# Uma Discussão Sobre o Método de Newton

Fernando Ricardo Moreira [1]
Rodrigo Couto Santos[2]
Raimundo Rodrigues Gomes Filho [3]
Helder Barbosa Paulino [4]

**Resumo**

O Método de Newton é usado em quase todas as áreas da matemática aplicada e é parte do programa da disciplina de cálculo numérico visto pelos alunos dos cursos da área de exatas. Damos uma idéia de como o Método surgiu e de como veio a consolidar-se ao longo de quase quatro séculos de descobrimento. Falamos um pouco a respeito de dois tipos de convergência: a linear e a quadrática. Na última parte fizemos alguns exemplos e por último demonstramos um teorema que diz que o Método converge quadraticamente desde que o ponto inicial da sequência de Newton seja tomado numa vizinhança apropriada da solução.

Palavras-Chave : Convergência, Sequência, Taxa de Convergência.

## A Discussion About the Method of Newton

**Abstract**

The Newton method is used in almost all areas of applied mathematics and is part of the program of the discipline of calculation number seen by the students of the courses in the area of exact. We an idea of how the method has emerged and how to consolidate over nearly four centuries of discovery. We talk a just about two types of convergence: the linear and quadratically. In the last part we have made some examples and finally demonstrate a theorem that says that the method converges quadraticly provided the starting point of the sequence of Newton is taken in a neighborhood of the appropriate solution.

Keywords: Convergence, Sequence, Convergence Rate.

## 1 Método de Newton

### 1.1 Histórico e Idéias sobre o Método

A idéia básica do Método de Newton é bastante simples, consiste na linearização de uma função derivável. Suponha *f*: R → 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 $x_0$:

$$f(x) \approx f(x_0) + f_0'(x_0)(x - x_0)$$

Note que pela equação anterior, temos

$$0 = f(x) \approx f(x_0) + f_0'(x_0)(x - x_0)$$

[1] Matemático, UFG - Jataí, E-mail: moreirafrmat@hotmail.com
[2] Engenheiro Agrícola, UFG - Jataí, E-mail: rodrigoambiencia@gmail.com
[3] Agrônomo, UFG - Jataí, E-mail: rrgomesfilho@hotmail.com
[4] Agrônomo, UFG - Jataí, E-mail: helderlino51@yahoo.com.br

Portanto podemos assumir que

$$0 = f(x_0) + f_0'(x_0)(x - x_0).$$

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.

**Definição 1.1.** A sequência do método de Newton para resolver $f(x) = 0$ onde *f* é uma função derivável é a sequência dada pela seguinte regra de recorrência:

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

**Observação 1.2.** Para esta definição fazer sentido temos que garantir que $x_k \in \mathrm{I}$, isto é , que $x_k$ esteja no domínio de *f* e que $f'(x_k) \neq 0$ para todo $k = 0, 1, 2, ...$

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 [5]). 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 [3] provou a convergência para o caso n-dimensional sem a hipótese de existência de solução. Bennet em [1] extendeu o resultado para o caso de dimensão infinita. Mais recentemente, em 1948, L. V. Kantorovich em [4] provou a existência de solução e a convergência do método para operadores *T*: $B_1 \rightarrow B_2$ onde $B_1$ e $B_2$ 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 [6, 7, 8].

## 1.2 Taxa de Convergência

Queremos tratar agora da taxa de convergência. Definiremos dois tipos de convergência: a linear e a quadrática. Vamos mostrar que a velocidade da convergência quadrática é bem maior do que a da linear. Daremos um sentido preciso sobre o que significa velocidade de uma convergência. Para isso consideremos $\{x_k\} \in R$ uma sequência convergente. Digamos, $\lim_{x \to \infty} x_k = x_* \in R$.

**Definição 1.3.** (Sobre a Taxa de Convergência e Erro cometido)
i) $\{x_k\}$ converge linearmente para $x_*$ se existe $c \in [0,1]$ tal que

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

ii) $\{x_k\}$ converge quadraticamente para $x_*$ se existe *c* > 0 tal que

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

iii) o erro cometido na *k*-ésima iteração é

$$e_k = |x_k - x_*|, \qquad k = 0, 1, 2, ...$$

Agora daremos sentido ao que referimos anteriormente como velocidade de uma convergência.

