UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

O Método de Newton

Uma Análise de Convergência Local e Semi-Local - o Teorema de Kantorovich

por

Fernando Ricardo Moreira

Orientador: Dr. Orizon Pereira Ferreira

Dissertação de Mestrado em Matemática
Goiânia - Goiás
2006

Universidade Federal de Goiás
Instituto de Matemática e Estatística
Coordenação de Pós-Graduação em Matemática

O Método de Newton

Uma Análise de Convergência Local e

Semi-Local - o Teorema de Kantorovich

por

Fernando Ricardo Moreira

Área de Concentração : **Otimização**
Orientador: **Dr. Orizon Pereira Ferreira**

Dissertação submetida à Banca Examinadora designada pelo Conselho Diretor do Instituto de Matemática e Estatística da Universidade Federal de Goiás, como parte dos requisitos necessários à obtenção do grau de Mestre em Matemática.

Goiânia - Goiás
2006

Dedico este trabalho a minha família, em especial a minha amada esposa Samantha e ao nosso bebê que ainda está na barriga da mamãe.

Agradecimentos

Ao Deus da minha salvação, pois me tem dado mais do que eu poderia imaginar e me tem feito andar em lugares altos.

Ao grande amor da minha vida, minha querida esposa Samantha por todo amor dedicado a mim e ao nosso querido bebê, o Pedro, que já estamos ansiosos para vê-lo.

Ao Prof. Dr. Orizon Pereira Ferreira por uma orientação impecável, pela dedicação dispensada a mim e por acreditar que eu desempenharia um bom trabalho.

Aos meus professores do IME/UFG que muito contribuíram para a minha formação e também aos funcionários do IME/UFG pela amizade.

Aos meus pais, irmãs e sobrinhas que sempre acreditaram e tiveram bastante orgulho de mim.

A Igreja Assembléia de Deus de Vila União pelo amparo espiritual.

Aos meus amigos Alex e Cristiane, Dayves e Lorena, Diogo e Camila, Giovanni e Kiara, Marcos e Cláudia, Sérgio e Dayse, Weder e Aczia.

Ao CNPq pelo apoio financeiro.

*11 - As palavras dos sábios são como aguilhões, e como pregos bem fixados as sentenças
coligidas, dadas pelo único Pastor. 12 - Demais, filho meu, atenta: não há limites para
fazer livros, e o muito estudar é enfado da carne. 13 - De tudo o que se tem ouvido,
a suma é: Teme a Deus e guarda os seus mandamentos; porque isto é o dever de todo
homem. 14 - Porque Deus há de trazer a juízo todas as obras, até as que estão escondidas,
quer sejam boas, quer sejam más.*

Bíblia Sagrada - Livro de Eclesiastes.

Sumário

Resumo

A busca por soluções de equações não-lineares em espaços de Banach, é objeto de interesse em várias áreas da ciência e engenharias. Devido a sua velocidade de convergência e eficiência, o método de Newton tem sido bastante utilizado para o propósito de obter soluções dessas equações. Nesta dissertação apresentamos uma análise de convergência local do método de Newton. Esta análise tem a desvantagem de exigir o conhecimento prévio de um zero do operador em consideração e hipóteses sobre o comportamento do operador nesse zero, mas por outro lado ela fornece informações sobre a taxa de convergência e unicidade local de solução. Também apresentamos uma análise de convergência semi-local, mais especificamente, apresentamos uma prova do teorema de Kantorovich. Este teorema não exige conhecimento prévio de zeros do operador e tem a vantagem de fazer hipóteses apenas sobre o ponto inicial, garantindo convergência quadrática, existência e unicidade local de zeros. A análise de convergência a ser apresentada, tanto para o caso local e quanto para o semi-local, é baseada na definição de um conjunto invariante pela aplicação de Newton.

Abstract

The search for solutions of nonlinear equations in Banach spaces, is object of interest in several areas of science and engineerings. Had it is speed of convergence and efficiency, the method of Newton has been sufficiently used for the intention to get solutions of these equations. In this dissertação we present a local analysis of convergence of the method of Newton. This analysis has the disadvantage to demand the previous knowledge of a zero of the operator in consideration and hypotheses on the behavior of the operator in this zero, but for other side it supplies to information on the convergence rate and local unicity of solution. Also we present a semilocal analysis of convergence, more specifically, present a proof of Kantorovich's theorem. This theorem does not demand previous knowledge of zeros of the operator and has the advantage to make hypotheses only on the initial point, guaranteeing quadratic convergence, existence and local unicity of zeros. The analysis of to be presented convergence, as much for the local case and how much for the semi-local, it is based on the definition of a set invariant for the application of Newton.

Capítulo 1

Introdução

A idéia básica do Método de Newton, para encontrar uma raíz de uma função derivável não-linear, é bastante simples. Ela consiste em substituir a função não-linear por sua aproximação linear. Mais precisamente, queremos resolver a equação $f(x) = 0$, onde f é uma função não-linear. Começando com um ponto inicial x_0 do domínio de f tal que a derivada $f'(x_0)$ seja não nula, construímos a aproximação linear de f, isto é,

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

Então substituímos a equação $f(x) = 0$ pela equação linear

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

que tem uma única solução, digamos

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

Agora, com hipóteses adequadas sobre f e o ponto inicial x_0, podemos provar que x_1 pertence ao domínio de f e que $f'(x_1) \neq 0$ e o processo anterior pode ser repetido e assim sucessivamente. Portanto, podemos definir uma seqüência $\{x_k\}$ no domínio de f dada por

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

Novamente, hipóteses adequadas sobre f e o ponto inicial x_0 não só garantem a boa definição da seqüência $\{x_k\}$, como convergência quadrática da seqüência para uma raíz da função em consideração.

Vamos falar um pouco sobre o desenvolvimento do método de Newton. O método descrito acima foi proposto inicialmente 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 J. Fourier, L. A. Cauchy entre outros. Em 1818, Fourier provou que o método convergia quadraticamente desde que o ponto inicial fosse tomado em uma vizinhança da solução procurada (faremos uma demonstração desse fato), 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 algumas contribuições para o método. Fine em [8] 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 [11] provou a existência de solução e a convergência do método para operadores 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 derivável qualquer. Os resultados de Bennet e Kantorovich merecem um destaque especial, pois foram obtidos antes da descoberta dos fundamentos da Análise Funcional, portanto podemos considerá-los como alguns dos precursores dessa área tão importante e bela da matemática. Para mais informações sobre o desenvolvimento do método de Newton veja [14, 18, 20] e suas referências.

Sabemos que tratar questões sobre a resolução de equações, ou mesmo mostrar que um dada equação possui solução dentro de um certo espaço considerado, é objeto de interesse em diversas áreas da Ciência Pura e Engenharia. O método de Newton é uma poderosa ferramenta na obtenção de soluções de equações dos mais variados tipos. A grande importância desse método reside no fato de que sob algumas hipóteses é garantida a convergência, a uma taxa relativamente alta, para uma solução. O método de Newton também é usado para mostrar outros teoremas importantes na matemática. Por exemplo: teoremas de existência e unicidade de soluções para certas equações diferenciais veja por exemplo [12, 15], o teorema da função inversa e implícita [13] e o teorema do mergulho isométrico veja [16].

Procurando saber mais sobre a solução de um problema em equações diferenciais foi que L. V. Kantorovich chegou ao seu teorema, que é o objeto de maior interesse nesta dissertação. Desde então, o teorema de Kantorovich tem se tornado uma ferramenta fundamental em análise não-linear e tem sido extensivamente usado na análise numérica, veja por exemplo [1, 5, 17] e sua referências. Wang [21] mostra em 1999, um teorema da função inversa em espaços de Banach usando um teorema do "tipo Kantorovich". Em 2002, Ferreira e Svaiter [6] estenderam o teorema de Kantorovich para variedades Riemannianas utilizando uma nova técnica. Mais recentemente, Potra [19] em 2005 usou o teorema de Kantorovich para fazer uma análise do método de pontos interiores.

Esta dissertação está organizada da seguinte forma. No capítulo 2 nós escrevemos alguns resultados básicos sobre espaços de Banach e derivação de operadores. Ainda nesse capítulo, demonstramos o lema de Banach e damos uma estimativa para o erro cometido em aproximar um operador, com derivada Hölder, pela sua expressão linearizada. O capítulo 3 é onde se concentra a discussão sobre o método de Newton. Na Seção 3.1 exploramos alguns exemplos, que clareiam nosso entendimento sobre quais são os problemas que podem aparecer na construção da seqüência de Newton e na análise da convergência. Na Seção 3.2 demonstramos, para o caso local, um teorema sobre a boa definição e convergência da seqüência de Newton para uma solução dada. Baseamos, a Seção 3.2 na parte do livro de Dennis e Schnabel [4] que trata sobre a questão da convergência local do método de Newton. As provas desta seção foram ligeiramente modificadas, procuramos aplicar a técnica que consite em obter um conjunto invariante pela iteração de Newton. Esta técnica também é usada na nossa referencia básica Ferreira e Svaiter [7]. Na Seção 3.3 mostramos a unicidade de solução numa bola de certo raio, e que essa bola é a maior que contém uma única solução, baseado nas idéias contidas no artigo de Huang[10]. Por fim, na Seção 3.4 mostramos que, sob certas hipóteses, podemos acelerar a taxa de convergência do método, baseado no artigo de Gerlach [9]. Finalmente, no Capítulo 4 provamos o teorema de Kantorovich. Todas as idéias contidas neste capítulo foram baseadas na artigo Ferreira e Svaiter [7]. Na introdução deste capítulo enunciamos o teorema de Kantorovich e na Subseção 4.0.1, obtemos resultados importantes sobre uma função quadrática, definida a posteriori. Na Seção 4.1 garantimos a boa definição da seqüência de Newton e a sua convergência para uma solução. E finalmente, na Seção 4.2 mostramos que a solução é única, em uma bola de certo raio, e que a taxa de convergência para uma solução poderá ser quadrática.

