □

Queremos mostrar que a seqüência $\{x_k\}$ dada por (2) está bem definida. Para isso considere a seguinte definição.

Definition 6. Seja $F : U \to \mathcal{B}_2$ um operador Fréchet derivável, onde $U \subset \mathcal{B}_1$ é um aberto e $\mathcal{B}_1$, $\mathcal{B}_2$ são espaços de Banach. A iteração de Newton $\mathcal{N}_F : B(x_*, 1/(2L\beta)^{1/p}) \to \mathcal{B}_1$ é dada por

$$\mathcal{N}_F(x) = x - F'(x)^{-1}F(x).$$

Note que, a seqüência de Newton definida por (2) pode ser escrita de maneira equivalente da seguinte forma

$$x_{k+1} = \mathcal{N}_F(x_k), \qquad k = 0, 1, \ldots . \tag{5}$$

Até agora, o único resultado que temos é que a derivada $F'(x)$ é não singular para todo x na bola $B(x_*, 1/(2L\beta)^{1/p})$. Assim, dado $x_0 \in B(x_*, 1/(2L\beta)^{1/p})$, segue-se que

$$x_1 = \mathcal{N}_F(x_0),$$

está bem definido. Como não podemos garantir que x_1 esteja nessa bola, o próximo ponto da seqüência pode não estar bem definido. Para garantir que possamos aplicar indefinidamente a iteração de Newton sobre $x_0 \in B(x_*, 1/(2L\beta)^{1/p})$, e assim garantir que a seqüência de Newton $\{x_k\}$ está bem definida, precisaremos do seguinte lema.

Lemma 7. *Seja $F : U \to \mathcal{B}_2$ um operador Fréchet derivável sobre o aberto convexo $U \subset \mathcal{B}_1$, onde $\mathcal{B}_1$ e $\mathcal{B}_2$ são espaços de Banach. Suponha que existam $L > 0$ e $x_* \in U$ que satisfaçam as mesmas condições do Teorema 3. Então, para todo $x \in B(x_*, 1/(2L\beta)^{1/p})$ valem as seguintes afirmações:*

- $\|\mathcal{N}_F(x) - x_*\| \leqslant \dfrac{2L\beta}{p+1}\|x - x_*\|^{p+1};$

- $\|\mathcal{N}_F(x) - x_*\| \leqslant \dfrac{1}{p+1}\|x - x_*\|.$

Além disso, vale a seguinte inclusão

$$\mathcal{N}_F\left(B(x_*, 1/(2L\beta)^{1/p})\right) \subseteq B(x_*, 1/(2L\beta)^{1/p}).$$

Demonstração. Tome $x \in B(x_*, 1/(2L\beta)^{1/p})$. Primeiro observe que

$$\begin{aligned}\|\mathcal{N}_F(x) - x_*\| &= \|x - x_* - F'(x)^{-1}F(x)\| = \|F'(x)^{-1}(F(x_*) - F(x) - F'(x)(x_* - x))\| \\ &= \|F'(x)^{-1}E_F(x, x_*)\| \leqslant \|F'(x)^{-1}\|\,\|E_F(x, x_*)\|.\end{aligned}$$

onde usamos a definição de $\mathcal{N}_F(x)$ e que $F(x_*) = 0$. Agora note que, o Lema 5 implica que $\|F'(x)^{-1}\| \leqslant 2\beta$ e pode ser demonstrado que $\|E_F(x, y)\| \leqslant L\|y - x\|^{p+1}/(p+1)$, assim substituindo estas duas desigualdades na equação acima obtemos o ítem i.

Como $x \in B(x_*, 1/(2L\beta)^{1/p})$, temos $\|x - x_*\| < 1/(2L\beta)^{1/p}$. Assim, pelo ítem i, obtemos que o ítem ii também vale.

Finalmente, dado $x \in B(x_*, 1/(2L\beta)^{1/p})$ obtemos do ítem ii que

$$\|\mathcal{N}_F(x) - x_*\| \leqslant \frac{1}{p+1}\frac{1}{(2L\beta)^{1/p}} < \frac{1}{(2L\beta)^{1/p}},$$

e como x é arbitrário, temos que a inclusão do lema é verdadeira.