• <u>Convergência Linear</u>

$$e_k \leq ce_{k-1} \Rightarrow e_k \leq c^k e_0, \qquad k = 0, 1, 2, ...$$

Suponha que na *k*-ésima iteração tenhamos

$$e_k \leq c^k e_0 \leq 10^{-n} \tag{2}$$

Esta desigualdade significa que o erro $e_k = |x_k - x_*|$ é da ordem de $10^{-n}$, em outras palavras, "que $x_k$ tem pelo menos *n* dígitos exatos". De (2) temos,

$$\log_{10}\left(c^k e_0\right) \leq \log_{10}\left(10^{-n}\right) \Leftrightarrow \log_{10}\left(c^k\right) + \log_{10}(e_0) \leq -n \Leftrightarrow k\log_{10}(c) \leq -n - \log_{10}(e_0),\ c \in (0,1)$$

$$\Leftrightarrow k(-\log_{10}(c)) \geq n + \log_{10}(e_0) \Leftrightarrow k \geq \frac{n}{-\log_{10}(c)} + \frac{\log_{10}(e_0)}{-\log_{10}(c)} \tag{3}$$

De (3) segue que, para dobrar o número de dígitos exatos será necessário dobrar o número de iterações, isto é

$$k \geq \frac{n}{-\log_{10}(c)} + \frac{\log_{10}(e_0)}{-\log_{10}(c)} \Leftrightarrow 2k \geq \frac{2n}{-\log_{10}(c)} + \frac{\log_{10}(e_0)}{-\log_{10}(c)}$$

• <u>Convergência Quadrática</u>

$$e_k \leq ce_{k-1}^2 \Rightarrow e_k \leq c^{2^k-1} e_0^{2^k} \tag{4}$$

De fato, por indução temos:

$$k = 0 \Rightarrow e_0 \leq c^{2^0-1} e_0^{2^0} = e_0$$

Suponhamos que (4) seja valida para *k* e vamos provar que vale para *k + 1*. Note que

$$e_{k+1} \leq c(e_k^2) \leq c(c^{2^k-1} e_0^{2^k})^2 = c^{2^{k+1}-1} e_0^{2^{k+1}}$$

portanto vale para k + 1.

Suponha que na *k*-ésima iteração tenhamos

$$e_k \leq c^{2^k-1} e_0^{2^k} \leq 10^{-n}$$

De (5) temos,

$$\log_{10}\left(c^{2^k-1}e_0^{2^k}\right) \leq \log_{10}\left(10^{-n}\right) \Leftrightarrow (2^k-1)\log_{10}(c) + 2^k \log_{10}(e_0) \leq -n$$
$$\Leftrightarrow 2^k(\log_{10}(c) + \log_{10}(e_0)) \leq -n + \log_{10}(c) \Leftrightarrow 2^k(\log_{10}(ce_0)) \leq -n + \log_{10}(c)$$
$$\Leftrightarrow 2^k \geq \frac{n}{-\log_{10}(ce_0)} + \frac{\log_{10}(c)}{-\log_{10}(ce_0)} \tag{6}$$

Onde assumimos, sem nenhuma perda de generalidade, que $ce_0 \in (0,1)$.

De (6) segue que para dobrar o número de dígitos exatos será necessário apenas mais uma iteração, isto é

$$2^k \geq \frac{n}{-\log_{10}(ce_0)} + \frac{\log_{10}(c)}{-\log_{10}(ce_0)} \Leftrightarrow 2^{k+1} \geq \frac{2n}{-\log_{10}(ce_0)} + \frac{\log_{10}(c)}{-\log_{10}(ce_0)}$$

**Observação 1.4.** Suponhamos que *f*: I → R é uma função derivável. Seja $\{x_k\}$ a sequência gerada pelo método de Newton. Suponha que $\{x_k\}$ está bem definida, isto é, $x_k \in$ I e $f'(x_k) \neq 0$ $\forall$ $k \in$ N e que converge para algum $x_*$, então $f(x_*) = 0$.

De fato, a equação (4) é equivalente à equação

$$f(x_k) + f'_0(x_{k+1} - x_k) = 0$$

Agora fazendo $k \to +\infty$ na última igualdade temos que $f(x_*) + f'_0(x_* - x_*) = 0$. Portanto $f(x_*) = 0$. (Assumimos que *f* é contínua em $\bar{I}$, o menor conjunto fechado contento o conjunto $I$.)