Capítulo 2

Conceitos Básicos e Resultados Preliminares

2.1 Espaços de Banach

Nesta seção definimos espaços de Banach, nossos espaços ambiente de trabalho nesta dissertação. Enunciamos e demonstramos algumas proposições e lemas que nos auxiliarão na obtenção de alguns resultados importantes. Definimos a derivada de Fréchet, que é a derivada para operadores em espaços de Banach, e enunciamos alguns resultados que envolvem esse tipo de derivação.

Começamos com a seguinte definição.

Definição 2.1. *Sejam $(\mathcal{V}, ||.||)$ um espaço vetorial normado e $\{x_n\} \subset \mathcal{V}$ uma seqüência. Dizemos que $\{x_n\}$ converge para $x_* \in \mathcal{V}$ se dado $\epsilon > 0$ existe n_0 tal que*

$$||x_n - x_*|| < \epsilon \quad \forall \quad n \geqslant n_0.$$

Uma seqüência $\{x_n\}$ é chamada seqüência de Cauchy se dado $\epsilon > 0$ existe n_0 tal que

$$||x_m - x_n|| < \epsilon \quad \forall \quad m, n \geqslant n_0.$$

É fácil observar que toda seqüência convergente é de Cauchy, mas nem toda seqüência de Cauchy é convergente. Veja o exemplo a seguir.

Exemplo 2.2. *Seja* $\{x_n\}$ *uma seqüência de números racionais definida por*

$$x_n = \left(1 + \frac{1}{n}\right)^n, \qquad n = 1, 2, \ldots .$$

Temos que $\{x_n\}$ *é de Cauchy em* $\mathbb{R}$*, pois* $\lim x_n = e$*, onde* $e \in \mathbb{R}$ *é o número de Euler. Portanto, temos que* $\{x_n\}$ *é também de Cauchy em* $\mathbb{Q}$ *e não é convergente em* $\mathbb{Q}$*.*

Definição 2.3. *Um espaço vetorial* $\mathcal{B}$ *é dito Espaço de Banach se toda seqüência de Cauchy em* $\mathcal{B}$ *é convergente, isto é, dada uma seqüência de Cauchy* $\{x_n\} \subset \mathcal{B}$*, existe* $x_* \in \mathcal{B}$ *tal que* $\lim x_n = x_*$*.*

Exemplo 2.4. *O espaço euclideano* $\mathbb{R}^n$ *e o espaço das funções contínuas com a métrica do supremo são espaços de Banach. O primeiro exemplo tem dimensão finita e o segundo não.*

Consideremos o conjunto $\mathcal{L}(C, \mathcal{B})$ das transformações lineares e contínuas de C em $\mathcal{B}$, onde C é um espaço vetorial métrico qualquer e $\mathcal{B}$ um espaço de Banach. Observe que $||.|| \ : \ \mathcal{L}(C, \mathcal{B}) \to \mathbb{R}$ é contínua, onde

$$\|\mathcal{F}\| = \sup_{\|x\| \neq 0} \frac{\|\mathcal{F}(x)\|}{\|x\|}, \qquad x \in C, \ \mathcal{F} \in \mathcal{L}(C, \mathcal{B}).$$

Nosso intuito agora é obter um lema que nos auxiliará na demonstração de um teorema muito importante para resultados futuros nessa dissertação.

Lema 2.5. *Sejam* $\mathcal{B}$ *um espaço de Banach e* $\{t_k\}$ *uma seqüência de números reais monótona crescente e convergente para* t_**. Considere uma seqüência* $\{x_k\} \subset \mathcal{B}$ *que satisfaça*

$$||x_{k+1} - x_k|| \leqslant t_{k+1} - t_k, \qquad \forall \ k \ \in \mathbb{N}.$$

Então a seqüência $\{x_k\}$ *é de Cauchy. Em particular,* $\{x_k\}$ *converge, digamos para* $x_* \in \mathcal{B}$*. Além disso, vale a desigualdade*

$$||x_* - x_k|| \leqslant t_* - t_k, \qquad \forall \ k \ \in \mathbb{N}.$$

Demonstração. Seja $n \in \mathbb{N}$. Usando a desigualdade triangular temos

$$\begin{aligned}
||x_{k+n} - x_k|| &\leqslant ||x_{k+n} - x_{k+n-1}|| + ||x_{k+n-1} - x_{k+n-2}|| + \cdots + ||x_{k+2} - x_{k+1}|| + ||x_{k+1} - x_k|| \\
&\leqslant (t_{k+n} - t_{k+n-1}) + (t_{k+n-1} - t_{k+n-2}) + \cdots + (t_{k+2} - t_{k+1}) - (t_{k+1} - t_k) \\
&= t_{k+n} - t_k \leqslant t_* - t_k.
\end{aligned}$$

Como $\{t_k\}$ é uma seqüência convergente para t_*, segue-se que $t_* - t_k$ pode ser tomado suficientemente pequeno para k arbitrariamente grande. Logo a equação acima implica que $\{x_k\}$ é uma seqüência de Cauchy.

Para obtermos a desigualdade requerida pelo lema, é suficiente fazer $n \to \infty$ na inequação acima. □

Lema 2.6. *Sejam $\mathcal{B}$ um espaço de Banach e C um espaço vetorial métrico qualquer. Então $\mathcal{L}(C, \mathcal{B})$ também é um espaço de Banach.*

Demonstração. Seja $\{T_n\}$ uma seqüência de Cauchy em $\mathcal{L}(C, \mathcal{B})$. Observemos que, para todo $x \in C$ e $m, n \geqslant 1$, temos

$$||T_m(x) - T_n(x)|| \leqslant ||T_m - T_n||\, ||x||.$$

Logo fixado $x \in C$ a seqüência $\{T_n(x)\}$ é de Cauchy em $\mathcal{B}$. Portanto existe $y \in \mathcal{B}$ tal que $T_n(x) \to y$. Definindo $T(x) = y$ temos que T é linear, pois T_n é linear e contínua para todo $n \in \mathbb{N}$. Da desigualdade triangular, obtemos

$$\big|\, ||T_m|| - ||T_n||\, \big| \leqslant ||T_m - T_n||.$$

Logo $\{||T_n||\}$ é de Cauchy em $\mathbb{R}$ e assim é convergente. Portanto existe $C > 0$ tal que $||T_n|| \leqslant C$ para todo n. Pela continuidade da norma temos

$$||T(x)|| = \lim_{n\to\infty} ||T_n(x)|| \leqslant \lim_{n\to\infty} ||T_n||\, ||x|| \leqslant C||x||$$

Logo T é limitado e portanto contínuo. Agora queremos concluir que $T_n \to T$. Como

$$||\,(T - T_n)\,(x)|| = \lim_{m\to\infty} ||\,(T_m - T_n)\,(x)|| \leqslant \lim_{m\to\infty} \sup_{x\in\mathcal{B}} ||T_m - T_n||\, ||x||,$$

temos que

$$||T-T_n|| = \sup_{x\neq 0} \frac{||\big(T-T_n\big)(x)||}{||x||} \leqslant \lim_{m\to\infty} ||T_m - T_n||,$$

que pode ser feito suficientemente pequeno tomando n arbitrariamente grande. Logo, temos que $\{T_n\}$ é convergente. Portanto $\mathcal{L}(C,\mathcal{B})$ é espaço de Banach. $\square$

Com a mesma notação do lema anterior, passaremos agora a demonstrar o seguinte teorema.

Teorema 2.7. *Seja $A \in \mathcal{L}(C,\mathcal{B})$ inversível tal que $||A^{-1}(B-A)|| < 1$. Então $B \in \mathcal{L}(C,\mathcal{B})$ é inversível e valem as desigualdades*

$$||B^{-1}A|| \leqslant \frac{1}{1-||A^{-1}(B-A)||}, \qquad ||B^{-1}|| \leqslant \frac{\|A^{-1}\|}{1-||A^{-1}(B-A)||}. \tag{2.1}$$

Demonstração. Primeiro vamos mostrar que se E é um operador linear qualquer sobre um espaço de Banach é tal que $\|E\| < 1$ então $I - E$ é inversível e vale

$$\|(I-E)^{-1}\| \leqslant \frac{1}{1-\|E\|}.$$

Para isso, considere as seguintes seqüências $\{S_k\}$ e $\{t_k\}$ definidas, respectivamente, por:

$$S_k = I + E + E^2 + \cdots + E^k, \qquad t_k = 1 + \|E\| + \|E\|^2 + \cdots + \|E\|^k.$$

Primeiro note que

$$\begin{aligned} \|S_{k+1} - S_k\| = &\ \ \|(I + E + \cdots + E^{k+1}) - (I + E + \cdots + E^k)\| \\ &\leqslant \|E\|^{k+1} = t_{k+1} - t_k. \end{aligned}$$

Agora, como $\|E\| < 1$, temos que $\{t_k\}$ é uma seqüência monótona crescente e convergente, com limite $t_* = 1/(1 - \|E\|)$. Portanto, pelo Lema 2.5 temos que $\{S_k\}$ é uma seqüência de Cauchy em $\mathcal{L}(C,\mathcal{B})$ que, pela proposição anterior, é espaço de Banach. Logo existe $\lim\limits_{n\to\infty} S_n$. Agora, observe que

$$(I-E)\, S_k = (I-E).(I + E + \cdots + E^k) = I - E^{k+1}. \tag{2.2}$$

$$S_k.(I-E) = (I+E+\cdots+E^k).(I-E) = I - E^{k+1}. \tag{2.3}$$

Por outro lado, temos que $\lim_{k\to\infty} I - E^k = I$, pois

$$||I-(I-E^k)|| = ||E^k|| \leqslant ||E||^k \quad \text{e} \quad \lim_{k\to\infty} ||E||^k = 0.$$

Assim, pelas equações (2.2) e (2.3) temos que $\lim_{k\to\infty} S_k = (I-E)^{-1}$. Note ainda que

$$||(I-E)^{-1}|| = \;\; ||\lim_{k\to\infty} S_k|| \leqslant \lim_{k\to\infty} \left(||I|| + ||E|| + \cdots + ||E^k||\right) = \lim_{k\to\infty} t_k = \frac{1}{1-||E||}.$$

Agora, tomando $E = -A^{-1}(B-A)$ e observando a hipótese $\|A^{-1}(B-A)\| < 1$, temos que $(I-E) = A^{-1}B$ é inversível e vale a estimativa para a norma da inversa $B^{-1}A$. Para concluir nossa demonstração basta mostrar que vale a estimativa para a norma de B^{-1}. Para isso é suficiente multiplicar primeira inequação em (2.1) por $\|A^{-1}\|$ e observar que $\|B^{-1}\| \leqslant \|B^{-1}A\|\,\|A^{-1}\|$. □

2.2 Derivada de Fréchet

Nesta seção definimos a derivada de operadores em espaços de Banach.

Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ espaços de Banach, $\mathcal{L}(\mathcal{B}_1, \mathcal{B}_2)$ o espaço dos operadores lineares contínuos de $\mathcal{B}_1$ em $\mathcal{B}_2$, $U \subset \mathcal{B}_1$ e $F : U \to \mathcal{B}_2$ um operador não linear qualquer.

Definição 2.8. *Um operador* $\mathrm{T} \in \mathcal{L}(\mathcal{B}_1, \mathcal{B}_2)$ *é a derivada de Fréchet de* F *em* $x_0 \in U$ *se para todo* $u \in \mathcal{B}_1$ *tal que* $u + x_0 \in U$ *vale*

$$\lim_{\|u\|\to 0} \frac{1}{\|u\|} \|F(x_0+u) - F(x_0) - Tu\| = 0.$$

Dizemos que F é Fréchet derivável em U se F for Fréchet derivável em todo ponto $x \in U$. Denotaremos a derivada de Fréchet de um operador F em x por $F'(x)$. Se $F' : C \to \mathcal{L}(\mathcal{B}_1, \mathcal{B}_2)$ é contínua em U dizemos que F é de classe C^1 em U, e usamos a seguinte notação $F \in C^1(U)$. Se o operador F' for derivável em U, e o operador $F'' : C \to \mathcal{L}(\mathcal{B}_1, \mathcal{L}(\mathcal{B}_1, \mathcal{B}_2))$ for contínuo, então dizemos que F é de classe C^2 em U e denotamos por $F \in C^2(U)$.

É importante ressaltar que, para a derivação de Fréchet, valem resultados semelhantes aos da derivação no espaço $\mathbb{R}^n$, isto é, valem as seguintes proposições:

Proposição 2.9. *(Regra da Cadeia) Sejam $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ espaços de Banach, $U \subset \mathcal{B}_1$, $V \subset \mathcal{B}_2$ conjuntos abertos. Sejam $F : U \to \mathcal{B}_2$, $G : V \to \mathcal{B}_3$ operadores contínuos tais que $F(U) \subset V$, $y_0 = F(x_0)$ e as derivadas de Fréchet $F'(x_0)$ e $G'(y_0)$ existem. Então GoF é Fréchet derivável e*

$$(GoF)'(x_0) = G'(y_0)oF'(x_0).$$

Proposição 2.10. *(Teorema da Função Inversa) Sejam $U \subset \mathcal{B}_1$ um conjunto aberto e $F : U \to \mathcal{B}_2$ um homeomorfismo de U sobre o conjunto aberto $V = F(U) \subset \mathcal{B}_2$. Suponha que F seja Fréchet derivável em $x_0 \in U$ e que $F'(x_0) : \mathcal{B}_1 \to \mathcal{B}_2$ seja um homeomorfismo linear. Então o operador inverso $G = F^{-1} : V \to \mathcal{B}_1$ é Fréchet derivável em $y_0 = F(x_0)$ e*

$$G'(y_0) = F'(x_0)^{-1}.$$

Proposição 2.11. *(Teorema Fundamental do Cálculo) Sejam $\mathcal{B}_1, \mathcal{B}_2$ espaços de Banach, $U \subset \mathcal{B}_1$ um conjunto aberto e convexo e $F : U \to \mathcal{B}_2$ um operador Fréchet derivável. Para todos $x, y \in U$ temos que*

$$\int_0^1 \frac{d}{dt}[F(x + t(y - x))]dt = F(y) - F(x).$$

Proposição 2.12. *(Desigualdade do Valor Médio) Sejam $\mathcal{B}_1, \mathcal{B}_2$ espaços de Banach, $U \subset \mathcal{B}_1$ um conjunto aberto. Sejam x_0, $x_0 + h \in U$, tal que o segmento*

$$[x_0, x_0 + h] = \{x_0 + \tau h, 0 \leqslant \tau \leqslant 1\} \subset U.$$

Seja $F : U \to \mathcal{B}_2$ Fréchet derivável para todo $x \in [x_0, x_0 + h]$. Então

$$\|F(x_0 + h) - F(x_0)\| \leqslant \sup\{\|F'(x_0)\| : x \in [x_0, x_0 + h]\}\ \|h\|.$$

Para mais informações sobre a derivada de Fréchet e as demonstrações das proposições acima veja [12].

Sabemos que o método de Newton é baseado na linearização de uma função diferenciável e também que existe um erro cometido devido a essa linearização. Por isso é importante, para questões de convergência do método, controlar esse erro de linearização, isto é, obter uma estimativa do tamanho desse erro. Para isso, considere um operador $F : U \to \mathcal{B}_2$ Fréchet derivável em $\text{int}(U) \subset \mathcal{B}_1$, onde $\mathcal{B}_1$, $\mathcal{B}_2$ são espaços de Banach. Com as mesmas notações acima, vamos demonstrar uma estimativa para esse erro. Antes, daremos as seguinte definição.

Definição 2.13. *Um operador $F : U \to \mathcal{B}_2$ é chamado Hölder contínuo de ordem $p \in (0, 1]$ com constante $L > 0$ se*

$$\|F(x) - F(y)\| \leqslant L\|x - y\|^p, \qquad \forall\, x, y \in U.$$

Denotamos por $\text{Hol}_L^p(U)$ o conjunto dos operadores Holder contínuos em U de ordem p e com constante L. Em particular quando $p = 1$ na definição acima dizemos que F é Lipschitz e denotamos o conjunto dos operadores Lipschtiz em U com constante L por $\text{Lip}_L(U)$.

Lema 2.14. *Sejam $U \subset \mathcal{B}_1$ um conjunto convexo e $F : U \to \mathcal{B}_2$ um operador contínuo e Fréchet derivável em int(U), tal que $F' \in Hol_L^p(int(U))$. Então, temos a seguinte estimativa do erro de linearização*

$$\|E_F(x, y)\| = \|F(y) - F(x) - F'(x)(y - x)\| \leqslant \frac{L}{p+1}\|y - x\|^{p+1}, \qquad x \in int(U)\,, y \in U.$$

Demonstração. Dados $x \in \text{int}(U)\,, y \in U$. Pela convexidade de U temos $x + t(y - x) \in \text{int}(U)$, para todo $t \in [0, 1)$. Assim pelo Teorema 2.11, temos que

$$F(y) - F(x) - F'(x)(y - x) = \int_0^1 [F'(x + t(y - x)) - F'(x)](y - x)dt.$$

Portanto, como $F' \in \text{Hol}_L^p(\text{int}(U))$ a igualdade acima implica que

$$\begin{aligned} \|F(y) - F(x) - F'(x)(y - x)\| &\leqslant \int_0^1 \|F'(x + t(y - x)) - F'(x)\|\, \|y - x\|dt \\ &\leqslant \int_0^1 L\|y - x\|^{p+1}t^p dt = \frac{L}{p+1}\|y - x\|^{p+1}. \end{aligned}$$

que é a desigualdade esperada. □

Capítulo 3

O Método de Newton

Nesta seção é feita uma discussão sobre o método de Newton. Vamos tratar 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 relacionada 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.

3.1 Discussão preliminar sobre o Método de Newton

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. Começamos com a seguinte definição. O método de Newton para resolver a equação $f(x) = 0$ é definido por

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

Primeiro suponha que a seqüência definida pela equação 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. Para ficar mais claro o que estamos dizendo, considere os seguintes exemplos.

Exemplo 3.1. *Seja $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^3$. Sabemos que $x = 0$ é a única raíz de f. Note que $f'(x) = 3x^2 \neq 0$ se $x \neq 0$. Portanto, para todo ponto inicial $x_0 \neq 0$ temos que a seqüência de Newton está bem definida e vale*

$$x_{k+1} = \frac{2}{3} x_k\,, \qquad k = 0, 1, 2, \ldots .$$

Portanto, temos que independente da escolha do ponto inicial, a seqüência de Newton para resolver $f(x) = 0$ está bem definida e converge para a única raíz da função, $x_* = 0$. Para referência futura note que a taxa de convergência da seqüência é linear e que $f'(x_*) = 0$.