□

Lemma 8. *Façamos agora a demonstração do Teorema 3.*

Demonstração. Mostraremos por indução que a seqüência (2), ou equivalentemente que (5), está bem definida. Tome $x_0 \in B(x_*, 1/(2L\beta)^{1/p})$. Pelo Lema 5, temos que $F'(x_0)$ é não singular e assim x_1 está bem definido e, pelo Lema 7, vale

$$x_1 = \mathcal{N}_F(x_0) \in B(x_*, 1/(2L\beta)^{1/p}).$$

Suponha agora que x_k está bem definido e que $x_k \in B(x_*, 1/(2L\beta)^{1/p})$. Como $x_k \in B(x_*, 1/(2L\beta)^{1/p})$, pelo Lema 5, temos que $F'(x_k)$ é não singular e portanto x_{k+1} está bem definido e, pelo Lema 7, vale

$$x_{k+1} = \mathcal{N}_F(x_k) \in B(x_*, 1/(2L\beta)^{1/p}).$$

Logo, a seqüência $\{x_k\}$ está bem definida e contida na bola $B(x_*, 1/(2L\beta)^{1/p})$. Temos que $\{x_k\} \subset B(x_*, 1/(2L\beta)^{1/p})$. Assim, tomando $x = x_k$ no ítem $ii)$ do Lema 7 e usando a equação (5), obtemos a seguinte desigualdade

$$\|x_{k+1} - x_*\| \leq \frac{1}{p+1} \|x_k - x_*\|, \qquad k = 0, 1, 2, ...$$

Como $p > 0$, a desigualdade anterior implica imediatamente que $\{x_k\}$ converge para x_*. Agora vamos demonstrar a desigualdade (3). Sabemos que $\{x_k\} \in B(x_*, 1/(2L\beta)^{1/p})$. Aplicando o Lema 7 ítem $i)$ com $x = x_k$ e usando a equação (5), obtemos a desigualdade desejada. Assim, concluímos a demonstração do Teorema 3 .

$\square$

Portanto, podemos observar que a hipótese sobre a não singularidade de F no ponto solução x_* assegurou que a taxa convergência da seqüência fosse, maior que a linear, podendo até mesmo ser quadrática, quando $p = 1$.

Corollary 9. *Seja $F : U \to \mathcal{B}_2$ um operador de classe $C^2(U)$, onde $U \subset \mathcal{B}_1$ é um conjunto aberto e convexo, $\mathcal{B}_1$, $\mathcal{B}_2$ são espaços de Banach e $\mathcal{B}_1$ é de dimensão finita. Suponha que exista $x_* \in U$ tal que $F(x_*) = 0$ e $F'(x_*)$ é não singular. Então existe $\rho > 0$ tal que $B(x_*, \rho) \subset U$ e a sequência $\{x_k\}$ gerada pelo método de Newton para resolver $F(x) = 0$,* *com ponto inicial* $x_0 \in B(x_*, \rho)$

$$x_{k+1} = x_k - F'(x_k)^{-1} F(x_k), \quad k = 0, 1, ...$$

está bem definida, contida na bola $B(x_*, \rho)$ *e converge para* x_**. Além disso, a sequência* $\{x_k\}$ *converge quadraticamente para* x_**, isto é,*

$$\|x_{k+1} - x_*\| \leqslant \frac{1}{2\rho} \|x_k - x_*\|^2, \quad k = 0, 1, 2, ...$$

3 Unicidade de Solução - Caso Local

Nesta seção vamos tratar da unicidade local de solução para a equação $F(x) = 0$ onde F é um operador que satisfaz as hipóteses do Teorema 3. Para mais detalhes sobre as demonstrações dos resultados que iremos enunciar veja [4].

Começamos com o seguinte teorema.

Theorem 10. *Seja* $F : U \to \mathcal{B}_2$ *um operador Fréchet derivável sobre o aberto convexo* $U \subset \mathcal{B}_1$*, onde* $\mathcal{B}_1$ *e* $\mathcal{B}_2$ *são espaços de Banach. Suponha que existam* $L > 0$*,* $p \in (0, 1]$ *e* $x_* \in U$ *tais que*

1. $F(x_*) = 0$ e $F'(x_*)$ é não singular ;
2. $B\left(x_*, ((p+1)/(\beta L))^{1/p}\right) \subset U$, onde $\beta = \|F'(x_*)^{-1}\|$;
3. $F' \in \mathrm{Hol}_L^p\left(B(x_*, ((p+1)/(\beta L))^{1/p})\right)$.