Ilustraremos agora com alguns exemplos numéricos.

**Exemplo 1.5.** Considere a função *f*: R → R dada por $f(x) = \dfrac{x}{\sqrt{1+x^2}}$

Note que f(x) = 0 implica que x = 0 e

$$f'(x) = \frac{1}{(1+x^2)^{3/2}} \neq 0 \qquad \forall x \in R$$

Agora observe que

$$x_{k+1} = -x_k^3, \qquad \forall\ k = 0,1,2,...$$

Assim:

• Para $x_0$ = 1 a sequência de Newton oscila entre 1 e –1; Para $x_0$ < 1 a sequência de Newton converge para 0 com taxa cúbica! Para $x_0$ > 1 a sequência de Newton diverge.

**Observação 1.6.** Note que $f'(0) = 1 \neq 0$, e também que $|x_{k+1}| = |x_k|^3$ e portanto $\{x_k\}$ tem convergência cúbica para $|x_0| < 1$.

**Exemplo 1.7.** Considere a função *f*: R → R dada por $f(x) = x^2$.
Note que f(x) = 0 implica que x = 0. Temos que $f'(x) = 2x \neq 0$ para x≠0 e

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

Veja que $\forall\, x_0 \in R - \{0\}$ a seqüência converge, isto é, $\lim_{k\to\infty} x_k = 0$. Como $|x_{k+1}| = (1/2)|x_k|$ temos que a convergência é linear. Observe ainda que neste exemplo *f'(0) = 0.*

**Exemplo 1.8.** Considere a função $f : (0,+\infty) \to R$ dada por

$$f(x) = 1 - \frac{1}{x}$$

Note que f(x) = 0 implica que x = 1. Temos que $f'(x) = \frac{1}{x^2} \neq 0 \quad \forall\, x \in R$ e

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

Observe que:
• $x_0$ = 2 implica que $x_1 = 0 \notin (0,+\infty)$ e $\forall\, x_0 > 2$ temos $x_1 \notin (0,+\infty)$. Desta forma a sequência não está bem definida; Se 0 < $x_0$ < 2 então a sequência está bem definida e converge para $x_* = 1$.

**Exemplo 1.9.** Considere a função *f*: R → R dada por *f(x) = x(x – 1)(x + 1).*
Note que *f(x)* = 0 implica que as raízes da equação são 0, 1 e –1. Observe que *f(x)* = $x^3 - x$. Com isso temos que $f'(x) = 3x^2 - 1$ e

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{x_k^3 - x_k}{3x_k^2 - 1} = \frac{2x_k^3}{3x_k^2 - 1}, \qquad \forall\, k = 0,1,2,...$$

Assim:

*i*) Para $x_0 = 1/\sqrt{5}$ a sequência de Newton oscila entre $1/\sqrt{5}$ e $-1/\sqrt{5}$; Para $x_0$ = 1/2 temos que $x_1$ = –1. Portanto, a sequência de Newton é finita (converge); Para $x_0$ = –1/2 temos que $x_1$ = 1. Portanto a sequência de Newton é finita (converge);

*ii*) Tomando $x_0$ = –0, 465444..., isto é, a raiz real do polinômio de terceiro grau $2\sqrt{3x^3} - 3x^2 + 1$ temos que

$$\frac{1}{\sqrt{3}} = \frac{2x_0^3}{3x_0^2 - 1}$$

logo $x_1 = 1/\sqrt{3}$. Assim, o método de Newton gera um ponto singular da derivada, neste caso, a sequência de Newton não está bem definida. Note que

$$-1/\sqrt{3} = -0{,}577350... < x_0 = -0{,}46544... < -1/\sqrt{5} = -0{,}447213...$$

Para finalizar demonstraremos um teorema sobre a convergência local da sequência gerada pelo método de Newton para encontrar uma solução da equação *f(x) = 0*. Mostraremos que é possível obter convergência quadrática desde que o ponto inicial seja tomado numa vizinhança da solução $x_*$ e que $f'(x_*)$ seja não singular.