Exemplo 3.2. *Considere a função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x/\sqrt{1 + x^2}$.*

Note que $x = 0$ é a única raíz da equação $f(x) = 0$ e que

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

Agora observe que para todo ponto inicial $x_0 \neq 0$ a seqüência de Newton está bem definida e vale

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

Assim:

i) *se* x_0 *for tomado tal que* $|x_0| \geqslant 1$*, a seqüência de Newton não convergirá;*

ii) *se* x_0 *for tomado tal que* $|x_0| < 1$*, a seqüência de Newton convergirá para* $x_* = 0$ *com taxa cúbica.*

Portanto, temos que a seqüência de Newton está bem definida para todo valor inicial que tomarmos, mas a convergência dela é condicionada ao ponto inicial x_0 *ser tomado tal que* $|x_0| < 1$*. Logo, temos que a escolha do ponto inicial é importante para garantir a convergência da seqüência de Newton.*

Exemplo 3.3. *Considere a função* $f : (0, +\infty) \to \mathbb{R}$ *dada por* $f(x) = 1 - 1/x$*.*

Note que $f(x) = 0 \Leftrightarrow x = 1$ *e que* $f'(x) = 1/x^2 \neq 0$ *para todo* x*. A seqüência de Newton para encontrar a solução de* $f(x) = 0$ *é dada por*

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

Note que:

i) *se* $x_0 = 2$ *implica que* $x_1 = 0 \notin (0, +\infty)$ *e para todo* $x_0 > 2$ *temos* $x_1 \notin (0, +\infty)$*. Desta forma a seqüência de Newton não está bem definida;*

ii) *se* $0 < x_0 < 2$ *então a seqüência está bem definida e converge para* $x_* = 1$*.*

Agora se $x_0 \geqslant 2$ *o ponto* x_1 *não está no domínio da função. Logo, para este caso, temos que a boa definição da seqüência e a convergência dela é condicionada à condição do ponto inicial ser tomado próximo da solução.*

Exemplo 3.4. *Considere a função* $f : \mathbb{R} \to \mathbb{R}$ *dada por* $f(x) = x(x-1)(x+1) = x^3 - x$*.*

Note que as raízes da equação $f(x) = 0$ *são* 0*,* 1 *e* -1*. Observe que* $f'(x) = 3x^2 - 1$ *e a seqüência de Newton para resolver* $f(x) = 0$ *é dada por*

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

Assim:

i) *se* $|x_0| = 1/\sqrt{5}$ *a seqüência de Newton oscila entre* $1/\sqrt{5}$ *e* $-1/\sqrt{5}$*;*

ii) *se* $x_0 = 1/2$ *temos que* $x_1 = -1$*. Portanto, a seqüência de Newton é finita (converge!);*

iii) *se* $x_0 = -1/2$ *temos que* $x_1 = 1$*. Portanto, a seqüência de Newton é finita (converge!);*

iv) *Tomando* $x_0 = -0,465444\ldots$*, isto é, a raíz real do polinômio de terceiro grau* $p(x) = 2\sqrt{3}x^3 - 3x^2 + 1$*, 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 seqüência de Newton não está bem definida.*

Portanto, se $|x_0| = 1/2$ *temos que a seqüência de Newton converge em um único passo! Note também que, se* x_0 *for tomado como sendo a raíz real do polinômio* $q(x) = 2\sqrt{3}x^3 + 3x^2 - 1$*, teremos* $x_1 = -1/\sqrt{3}$ *que também é um ponto singular da derivada, logo a seqüência de Newton não estaria bem definida com essa escolha de* x_0*.*

3.2 Convergência do Método - Caso local

Nesta seção consideramos um operador $F : U \to \mathcal{B}_2$ Fréchet derivável sobre $\text{int}(U)$, 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.

Teorema 3.5. *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 $L > 0$, $p \in (0, 1]$ e $x_* \in U$ tais que*

i) $F(x_*) = 0$ *e* $F'(x_*)$ *é não singular ;*

ii) $B(x_*, 1/(2L\beta)^{1/p}) \subset U$, *onde* $\beta = \|F'(x_*)^{-1}\|$;

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

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

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad k = 0, 1, \ldots, \tag{3.1}$$

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,

$$\|x_{k+1} - x_*\| \leqslant \frac{2L\beta}{p+1}\|x_k - x_*\|^{p+1}, \qquad k = 0, 1, 2, \ldots . \tag{3.2}$$

Para facilitar a demonstração do Teorema 3.5 precisaremos de dois resultados auxiliares. Primeiro o seguinte lema.

Lema 3.6. *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 3.5 .*

Então, para todo $x \in B(x_*, 1/(2L\beta)^{1/p})$ a derivada $F'(x)$ é não singular e vale a seguinte estimativa

$$\|F'(x)^{-1}\| \leqslant 2\|F'(x_*)^{-1}\| = 2\beta. \tag{3.3}$$

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

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

Portanto, pelo Teorema 2.7 temos que $F'(x)$ é inversível. Ainda, pelo mesmo teorema, temos que é valida a seguinte estimativa

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

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 (3.1) está bem definida. Para isso considere a seguinte definição.

Definição 3.7. *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 (3.1) pode ser escrita de maneira equivalente da seguinte forma

$$x_{k+1} = \mathcal{N}_F(x_k), \qquad k = 0, 1, \ldots . \tag{3.4}$$

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.

Lema 3.8. *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 3.5 . Então, para todo $x \in B(x_*, 1/(2L\beta)^{1/p})$ valem as seguintes afirmações:*

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

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

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 3.6 implica que $\|F'(x)^{-1}\| \leqslant 2\beta$ e o Lema 2.14 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. $\square$

Façamos agora a demonstração do Teorema 3.5.

Demonstração. Mostraremos por indução que a seqüência (3.1), ou equivalentemente que (3.4), está bem definida. Tome $x_0 \in B(x_*, 1/(2L\beta)^{1/p})$. Pelo Lema 3.6, temos que $F'(x_0)$ é não singular e assim x_1 está bem definido e, pelo Lema 3.8, 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 3.6, temos que $F'(x_k)$ é não singular e portanto x_{k+1} está bem definido e, pelo Lema 3.8, 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 3.8 e usando a equação (3.4), obtemos a seguinte desigualdade

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

Como $p > 0$, a desigualdade anterior implica imediatamente que $\{x_k\}$ converge para x_*.

Agora vamos demonstrar a desigualdade (3.2). Sabemos que $\{x_k\} \in B(x_*, 1/(2L\beta)^{1/p})$. Aplicando o Lema 3.8 ítem *i*) com $x = x_k$ e usando a equação (3.4), obtemos a desigualdade desejada. Assim, concluímos a demonstração do Teorema 3.5 . $\square$

A seguir enunciaremos o teorema anterior para o caso $p = 1$, veja [4].

Teorema 3.9. *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 $x_* \in U$ tais que*

i) $F(x_*) = 0$ *e* $F'(x_*)$ *é não singular ;*

ii) $B(x_*, 1/(2L\beta)) \subset U$*, onde* $\beta = \|F'(x_*)^{-1}\|$*;*

iii) $F' \in Lip_L(B(x_*, 1/(2L\beta)))$.

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

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

está bem definida, contida na bola $B(x_*, 1/(2L\beta))$ *e converge para* x_**. Além disso, a seqüência* $\{x_k\}$ *converge quadraticamente para* x_* *como se segue*

$$\|x_{k+1} - x_*\| \leqslant L\beta\|x_k - x_*\|^2, \qquad k = 0, 1, 2, \ldots.$$

Demonstração. Segue do Teorema 3.5 notando apenas que $Hol^1_L(U) = Lip_L(U)$. □

Corolário 3.10. *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 seqüência* $\{x_k\}$ *gerada pelo método de Newton para resolver* $F(x) = 0$*, com ponto inicial* $x_0 \in B(x_*, \rho)$

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

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

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

Demonstração. Como U é aberto existe $r > 0$ tal que $B[x_*, r] \subset U$ e por ser $\mathcal{B}_1$ de dimensão finita, a bola $B[x_*, r]$ é compacta. Agora, como F'' é contínua em U, existe $L > 0$ tal que

$$L = \sup_{x \in B[x_*, r]} \|F''(x)\|.$$

Pelo Desigualdade do Valor Médio, Proposição 2.12, temos que

$$\|F'(x) - F'(y)\| \leqslant L\|x - y\|, \qquad \forall\, x, y \in B[x_*, r].$$

Portanto, $F' \in Lip_L(B(x_*, r))$. Defina $\beta = \|F'(x_*)^{-1}\|$. Logo, tomando

$$\rho = \min\{r, 1/(2L\beta)\},$$

a conclusão segue diretamente do Teorema 3.9. □

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 $p = 1$. Observando o Exemplo 3.1 temos um caso em que f' é singular na solução e a taxa de convergência é menor do que a taxa obtida no Teorema 3.5.

3.3 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 3.5. Para mais detalhes sobre as demonstrações dos resultados que iremos enunciar veja [10].

Começamos com o seguinte teorema.

Teorema 3.11. *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*

i) $F(x_*) = 0$ *e* $F'(x_*)$ *é não singular ;*

ii) $B\left(x_*, ((p+1)/(\beta L))^{1/p}\right) \subset U$, *onde* $\beta = \|F'(x_*)^{-1}\|$;

iii) $F' \in Hol_L^p\left(B(x_*, ((p+1)/(\beta L))^{1/p})\right)$.

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.

Demonstração. 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

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

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, pelo Lema 2.7 e a 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

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

onde $\beta = \|f'(0)^{-1}\|$. Note que as únicas 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 3.11, com $x_* = t_2 = 0$. Também vemos facilmente que $x_* = 0$ é a única raíz 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 raíz de f. Portanto $B(x_*, ((p+1)/(\beta L))^{1/p})$ é a maior bola contendo uma única singularidade do operador F. $\square$

3.4 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:

h1) $f \in C^m(I)$, onde $m > 1$ é um número natural;

h2) 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.

Teorema 3.12. *Sejam I um intervalo aberto e $f : I \to \mathbb{R}$ uma função que satisfaz as hipóteses* **h1**, *e* **h2**. *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*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, \ldots, \tag{3.5}$$

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 3.10, 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

$$|f'(x)| \geqslant c_0 > 0 \qquad |f^{(m)}(x)| \leqslant c_1, \qquad \forall\, x \in [x_* - \bar{\rho}, x_* + \bar{\rho}]. \tag{3.6}$$

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 (3.5), 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 (3.6) 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. □

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.