Theorem 11. *Então* x_* *é o único zero de* F *em* $B(x_*, ((p+1)/(\beta L)^{1/p}))$*. Além disso, temos que* $B(x_*, ((p+1)/(\beta L))^{1/(\beta p)})$ *é a maior bola em* U *contendo um único zero de* F*.*

Sejam $y_* \in B(x_*, ((p+1)/(\beta L))^{1/p})$ tal que $F(y_*) = 0$. Como a bola aberta $B(x_*, ((p+1)/(\beta L))^{1/p})$ é convexa, temos que

$$x_* + t(y_* - x_*) \in B(x_*, ((p+1)/(\beta L))^{1/p}), \qquad \forall\, t \in [0, 1].$$

Por hipótese $F' \in \mathrm{Hol}_L^p(B(x_*, ((p+1)/(\beta L))^{1/p}))$, deste modo temos

$$\left\| F'(x_*)^{-1} \left[\int_0^1 F'(x_* + t(y_* - x_*))dt - F'(x_*) \right] \right\| \leqslant \|F'(x_*)^{-1}\| L \|y_* - x_*\|^p \int_0^1 t^p dt$$
$$= \frac{\beta L}{p+1} \|y_* - x_*\|^p.$$

Agora, como $y_* \in B(x_*, ((p+1)/(\beta L))^{1/p})$, temos da desigualdade anterior que

$$\left\| F'(x_*)^{-1} \left[\int_0^1 F'(x_* + t(y_* - x_*))dt - F'(x_*) \right] \right\| < 1.$$

Portanto, pela desigualdade acima, segue-se que o operador linear

$$T := \int_0^1 F'(x_* + t(y_* - x_*))dt,$$

é inversível. Por outro lado, pelo Teorema Fundamental do Cálculo, temos

$$\int_0^1 F'(x_* + t(y_* - x_*))dt(y_* - x_*) = \int_0^1 \frac{d}{dt}[F(x_* + t(y_* - x_*))]dt = F(y_*) - F(x_*).$$

Como $F(x_*) = F(y_*) = 0$, obtemos da equação acima e definição de T que

$$T(y_* - x_*) = 0.$$

A igualdade anterior implica que $x_* = y_*$, pois T é um operador linear e inversível. Assim, x_* é o único zero de F em $B(x_*, ((p+1)/(\beta L))^{1/p})$. Agora vamos mostrar que $B(x_*, ((p+1)/(\beta L))^{1/p})$ é a maior bola que contém um único zero de F. Para mostrar que esta afirmação é verdadeira, considere

$$L > 0$$

$p \in (0, 1]$ e a função $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} -\dfrac{1}{\beta}t + \dfrac{L}{p+1}t^{p+1}, & t \geq 0 \\ -\dfrac{1}{\beta}t + \dfrac{L}{p+1}(-t)^{p+1}, & t < 0 \end{cases}$$

onde $\beta = \|f'(0)^{-1}\|$. Note que as \'unicas raízes reais de f são:

$$t_1 = -\left(\frac{p+1}{\beta L}\right)^{1/p}, \qquad t_2 = 0, \qquad t_3 = \left(\frac{p+1}{\beta L}\right)^{1/p}.$$

Observe que a função f satisfaz as hipóteses do Teorema 10, com $x_* = t_2 = 0$. Também vemos facilmente que $x_* = 0$ é a única raiz de f em $B(x_*, ((p+1)/(\beta L))^{1/p})$. Temos ainda que, para todo $\delta > 0$ a bola $B(x_*, ((p+1)/(\beta L))^{1/p} + \delta)$ contém mais que uma raiz de f. Portanto $B(x_*, ((p+1)/(\beta L))^{1/p})$ é a maior bola contendo uma única singularidade do operador F.

4 Aceleração da Taxa de Convergência - Caso Local

Vimos na seção anterior que a taxa de convergência do método de Newton pode ser quadrática. Nesta seção queremos responder a seguinte questão: Para qual classe de funções a taxa convergência do método de Newton pode ser acelerada? Veremos que supondo algumas hipóteses sobre f, a função para a qual se deseja aplicar o método, podemos construir uma nova função F que tem as mesmas raízes de f, e que a taxa de convergência do método para F é maior do que para f. Portanto, poderemos acelerar a convergência do método de Newton. Faremos apenas o caso em que f é uma função real.