**Teorema 1.10.** Seja *f: I* ⊂ → R uma função de classe C$^2$ (Ser de classe C$^2$ significa poder ser derivável 2 vezes ). Suponha que exista $x_* \in I$ tal que $f(x_*) = 0$ e $f'(x_*) \neq 0$. Então existem A e B > 0 e ε > 0 tal que para todo $x_0 \notin (x_* - \varepsilon,\ x_* + \varepsilon)$ a seqüência

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

está bem definida, converge para $x_*$ e

$$\left|x^* - x_{k+1}\right| \leq \frac{A}{2B}\left|x_* - x_k\right|^2 \qquad k = 0,1,2,...$$

Demonstração. Como $f'(x_*) \neq 0$ e f é de classe C$^2$ existem constantes A, B > 0 e $\delta > 0$ tais que $\left|f''(x)\right| \leq A \;\; e \;\; \left|f'(x)\right| \geq B$ para todo $x \in (x_* - \delta,\ x_* + \delta)$.

Seja $\varepsilon = \min\left\{\delta, \frac{B}{A}\right\}$, assim para todo $x \in (x_* - \varepsilon,\ x_* + \varepsilon)$ temos $\left|f''(x)\right| \leq A \;\; e \;\; \left|f'(x)\right| \geq B$. Dado $x_0 \in (x_* - \varepsilon,\ x_* + \varepsilon)$ suponhamos que $x_k \in (x_* - \varepsilon,\ x_* + \varepsilon)$. Então existe *c* entre $x_*$ e $x_k$ tal que

$$0 = f(x_*) = f(x_k) + f'(x_k)(x_* - x_k) + \frac{1}{2} f''(c)(x_* - x_k)^2$$

equivalentemente,

$$f'(x_k)x_k - f(x_k) - f'(x_k)x_* = \frac{1}{2} f''(c)(x_k - x_*)^2$$

ou ainda

$$f'(x_k)(x_k - f'(x_k)^{-1} f(x_k) - x_*) = \frac{1}{2} f''(c)(x_k - x_*)^2$$

Da última expressão temos,

$$x_{k+1} - x_* = \frac{1}{2}\frac{f''(c)}{f'(x_k)}(x_k - x_*)^2 \tag{7}$$

De (7) temos

$$\left|x_{k+1} - x_*\right| = \frac{1}{2}\frac{\left|f''(c)\right|}{\left|f'(x_k)\right|}\left|x_k - x_*\right|^2 \le \frac{A}{2B}\left|x_k - x_*\right|^2 \tag{8}$$

Agora mostraremos que a sequência converge. De (8) temos

$$\left|x_{k+1} - x_*\right| \le \frac{A}{2B}\left|x_k - x_*\right|^2 = \frac{A}{2B}\left|x_k - x_*\right|\left|x_k - x_*\right| \le \frac{A}{2B}\varepsilon\left|x_k - x_*\right| \le \frac{A}{2B}\frac{B}{A}\left|x_k - x_*\right| \le \frac{1}{2}\left|x_k - x_*\right|$$

Assim da última desigualdade temos a convergência da sequência concluindo a prova do teorema.

**Observação 1.11.** Na prova do teorema anterior assumimos que se $x_0 \in (x_* - \varepsilon,\ x_* + \varepsilon)$ então $x_k \in (x_* - \varepsilon,\ x_* + \varepsilon)$. Isso pode ser demonstrado assumindo a hipótese de que *f'* seja localmente Lipschitz, e como assumimos que *f* é de classe $C^2$ temos que *f'* é localmente Lipschitz.

**Referências Bibliográficas:**

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

[2] Dennis, J. E. Jr, Schnabel, R. B., *Numerical Methods for Unconstrained Optimization and Nonlinear Equations Classics* in Applied Mathematics, SIAM (1996).

[3] Fine, H., **On Newton's Method of Approximation**, Proc. Nat. Acad. Sci. USA 2(10), (1916), p. 546–552.

[4] Kantorovich, L. V., **On Newton's Method for functional Equations**, Dokl. Akad. Nauk SSSR, 59(7), (1948), p. 1237–1240.

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

[6] Mathews, J. H., **Bibliography for Newton's Method**. Disponível em <http://math.fullerton.edu/mathews/newtonsmethod/Newton'sMethodBib/Links/Newton'sMethodBibInk3.html>. Acesso em 05/abr/2009.

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

[8] Ypma, T, J., **Historical Developement of the Newton-Raphson Method**. SIAM Review, 37(4), (1995), p. 531–551.