Teorema 3.13. *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*

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

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

$$0 = mg'(x)f'(x) + g(x)f''(x), \qquad x \in J. \tag{3.7}$$

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 (3.7) é dada por

$$\begin{aligned} 0 = {} & mg^{(k)}(x)f'(x) + (m(k-1)+1)g^{(k-1)}(x)f''(x) + \cdots \\ & + \frac{(k-1)!}{j!(k-j)!}(j + m(k-j))g^{(k-j)}(x)f^{(j+1)}(x) + \cdots \\ & + (m+k-1)g'(x)f^{(k)}(x) + g(x)f^{(k+1)}(x), \qquad x \in J. \end{aligned} \tag{3.8}$$

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 (3.8) se reduz a

$$mg^{(m-1)}(x_*)f'(x_*) + g(x_*)f^{(m)}(x_*) = 0, \tag{3.9}$$

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

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

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

$$F^{(k)}(x) = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} f^{(j)}(x)g^{(k-j)}(x), \qquad x \in J. \tag{3.10}$$

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 (3.9) e (3.10), temos que

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

Assim, concluímos a nossa demonstração. □

Observe que a função $F(x) = f(x)/\sqrt[m]{f'(x)}$ satisfaz as hipóteses do Teorema 3.12, 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 [9]. Agora faremos um exemplo numérico para mostrar como essa aceleração funciona. Analisaremos, à luz das idéias anteriores, a função dada no Exemplo 3.2.

Exemplo 3.14. *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 3.12 e pela análise feita no Exemplo 3.2, 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*

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

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

Capítulo 4

O Teorema de Kantorovich

No capítulo anterior fizemos a análise do método de Newton para o caso local. Esta análise tem a desvantagem de exigir o conhecimento prévio de um zero do operador em consideração e hipóteses sobre o comportamento do operador nesse zero. Neste capítulo enunciamos e demonstramos o Teorema de Kantorovich sobre o Método de Newton. Este teorema não exige conhecimento prévio de zeros do operador e tem a vantagem de fazer hipóteses apenas sobre o ponto inicial. Ele garante convergência quadrática, existência e unicidade local de zeros. A demonstração é um tanto quanto longa, mas por outro lado ela é muito elegante e está organizada da seguinte forma: Primeiro vamos provar o Teorema de Kantorovich para um polinômio quadrático especifico, construído a partir da hipótese sobre o operador em consideração, mais especificamente, da constante de Lipschitz do operador e do primeiro passo de Newton, isto é, do passo de Newton a partir do ponto inicial. Este polinômio é conhecido como função majorante, a seqüência produzida pelo método de Newton aplicado a ele, a partir da origem, desempenha papel fundamental na prova do teorema no caso geral. A prova do caso geral está dividida em duas partes: na primeira parte estão provados alguns fatos básicos, sobre a inversa da derivada do operador em consideração, uma estimativa para o erro de linerização e a convergência da seqüência gerada pelo método de Newton para um zero do operador. Na segunda parte estão provados resultados sobre a taxa de convergência e unicidade local de zeros do operador. A seguir segue o enunciado do teorema de Kantorovich.

Teorema 4.1. *Sejam $\mathcal{B}_1$, $\mathcal{B}_2$ Espaços de Banach, $\Omega \subseteq \mathcal{B}_1$ um conjunto aberto e convexo, $F : \Omega \to \mathcal{B}_2$ de classe C^1 sobre Ω e contínua em $\bar{\Omega}$ o fecho de Ω. Tome $x_0 \in int(\Omega)$ tal que $F'(x_0)$ seja não singular. Suponhamos que:*

h1) $\|F'(x_0)^{-1}[F'(y) - F'(x)]\| \leqslant L\,||y - x||\,, \quad \forall \;\; x, y \in \Omega;$

h2) $\exists\, b > 0$ *tal que* $\|F'(x_0)^{-1}F(x_0)\| \leqslant b,$ *e* $\ell := Lb \leqslant 1/2;$

h3) $B(x_0, t_*) \subset \Omega$, *onde* $t_* = (1 - \sqrt{1 - 2\ell})/L$ *é a menor raíz do polinômio quadrático*

$$f(t) = \frac{L}{2}t^2 - t + b. \tag{4.1}$$

Então as sequências $\{t_k\}$ e $\{x_k\}$ geradas pelo Método de Newton para resolver as equações $f(t) = 0$ e $F(x) = 0$ com pontos iniciais $t_0 = 0$ e x_0 dadas, respectivamente, por:

$$t_{k+1} = t_k - f'(t_k)^{-1}f(t_k), \qquad x_{k+1} = x_k - F'(x_k)^{-1}F(x_k); \qquad k = 0, 1, \ldots,$$

satisfazem as seguintes condições:

i) $\{t_k\}$ *está bem definida, é monótona crescente e converge para* t_**;*

ii) $\{x_k\}$ *está bem definida, contida na bola $B(x_0, t_*)$ e converge para $x_* \in B[x_0, t_*]$ que é o único zero de F em $B[x_0, t_*]$;*

iii) $\|x_* - x_k\| \leqslant t_* - t_k, \qquad \|x_* - x_{k+1}\| \leqslant \frac{L}{2(1-Lt_k)}\|x_* - x_k\|^2; \qquad k = 0, 1, \ldots;$

iv) $\{t_k\}$ *e* $\{x_k\}$ *convergem Q-linearmente, mais precisamente,*

$$t_* - t_{k+1} \leqslant \frac{1}{2}(t_* - t_k), \qquad \|x_* - x_{k+1}\| \leqslant \frac{1}{2}\|x_* - x_k\|, \qquad k = 0, 1, \ldots.$$

Se além disso, $\ell = Lb < 1/2$. Então $\{t_k\}$ e $\{x_k\}$ convergem Q-quadraticamente, isto é,

$$t_* - t_{k+1} = \frac{1 - q^{2^k}}{1 + q^{2^k}} \frac{L}{2\sqrt{1 - 2\ell}}(t_* - t_k)^2 < \frac{L}{2\sqrt{1 - 2\ell}}(t_* - t_k)^2, \qquad k = 0, 1, \ldots,$$

$$\|x_* - x_{k+1}\| = \frac{1 - q^{2^k}}{1 + q^{2^k}} \frac{L}{2\sqrt{1 - 2\ell}}\|x_* - x_k\|^2 \leqslant \frac{L}{2\sqrt{1 - 2\ell}}\|x_* - x_k\|^2, \qquad k = 0, 1, \ldots,$$

onde $q = t_*/t_{**}$. Se temos ainda $B(x_0, t_{**}) \subset \Omega$, então x_* é o único zero de F nesta bola.

4.0.1 Prova do Teorema para a função quadrática

Nesta seção vamos provar os resultados do Teorema de Kantorovich que envolvem a seqüência $\{t_k\}$ gerada pelo método de Newton para resolver a equação $f(t) = 0$, com ponto inicial $t_0 = 0$, onde f é a função quadrática definida na equação (4.1).

Começamos com a seguinte definição:

Definição 4.2. *A iteração de Newton para o polinômio f é a função $\mathcal{N}_f : [0, t_*) \to \mathbb{R}$ definida por:*

$$\mathcal{N}_f(t) = t - \frac{f(t)}{f'(t)}. \tag{4.2}$$

Proposição 4.3. *É verdade que*

$$t < \mathcal{N}_f(t) < t_* \qquad \forall\, t \in [0, t_*). \tag{4.3}$$

Como consequência, $\mathcal{N}_f([0, t_*)) \subset [0, t_*)$. Além disso, vale a seguinte desigualdade

$$t_* - \mathcal{N}_f(t) = -\frac{L}{2f'(t)}(t_* - t)^2 \leqslant \frac{1}{2}(t_* - t), \qquad \forall\, t \in [0, t_*). \tag{4.4}$$

Se adicionalmente $\ell = Lb < 1/2$, então

$$t_* - \mathcal{N}_f(t) < -\frac{L}{2f'(t_*)}(t_* - t)^2, \qquad \forall\, t \in [0, t_*). \tag{4.5}$$

Demonstração. Seja $t \in [0, t_*)$. A primeira desigualdade em (4.3) segue simplesmente observando que $f(t) > 0$ e $f'(t) < 0$. Para demonstrar a segunda desigualdade em (4.3), observe que

$$t_* - \mathcal{N}_f(t) = t_* - t + \frac{f(t)}{f'(t)} = -\frac{1}{f'(t)}\left(f(t_*) - f(t) - f'(t)(t_* - t)\right),$$

onde usamos que t_* é uma raíz de f. Pelo Teorema Fundamental do Cálculo temos, da última equação, que

$$t_* - \mathcal{N}_f(t) = \frac{1}{f'(t)} \int_0^1 \left(f'(t + u(t_* - t)) - f'(t)\right)(t_* - t)du.$$

Usando a expressão do polinômio f dado por (4.1) na última equação e após simples cálculos, temos

$$t_* - \mathcal{N}_f(t) = -\frac{L}{2f'(t)}(t_* - t)^2. \tag{4.6}$$

Finalmente, como $t \in [0, t_*)$ e $f'(t) = Lt-1 < 0$, a última igualdade implica $t_* - \mathcal{N}_f(t) > 0$, que é equivalente à desigualdade desejada. A inclusão segue imediatamente da primeira parte.

A igualdade em (4.4) já foi provada em (4.6). Agora note que se $t \in [0, t_*)$, temos que $f'(t) = Lt - 1 < 0$ e $f'(t_*) = Lt_* - 1 \leqslant 0$. Assim, para todo $t \in [0, t_*)$ temos

$$1 = \frac{Lt-1}{f'(t)} \geqslant \frac{Lt-1}{f'(t)} + \frac{1 - Lt_*}{f'(t)} = -\frac{L(t_* - t)}{f'(t)}. \tag{4.7}$$

Combinando a última desigualdade com a igualdade em (4.4) temos desigualdade em (4.4).