Para mostrar que a afirmação acima é verdadeira faremos as seguintes hipóteses sobre a função $f : I \to \mathbb{R}$, onde $I \subseteq \mathbb{R}$é um intervalo aberto. Hipóteses:

1. $f \in C^m(I)$, onde $m > 1$ é um número natural;

2. f possui uma raiz simples em $x_* \in I$, isto é $f(x_*) = 0$ e $f'(x_*) \neq 0$.

Para responder a questão levantada considere o seguinte teorema.

Theorem 12. *Sejam I um intervalo aberto e $f : I \to \mathbb{R}$ uma função que satisfaz as hipóteses 1, e 2 acima. Suponha que $f''(x_*) = f'''(x_*) = \cdots = f^{(m-1)}(x_*) = 0$ e que $f^{(m)}(x_*) \neq 0$. Então existe $\delta > 0$ e $x_0 \in [x_* - \delta, x_* + \delta]$ tal que a seqüência de Newton $\{x_k\}$, com ponto inicial x_0, dada por*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, \ldots, \tag{6}$$

está bem definida, contida no intervalo $[x_* - \delta, x_* + \delta]$, converge para x_* e satisfaz

$$|x_{k+1} - x_*| \leqslant C|x_k - x_*|^m, \qquad k = 0, 1, \ldots,$$

para alguma constante C.

Demonstração. Por hipótese, f é pelo menos $C^2(I)$, assim aplicando o Corolário **??**, com $I = U$ e $\mathcal{B}_1 = \mathcal{B}_2 = \mathbb{R}$, temos que existe $\rho > 0$ tal que a seqüência de Newton $\{x_k\}$, com ponto inicial x_0, está bem definida, contida no intervalo $[x_* - \rho, x_* + \rho] \subset I$ e converge para x_*. Resta mostrar a desigualdade do teorema.

Primeiro note que sendo f' e $f^{(m)}$ são contínuas, $f'(x_*) \neq 0$ e $f^{(m)}(x_*) \neq 0$, existe $0 < \bar{\rho} < \rho$ tal que

$$|f'(x)| \geqslant c_0 > 0 \qquad |f^{(m)}(x)| \leqslant c_1, \qquad \forall\, x \in [x_* - \bar{\rho}, x_* + \bar{\rho}]. \tag{7}$$

Sejam as seguintes constantes $C > 0$ e $\delta > 0$ definidas por

$$C := \frac{mc_1 + c_1}{m!c_0}, \qquad \delta := \min\left\{\bar{\rho},\ \sqrt[m]{\bar{\rho}/C}\right\}.$$

Agora, vamos mostrar a desigualdade por indução. Tome $x_0 \in [x_* - \delta, x_* + \delta]$. Pela expansão de Taylor de f e f' em x_0 temos que, existem ξ_1 e ξ_2 entre x_*e x_0 tais que

$$f(x_0) = f'(x_*)(x_0 - x_*) + \frac{f^{(m)}(\xi_1)}{m!}(x_0 - x_*)^m, \qquad f'(x_0) = f'(x_*) + \frac{f^{(m)}(\xi_2)}{(m-1)!}(x_0 - x_*)^{m-1},$$

são verdadeiras. Usando a definição da seqüência de Newton dada por 6, temos

$$x_1 - x_* = x_0 - x_* - \frac{f(x_0)}{f'(x_0)} = \frac{1}{f'(x_0)}\left[f'(x_0)(x_0 - x_*) - f(x_0)\right].$$

Substituindo as expressões acima, dadas pela expansão de Taylor, na equação anterior obtemos após algumas manipulações algébricas que

$$x_1 - x_* = \frac{1}{f'(x_0)}\left[\frac{f^{(m)}(\xi_2)}{(m-1)!} - \frac{f^{(m)}(\xi_1)}{m!}\right](x_0 - x_*)^m = \frac{mf^{(m)}(\xi_2) - f^{(m)}(\xi_1)}{f'(x_0)m!}(x_0 - x_*)^m.$$

Como $x_0 \in [x_* - \delta, x_* + \delta]$, obtemos de (7) e das definições de C e δ que

$$\left|\frac{mf^{(m)}(\xi_2) - f^{(m)}(\xi_1)}{f'(x_k)m!}\right| \leqslant \frac{mc_1 + c_1}{m!c_0} = C.$$