Finalmente, suponhamos que $\ell = Lb < 1/2$, ou equivalentemente $f'(t_*) = Lt_* - 1 \neq 0$. Note que, $-f'(t) > -f'(t_*) > 0$, para todo $t \in [0, t_*)$, isto juntamente com a igualdade em (4.4) implica que

$$t_* - \mathcal{N}_f(t) = -\frac{L}{2f'(t)}(t_* - t)^2 < -\frac{L}{2f'(t_*)}(t_* - t)^2,$$

o que concluímos a prova. □

Consideremos agora a seqüência

$$t_0 = 0 \quad \text{e} \quad t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)} = \mathcal{N}_f(t_k), \qquad k = 0, 1, \ldots. \tag{4.8}$$

Note que a seqüência $\{t_k\}$ definida no Teorema 4.1 é a mesma definida na equação anterior. Agora, usando a consideração anterior, mostraremos que é possível obter uma forma fechada para um termo qualquer t_k da seqüência $\{t_k\}$, isto é, mostraremos que é possível escrever t_k como função do índice k e das raízes da função quadrática f. Este resultado foi baseado no artigo de Chen[3].

Proposição 4.4. *Seja* $\{t_k\}$ *a seqüência definida em* (4.8). *Se* $\ell = Lb < 1/2$, *então*

$$t_k = t_* - \frac{q^{2^k}}{1 - q^{2^k}} \frac{2\sqrt{1 - 2\ell}}{L}, \qquad k = 0, 1, \ldots,$$

onde $q = t_*/t_{**}$ é o quociente entre as raízes do polinômio quadrático f. Como conseqüência obtemos que

$$f'(t_k) = -\frac{1+q^{2^k}}{1-q^{2^k}}\sqrt{1-2\ell}, \qquad k = 0, 1, \ldots.$$

Demonstração. Seja $t \in [0, t_*)$. Como $\ell = Lb < 1/2$ temos que o polinômio f possui duas raízes reais e distintas t_* e t_{**}. Assim, usando a expressão de f fatorado nas suas raízes, temos

$$f(t) = \frac{L}{2}(t - t_*)(t - t_{**}), \qquad f'(t) = \frac{L}{2}\left((t - t_*) + (t - t_{**})\right).$$

Usando a expressão de $\mathcal{N}_f$ e as duas igualdades anteriores obtemos que

$$t_* - \mathcal{N}_f(t) = t_* - t + \frac{f(t)}{f'(t)} = t_* - t + \frac{L/2(t-t_*)(t-t_{**})}{L/2((t-t_*)+(t-t_{**}))}.$$

Após simples manipulações algébricas na última equação obtemos a seguinte igualdade

$$t_* - \mathcal{N}_f(t) = -\frac{(t-t_*)^2}{(t-t_*)+(t-t_{**})}.$$

Analogamente podemos mostrar que a seguinte igualdade também vale

$$t_{**} - \mathcal{N}_f(t) = -\frac{(t-t_{**})^2}{(t-t_*)+(t-t_{**})}.$$

Portanto, fazendo o quociente entre as duas últimas equações, obtemos a igualdade

$$\frac{t_* - \mathcal{N}_f(t)}{t_{**} - \mathcal{N}_f(t)} = \left(\frac{t-t_*}{t-t_{**}}\right)^2.$$

Como $\{t_k\} \subset [0, t_*)$ e $t_* < t_{**}$, tomando $t = t_{k-1}$ na última igualdade obtemos que

$$\begin{aligned}\frac{t_* - t_k}{t_{**} - t_k} &= \left(\frac{t_{k-1}-t_*}{t_{k-1}-t_{**}}\right)^2 = \left(\frac{t_{k-2}-t_*}{t_{k-2}-t_{**}}\right)^{2^2} = \left(\frac{t_{k-3}-t_*}{t_{k-3}-t_{**}}\right)^{2^3} = \cdots \\ &= \left(\frac{t_0-t_*}{t_0-t_{**}}\right)^{2^k} = \left(\frac{t_*}{t_{**}}\right)^{2^k} = q^{2^k},\end{aligned}$$

pois, $t_0 = 0$. Isolando t_k, na última equação, obtemos a seguinte igualdade

$$t_k = \frac{1}{1-q^{2^k}}\left(t_* - q^{2^k}t_{**}\right).$$

Agora, novamente após simples manipulações na igualdade acima obtemos a igualdade

$$t_k = t_* - t_* + t_k = t_* - \frac{1}{1-q^{2^k}}\left(t_*(1-q^{2^k}) - t_* + q^{2^k}t_{**}\right) = t_* - \frac{q^{2^k}}{1-q^{2^k}}\left(t_{**} - t_*\right).$$

Portanto, observando que $t_{**} - t_* = 2\sqrt{1-2\ell}/L$, obtém-se a igualdade desejada na primeira parte.

Para provar a igualdade da segunda parte, primeiro note que $f'(t) = Lt - 1$. Portanto, substituindo o valor de t_k, dado pela primeira parte, na expressão de f', obtemos

$$f'(t_k) = Lt_k - 1 = L\left(t_* - \frac{q^{2^k}}{1-q^{2^k}}\frac{2\sqrt{1-2\ell}}{L}\right) - 1 = Lt_* - \frac{q^{2^k}}{1-q^{2^k}}2\sqrt{1-2\ell} - 1.$$

Agora, substituindo a expressão de t_* na equação anterior e realizando alguns simples cálculos, obtém-se a igualdade desejada. $\square$

Como a sequência $\{t_k\}$ definida no Teorema 4.1 pode ser reescrita na forma

$$t_0 = 0 \quad \text{e} \quad t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)} = \mathcal{N}_f(t_k), \qquad k = 0, 1, \ldots,$$

então, baseado na Proposição 4.3 e Proposição 4.4, podemos concluir que :

a) $\{t_k\}$ é bem definida, é monótona crescente e está contida no intervalo $[0, t_*)$;

b) a seqüência $\{t_k\}$ converge Q-linearmente para t_*, e vale a seguinte desigualdade

$$t_* - t_{k+1} \leqslant \frac{1}{2}(t_* - t_k), \qquad k = 0, 1, \ldots;$$

c) Se $\ell = Lb < 1/2$, isto é $f'(t_*) \neq 0$, ou equivalentemente se a função f possui duas raízes distintas t_* e t_{**}, então pela equação (4.5), a seqüência $\{t_k\}$ converge quadraticamente para t_*, e vale a seguinte desigualdade

$$t_* - t_{k+1} = \frac{1-q^{2^k}}{1+q^{2^k}}\frac{L}{2\sqrt{1-2\ell}}(t_* - t_k)^2 < \frac{L}{2\sqrt{1-2\ell}}(t_* - t_k)^2, \qquad k = 0, 1, \ldots,$$

e portanto $\{t_k\}$ converge Q-quadraticamente para t_*.

Portanto, todos os resultados do Teorema 4.1 referentes a sequência $\{t_k\}$ estão provados.

4.1 Convergência

Nesta seção demonstramos que a seqüência $\{x_k\}$ gerada pelo Método de Newton para resolver $F(x) = 0$ com ponto inicial x_0, como no Teorema 4.1, está bem definida e converge para um ponto $x_* \in B[x_0, t_*]$. Começamos com a seguinte proposição.

Proposição 4.5. *Se* $\|x - x_0\| \leqslant t < t_*$ *, então* $F'(x)$ *é não singular e vale a estimativa*

$$\|F'(x)^{-1}F'(x_0)\| \leqslant -\frac{1}{f'(t)}.$$

Em particular, $F'(x)$ *é não singular para todo* $x \in B(x_0, t_*)$.

Demonstração. Tome $x \in B[x_0, t]$, onde $0 \leqslant t < t_*$. Note que

$$\begin{aligned}\|F'(x_0)^{-1}F'(x) - I\| &= \|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| = L\|x - x_0\| \\ &\leqslant Lt < Lt_* = 1 - \sqrt{1 - 2\ell} < 1\end{aligned}$$

Portanto, pelo Teorema 2.7 temos que $F'(x_0)^{-1}F'(x)$ é não singular e portanto $F'(x)$ é não singular para todo $x \in B[x_0, t]$. Pelo Teorema 2.7 e a primeira desigualdade acima temos

$$\|F'(x)^{-1}F'(x_0)\| \leqslant \frac{1}{1 - \|F'(x_0)^{-1}F'(x) - I\|} \leqslant \frac{1}{1 - Lt} = -\frac{1}{f'(t)}.$$

Como queríamos demonstrar. □

Vimos que a idéia do Método de Newton consiste na linearização de uma função derivável. Entendemos por linearização de uma função a expansão de Taylor de primeira ordem dessa função. Esta expansão gera um erro. Queremos estudar o erro cometido na linearização da funções F e f e estabelecer uma relação entre eles. Por definição o erro de linearização de F é igual a

$$E(x, y) := F(y) - F(x) - F'(x)(y - x), \qquad x, \in \Omega, \quad y \in \bar{\Omega}.$$

e o erro de linearização do polinômio quadrático f é definido por

$$e(t, v) := f(v) - f(t) - f'(t)(v - t), \qquad t, v \in \mathbb{R}.$$

No lema a seguir relacionamos os erros de linerização de F e f.

Lema 4.6. *Tome $x, y \in \Omega$ e $t \in (0, t_*)$. Se $\|x - x_0\| \leqslant t$, temos*

$$\|F'(x_0)^{-1}E(x,y)\| \leqslant e(\|x - x_0\|, \|y - x\| + \|x - x_0\|) = \frac{L}{2}\|y - x\|^2. \tag{4.9}$$

Demonstração. Sejam $x, y \in \Omega$. Temos $x + u(y - x) \in \Omega$ se $u \in [0, 1]$, pois Ω é convexo. Portanto, usando o Teorema Fundamental do Cálculo e a hipótese **h1** obtemos

$$\begin{aligned}
\|F'(x_0)^{-1}E(x,y)\| &= \|F'(x_0)^{-1}\left(F(y) - [F(x) + F'(x)(y - x)]\right)\| \\
&\leqslant \int_0^1 \|F'(x_0)^{-1}\left[F'(x + u(y - x)) - F'(x)\right]\| \, \|y - x\| \, du \\
&\leqslant \int_0^1 Lu\|y - x\|^2 \, du \ = \frac{L}{2}\|y - x\|^2.
\end{aligned}$$

Como f é um polinômio de segundo grau e $f'' \equiv L$, temos

$$f(v) = f(u) + f'(u)(v - u) + \frac{L}{2}(v - u)^2,$$

e portanto, tomando $v = \|y - x\| + \|x - x_0\|$ e $u = \|x - x_0\|$ na equação anterior, e usando a definição de $e(u, v)$, temos a igualdade em (4.9). $\square$

Nosso trabalho agora será mostrar que a seqüência $\{x_k\}$ gerada pelo Método de Newton para encontrar um zero do operador F está bem definida. Lembre-se que temos a garantia de que F' é não singular na bola $B(x_0, t_*)$. Para simplificar a notação definamos a iteração de Newton $\mathcal{N}_F : B(x_0, t_*) \to \mathcal{B}_1$ como

$$\mathcal{N}_F(x) = x - F'(x)^{-1}F(x). \tag{4.10}$$

Temos que $\mathcal{N}_F$ está bem definida na bola $B(x_0, t_*)$, mas não temos nenhuma garantia de que $\mathcal{N}_F(x) \in B(x_0, t_*)$ dado que $x \in B(x_0, t_*)$, sequer temos a garantia de que $\mathcal{N}_F(x)$ está no domínio de F. Mostrar que a bola aberta $B(x_0, t_*)$ é invariante por $\mathcal{N}_F$, isto é, $\mathcal{N}_F(B(x_0, t_*)) \subset B(x_0, t_*)$ é fundamental para nosso propósito, na verdade é suficiente mostrar que um determinado subconjunto de $B(x_0, t_*)$ é invariante por $\mathcal{N}_F$. Para isso considere os conjuntos definidos por

$$K(t) := \left\{x \in B[x_0, t]\,;\ \|F'(x)^{-1}F(x)\| \leqslant -\frac{f(t)}{f'(t)}\right\};\quad t \in [0, t_*)$$

$$K := \bigcup_{t\in[0,t_*)} K(t).$$

Se $t \in [0, t_*)$ então $f'(t) \neq 0$ e $F'(x)$ é não singular em $B[x_0, t] \subset B(x_0, t_*)$. Logo a definição dos conjuntos acima é consistente.

Lema 4.7. *Para cada $t \in [0, t_*)$, $K(t) \subset B(x_0, t_*)$ e*

$$\mathcal{N}_F(K(t)) \subset K(\mathcal{N}_f(t)). \tag{4.11}$$

Como conseqüência, $K \subset B(x_0, t_*)$ e K é invariante por $\mathcal{N}_F$, isto é, $\mathcal{N}_F(K) \subset K$.

Demonstração. Note que $K(t) \subset K \subset B(x_0, t_*)$ para todo $t \in [0, t_*)$. Tome $t \in [0, t_*)$ e $x \in K(t)$. Pela definição de $\mathcal{N}_F$ e $\mathcal{N}_f$ e Proposição 4.3, temos que

$$\begin{aligned}\|\mathcal{N}_F(x) - x_0\| &\leqslant \|\mathcal{N}_F(x) - x\| + \|x - x_0\| \\ &\leqslant \|F'(x)^{-1}F(x)\| + t \leqslant -\frac{f(t)}{f'(t)} + t = \mathcal{N}_f(t) < t_*.\end{aligned}$$

Logo, $\mathcal{N}_F(x) \in B[x_0, \mathcal{N}_f(t)] \subset B(x_0, t_*)$ para todo $x \in K(t)$. Assim, temos que $\mathcal{N}_F(x)$ e $\mathcal{N}_f(t)$ estão nos domínios de F e f respectivamente. Agora, como $x \in K(t)$ e $t \in [0, t_*)$, temos

$$\|\mathcal{N}_F(x) - x\| = \|F'(x)^{-1}F(x)\| \leqslant -\frac{f(t)}{f'(t)} = \mathcal{N}_f(t) - t.$$

Aplicando o Lema 4.6, com $y = \mathcal{N}_F(x)$ e observando a inequação anterior, temos que

$$\|F'(x_0)^{-1}E(x, \mathcal{N}_F(x))\| \leqslant \frac{L}{2}\|\mathcal{N}_F(x) - x\|^2 \leqslant \frac{L}{2}\|\mathcal{N}_f(t) - t\|^2 = e(t, \mathcal{N}_f(t)).$$

Observe que, usando as expressões dos erros de linearização, $\mathcal{N}_F$ e $\mathcal{N}_f$, é fácil ver que $E(x, \mathcal{N}_F(x)) = F(\mathcal{N}_F(x))$ e $e(t, \mathcal{N}_f(t)) = f(\mathcal{N}_f(t))$. Assim, a última inequação se reduz a

$$\|F'(x_0)^{-1}F(\mathcal{N}_F(x))\| \leqslant f(\mathcal{N}_f(t)).$$

Por outro lado $\|\mathcal{N}_F(x) - x_0\| \leqslant \mathcal{N}_f(t) < t_*$. Logo pela Proposição 4.5 temos que $F'(\mathcal{N}_F(x))$ é não singular e vale a estimativa

$$\|F'(\mathcal{N}_F(x))^{-1}F'(x_0)\| \leqslant -\frac{1}{f'(t)} \leqslant -\frac{1}{f'(\mathcal{N}_f(t))},$$

pois, f' é crescente e pela Proposição 4.3 temos $t < \mathcal{N}_f(t)$. Agora, combinando as duas últimas inequações, podemos concluir que

$$\begin{aligned}\|F'(\mathcal{N}_F(x))^{-1}F(\mathcal{N}_F(x))\| &\leqslant \|F'(\mathcal{N}_F(x))^{-1}F'(x_0)\|\,\|F'(x_0)^{-1}F(\mathcal{N}_F(x))\| \\ &\leqslant -\frac{f(\mathcal{N}_f(t))}{f'(\mathcal{N}_f(t))},\end{aligned}$$

o que implica que $\mathcal{N}_F(x) \in K(\mathcal{N}_f(t))$ para todo $x \in K(t)$, isto é, $\mathcal{N}_F(K(t)) \subset K(\mathcal{N}_f(t))$.

Agora, tome $x \in K$. Pela definição de K temos que existe $t \in [0, t_*)$ tal que $x \in K(t)$. Pelo resultado anterior temos $\mathcal{N}_F(x) \in K(\mathcal{N}_f(t))$, e pelo fato de que $\mathcal{N}_f(t) \in [0, t_*)$ para todo $t \in [0, t_*)$, podemos concluir que $\mathcal{N}_F(K) \subset K$. Assim concluímos a prova do lema. □

Veja que a seqüência $\{x_k\}$ gerada pelo método de Newton para resolver a equação $F(x) = 0$ definida por $x_{k+1} = x_k - F'(x_k)^{-1}F(x_k)$ satisfaz

$$x_{k+1} = \mathcal{N}_F(x_k) \;\; ; \; k = 0, 1, 2, \dots \; . \tag{4.12}$$

Portanto temos a seguinte proposição.

Corolário 4.8. *A seqüência $\{x_k\}$ satisfaz a relação $x_k \in K(t_k)$, para todo $k = 0, 1, \dots$. Como consequência*

$$\|x_{k+1} - x_k\| \leqslant t_{k+1} - t_k, \qquad k = 0, 1, \dots \; .$$

Demonstração. Primeiro, mostraremos por indução, que $x_k \in K(t_k)$, para $k = 0, 1, \dots$. Por hipótese, no Teorema 4.1 temos que

$$\|F'(x_0)^{-1}F(x_0)\| \leqslant b = -\frac{f(0)}{f'(0)}.$$

Portanto $x_0 \in K(0) \subset K$. Por indução, suponha que $x_k \in K(t_k)$. Pelo resultado anterior temos que $\mathcal{N}_F(x_k) \in K(\mathcal{N}_f(t_k))$, ou seja $x_{k+1} \in K(t_{k+1})$, isto conclui a indução.

Como $x_k \in K(t_k)$ segue, da definição de $K(t_k)$, que

$$\|x_{k+1} - x_k\| = \|F'(x_k)^{-1}F(x_k)\| \leqslant -\frac{f(t_k)}{f'(t_k)} = t_{k+1} - t_k, \qquad k = 0, 1, \dots \; .$$

□

Proposição 4.9. *A seqüência* $\{x_k\}$ *está bem definida, contida em* $B(x_0, t_*)$*, converge para um ponto* $x_* \in B[x_0, t_*]$ *tal que* $F(x_*) = 0$ *e vale a desigualdade*

$$\|x_* - x_k\| \leqslant t_* - t_k, \qquad k = 0, 1, \ldots . \tag{4.13}$$

Demonstração. Pelo Corolário 4.8 temos que $x_k \in K(t_k)$, para $k = 0, 1, \ldots$. Como $t_k < t_*$, por definição concluímos que $K(t_k) \subset K \subset B(x_0, t_*)$. Mas o Lemma 4.7 implica que $\mathcal{N}_F(K) \subset K$. Isto juntamente com (4.12) implica que $\{x_k\}$ está bem definida e contida em $K \subset B(x_0, t_*)$.