Substituindo esta desigualdade, na equação acima, obtemos

$$|x_1 - x_*| \leqslant C|x_0 - x_*|^m,$$

e isto prova a desigualdade do teorema para $k = 0$. Agora, como $x_0 \in [x_* - \delta, x_* + \delta]$, segue-se da última desigualdade e definição de δ que

$$|x_1 - x_*| \leqslant \delta,$$

isto é, $x_1 \in [x_* - \delta, x_* + \delta]$. Analogamente, supondo que $x_k \in [x_* - \delta, x_* + \delta]$, podemos mostrar que a desigualdade vale para $k+1$ e que $x_{k+1} \in [x_* - \delta, x_* + \delta]$. Portanto a desigualdade vale para todo k e o teorema está provado.

$\square$

Nosso objetivo agora é encontrar um função F que possui as mesmas raízes de f, mas o método de Newton aplicado a F tem taxa de convergência maior que a taxa de convergência do método aplicado para f. Para isso considere o seguinte teorema.

Theorem 13. *Seja $f : I \to \mathbb{R}$ uma função que satisfaz as hipóteses h_1 e h_2. Suponha que $f'(x_*) > 0$, $f''(x_*) = f'''(x_*) = \cdots = f^{(m-1)}(x_*) = 0$ e que $f^{(m)}(x_*) \neq 0$. Então a função*

$$F(x) = \frac{f(x)}{\sqrt[m]{f'(x)}}, \tag{8}$$

está bem definida num certo intervalo contendo x_* e satisfaz $F(x_*) = 0$, $F'(x_*) > 0$, $F''(x_*) = F'''(x_*) = \cdots = F^{(m)}(x_*) = 0$.

Demonstração. Como $f'(x_*) > 0$, existe um intervalo $J \subset I$ que contém x_*, de modo que

$$g(x) := 1/\sqrt[m]{f'(x)}, \qquad x \in J,$$

está bem definida, isto é, $f'(x) > 0$ para todo $x \in J$. Assim a equação anterior pode ser reescrita equivalentemente da seguinte forma

$$1 = g(x)^m f'(x), \qquad x \in J.$$

Derivando a expressão anterior obtemos

$$0 = mg(x)^{m-1}g'(x)f'(x) + g(x)^m f''(x), \qquad x \in J,$$

e portanto, como $g(x) \neq 0$ em J, temos que

$$0 = mg'(x)f'(x) + g(x)f''(x), \qquad x \in J. \tag{9}$$

Como $f''(x_*) = 0$ e $f'(x_*) > 0$, temos $g'(x_*) = 0$. Derivando a última equação obtemos

$$0 = mg''(x)f'(x) + (m+1)g'(x)f''(x) + g(x)f'''(x), \qquad x \in J.$$

Agora, derivando a equação anterior, encontramos

$$0 = mg'''(x)f'(x) + (2m+1)g''(x)f''(x) + (m+2)g'(x)f'''(x) + g(x)f^{(4)}(x), \qquad x \in J.$$

Portanto, a derivada de ordem $k-1$ da equação (9) é dada por

$$\begin{aligned} 0 = \quad & mg^{(k)}(x)f'(x) + (m(k-1)+1)g^{(k-1)}(x)f''(x) \quad + \cdots \\ & + \tfrac{(k-1)!}{j!(k-j)!}(j + m(k-j))g^{(k-j)}(x)f^{(j+1)}(x) \qquad + \cdots \\ & + (m+k-1)g'(x)f^{(k)}(x) + g(x)f^{(k+1)}(x), \qquad x \in J. \end{aligned} \tag{10}$$

Por hipótese, temos que $f''(x_*) = \cdots = f^{(m-1)}(x_*) = 0$. Portanto a última equação implica que $g^{(k)}(x_*) = 0$ para

$1 \leqslant k \leqslant m-2$.Agora, fazendo $k = m-1$ e $x = x_*$ a equação (10) se reduz a

$$mg^{(m-1)}(x_*)f'(x_*) + g(x_*)f^{(m)}(x_*) = 0, \tag{11}$$

enquanto, para $k = m$ e $x = x_*$, tem-se que

$$mg^{(m)}(x_*)f'(x_*) + g(x_*)f^{(m+1)}(x_*) = 0.$$

Agora, vamos investigar a função $F(x) := f(x)g(x)$. A primeira e a segunda derivada de F são dadas respectivamente por

$$F'(x) = f(x)g'(x) + f'(x)g(x), \qquad F''(x) = f(x)g''(x) + 2f'(x)g'(x) + f''(x)g(x), \qquad x \in J.$$

Assim, após alguns cálculos, obtemos que a k-ésima derivada de F é dada por

$$F^{(k)}(x) = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} f^{(j)}(x)g^{(k-j)}(x), \qquad x \in J. \tag{12}$$

Por definição $F(x_*) = 0$. Agora, usando a hipótese sobre o comportamento de f em x_*, os resultados obtidos sobre o comportamento das derivadas de g no mesmo ponto x_* e a fórmula para a expressão da derivada de ordem k de F, obtemos após alguns cálculos que

$$F'(x_*) = f'(x_*)g(x_*) = f'(x_*)^{1-1/m} > 0, \qquad F^{(k)}(x_*) = 0, \qquad 2 \leqslant k \leqslant m-1.$$

Agora, pelas equações (11) e (12), temos que

$$F^{(m)}(x_*) = mf'(x_*)g^{(m-1)}(x_*) + f^{(m)}(x_*)g(x_*) = 0.$$

Assim, concluímos a nossa demonstração.

$\square$

Observe que a função $F(x) = f(x)/\sqrt[m]{f'(x)}$ satisfaz as hipóteses do Teorema 12, com $F(x_*) = 0$, $F'(x_*) \neq 0$, $F''(x_*) = F'''(x_*) = \cdots = F^{(m)}(x_*) = 0$. Portanto a taxa de convergência do método de Newton para F é de ordem pelo menos $m+1$.

Para mais detalhes sobre a aceleração do método de Newton veja [3]. Agora faremos um exemplo numérico para mostrar como essa aceleração funciona. Analisaremos, à luz das idéias anteriores, a função no exemplo a seguir.

Example 14. Considere a função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x/\sqrt{1+x^2}$.

Seja $x_* = 0$, a raíz de f. Note que

$$f'(x) = \frac{1}{(1+x^2)^{3/2}}; \quad f''(x) = -\frac{3x}{(1+x^2)^{5/2}}; \quad f'''(x) = \frac{-3+12x^2}{(1+x^2)^{7/2}};$$

e portanto $f'(x_*) \neq 0$, $f''(x_*) = 0$ e $f'''(x_*) \neq 0$. Logo, pelo Teorema 12, temos que existe uma vizinhança V de x_* tal que a taxa de convergência do método de Newton para f é cúbica. Para acelerar a convergência tome $F(x) = f(x)/\sqrt[3]{f'(x)}$ e note que

$$F(x) = x \qquad \text{e} \qquad x_+ = x - \frac{F(x)}{F'(x)} = 0.$$

Assim, o método de Newton aplicado à função F converge em um único passo!

Referências

[1] Fine, H., **On Newton's Method of Approximation**, Proc. Nat. Acad. Sci. USA 10, (1916), pp 546–552.

[2] Bennet, A.A., **Newon's Method in General Analysis**. Proc. Nat. Acad. Sci. USA 10, (1916) pp 592–598.

[3] Gerlach, J. **Accelerated Convergence in Newton's Method**, SIAM Review, 2, (1994), pp 272–276.

[4] Huang, Zhengda., **The Convergence Ball of Newton's Method and the Uniqueness Ball of Equations's under Hölder-TypeContinuos Derivatives**, Computers and Mathematics with Applications. 47, (2004), pp 247–251.

[5] Kantorovich, L. V., **On Newton's Method for functional Equations**, Dokl. Akad. Nauk SSSR,), (1948), pp 1237–1240.

[6] Lima, E. L. **Curso de Análise Vol. 2**, Projeto Euclides, IMPA (1995).

[7] Mathews, J. H., **Bibliography for Newton's Method**, math.fullerton.edu/mathews/new\\tonsmethod/ Newton'sMethodBib/Links/Newton'sMethodBib_lnk_3.html.

[8] Polyak, B. T., **Newton-Kantorovich Method and its Global Convergence**, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.(POMI) 312 (2004).

[9] Ypma, T, J., **Historical Developement of the Newton-Raphson Method**, SIAM Review, (4), (1995), pp. 531–551.