Sabemos que $\{t_k\}$ é uma seqüência de números reais monótona crescente que converge para t_*, e que a seqüência $\{x_k\} \subset B(x_0, t_*) \subset \mathcal{B}_1$, onde $\mathcal{B}_1$ é um espaço de Banach. Portanto, pelo Lema 2.5 e a desigualdade do Corolário 4.8, concluímos que a seqüência $\{x_k\}$ converge para um ponto $x_* \in B[x_0, t_*]$ e vale a desigualdade (4.13).

Resta provar que $F(x_*) = 0$. Primeiro note que, usando a desigualdade triangular

$$\|F'(x_k)\| \leqslant \|F'(x_0)\| + \|F'(x_k) - F'(x_0)\|.$$

Agora, após simples manipulações algébricas na última equação, temos

$$\|F'(x_k)\| \leqslant \|F'(x_0)\| + \|F'(x_0)\| \, \|F'(x_0)^{-1}[F'(x_k) - F'(x_0)]\|.$$

Como $\|x_k - x_0\| \leqslant t_k < t_*$ usando a hipótese **h1** do Teorema 4.1, temos

$$\|F'(x_0)^{-1}[F'(x_k) - F'(x_0)]\| \leqslant L\|x_k - x_0\| < Lt_k < Lt_*.$$

Portanto, combinando as duas últimas equações, temos que a seqüência $\{\|F'(x_k)\|\}$ é limitada. Por outro lado observe que

$$\|F(x_k)\| \leqslant \|F'(x_k)\| \, \|F'(x_k)^{-1}F(x_k)\| \leqslant \|F'(x_k)\| \, (t_{k+1} - t_k).$$

Logo, como a seqüência $\{\|F'(x_k)\|\}$ é limitada e $\{t_k\}$ é convergente, pela última desigualdade temos que

$$\lim_{k\to\infty} F(x_k) = 0.$$

Finalmente, como F é contínua em $B[x_0, t_*]$, e a seqüência $\{x_k\} \subset B(x_0, t_*) \subset \Omega$ converge para x_*, concluímos que

$$0 = \lim_{k\to\infty} F(x_k) = F(\lim_{k\to\infty} x_k) = F(x_k).$$

Como queríamos demonstrar. $\square$

4.2 Unicidade e Taxa de Convergência

O que fizemos até agora sobre a seqüência $\{x_k\}$ foi mostrar que ela está bem definida, contida na bola $B(x_0, t_*)$ e converge para um ponto $x_* \in B[x_0, t_*]$, que satisfaz $F(x_*) = 0$, isto é, x_* é solução da equação $F(x) = 0$. Nesta seção mostramos que a seqüência $\{x_k\}$ tem convergência pelo menos Q-linear para x_* e vale a unicidade de solução na bola $B(x_0, t_*)$. Também mostramos que quando supomos $\ell = Lb < 1/2$ a convergência de $\{x_k\}$ para x_* é pelo menos Q-quadrática e que vale a unicidade de solução na bola $B(x_0, t_{**})$ que é maior que a bola anterior. Começamos com o seguinte lema.

Lema 4.10. *Sejam $x, y \in B(x_0, t_{**})$ e $0 \leqslant t < t_*$. Se $B(x_0, t_{**}) \subset \Omega$ e além disso,*

$$x \in B(x_0, t), \qquad F(y) = 0.$$

Então é válida a seguinte desigualdade

$$\|y - \mathcal{N}_F(x)\| \leqslant -\frac{L}{2f'(t)}\|y - x\|^2. \tag{4.14}$$

Demonstração. Note que usando a definição de $\mathcal{N}_F$ e que $F(y) = 0$, temos que

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

Agora, colocando em evidência $-F'(x)^{-1}$ e usando a expressão do erro $E(x, y)$, temos

$$y - \mathcal{N}_F(x) = -F'(x)^{-1}E(x, y) = \left[-F'(x)^{-1}F'(x_0)\right]\left[F'(x_0)^{-1}E(x, y)\right].$$

Tomando a norma em ambos os lados da última igualdade, usando a Proposição 4.5 e o Lema 4.6, temos

$$\|y - \mathcal{N}_F(x)\| \leqslant \| - F'(x)^{-1}F'(x_0)\| \, \|F'(x_0)^{-1}E(x,y)\| \leqslant -\frac{1}{f'(t)}\,\frac{L}{2}\|y-x\|^2,$$

que é a desigualdade desejada. □

Corolário 4.11. *As seqüências $\{x_k\}$ e $\{t_k\}$ satisfazem a desigualdade*

$$\|x_* - x_{k+1}\| \leqslant -\frac{L}{2f'(t_k)}\|x_* - x_k\|^2\,, \quad k = 0, 1, \ldots .$$

Em particular, temos a convergência Q-linear de $\{x_k\}$ como se segue

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

Se $\ell = Lb < 1/2$, isto é, $f'(t_*) \neq 0$, a convergência de $\{x_k\}$ é Q-quadrática, mais precisamente temos que

$$\|x_* - x_{k+1}\| \leq \frac{1-q^{2^k}}{1+q^{2^k}}\frac{L}{2\sqrt{1-2\ell}}\|x_* - x_k\|^2 \leqslant \frac{L}{2\sqrt{1-2\ell}}\|x_* - x_k\|^2, \qquad k = 0, 1, \ldots . \tag{4.15}$$

Demonstração. Vamos provar a primeira desigualdade. Pelo Corolário 4.8 temos que $\|x_k - x_0\| \leqslant t_k$ para todo $k = 0, 1, \ldots$. Assim, aplicando o Lema 4.10 com $y = x_*$, $x = x_k$ e $t = t_k$ obtemos diretamente que

$$\|x_* - \mathcal{N}_F(x_k)\| \leqslant -\frac{L}{2f'(t_k)}\|x_* - x_k\|^2,$$

Usando (4.12), a inequação acima se reduz à desigualdade desejada.

Agora vamos provar a segunda desigualdade. Usando a Proposição 4.9 temos que $\|x_* - x_k\| \leqslant t_* - t_k$. Isto juntamente com a primeira desigualdade do corolário implica que

$$\|x_* - x_{k+1}\| \leqslant -\frac{L}{2f'(t_k)}(t_* - t_k)\|x_* - x_k\|.$$

Por outro lado, como $t_k < t_*$ temos $f'(t_k) = Lt_k - 1 < 0$ e $-f'(t_*) = 1 - Lt_* \geq 0$. Assim,

$$1 = \frac{Lt_k - 1}{f'(t_k)} \geqslant \frac{Lt_k - 1}{f'(_k)} + \frac{1 - Lt_*}{f'(t_k)} = -\frac{L(t_* - t_k)}{f'(t_k)}.$$

Combinando as duas últimas inequações obtemos a segunda desigualdade.

Finalmente vamos provar (4.15). Assuma que $\ell = Lb < 1/2$. Usando a primeira desigualdade do corolário e a segunda igualdade da Proposição 4.4 obtemos

$$\|x_* - x_{k+1}\| \leqslant -\frac{L}{2f'(t_k)}\|x_* - x_k\|^2 = \frac{1 - q^{2^k}}{1 + q^{2^k}} \frac{L}{2\sqrt{1 - 2\ell}}\|x_* - x_k\|^2, \qquad k = 0, 1, \ldots .$$

Assim, da última desigualdade e do fato de que $(1 - q^{2^k})/(1 + q^{2^k}) < 1$, temos que a inequação (4.15) é verdadeira, para todo $k \in \mathbb{N}$. Assim concluímos a demonstração. □

Corolário 4.12. *O limite x_* da seqüencia $\{x_k\}$ é o único zero do operador F em $B[x_0, t_*]$. Se vale a hipótese $\ell = Lb < 1/2$ e $B(x_0, t_{**}) \subset \Omega$, então x_* é o único zero do operador F na bola $B(x_0, t_{**})$.*

Demonstração. Suponha, por contradição, que exista $y_* \in B[x_0, t_*]$ tal que $F(y_*) = 0$. Mostraremos, por indução, que

$$\|y_* - x_k\| \leqslant t_* - t_k \qquad k = 0, 1, \ldots . \tag{4.16}$$

Primeiro note que, pela hipótese de contradição, $y_* \in B[x_0, t_*]$, portanto temos

$$\|y_* - x_0\| \leqslant t_* = t_* - t_0,$$

pois $t_0 = 0$. Suponha agora que a inequação (4.16) seja válida para k. Pelo Corolário 4.8 temos que $\|x_k - x_0\| \leqslant t_k$ para todo $k = 0, 1, \ldots$. Aplicando o Lema 4.10 para $x = x_k$, $y = y_*$, $t = t_k$ obtemos a seguinte desigualdade

$$\|y_* - \mathcal{N}_F(x_k)\| \leqslant -\frac{L}{2f'(t_k)}\|y_* - x_k\|^2.$$

E como a inequação (4.16) é válida para k, usando a desigualdade acima e (4.12), obtemos

$$\|y_* - x_{k+1}\| \leqslant -\frac{L}{2f'(t_k)}(t_* - t_k)^2.$$

Como $t_k < t_*$, usamos a equação (4.4) da Proposição 4.3 e a última inequação para obter

$$\|y_* - x_{k+1}\| \leqslant -\frac{L}{2f'(t_k)}(t_* - t_k)^2 = t_* - \mathcal{N}_f(t_k) = t_* - t_{k+1},$$

e a indução está concluída. Para mostrar a unicidade de solução da equação $F(x) = 0$ na bola $B[x_0, t_*]$, note que as seqüências $\{x_k\}$ e $\{t_k\}$ convergem respectivamente para x_* e t_*, assim tomando limite em (4.16), obtemos

$$\|y_* - x_*\| = 0,$$

e portanto $y_* = x_*$. Logo, a unicidade de solução da equação $F(x) = 0$ em $B[x_0, t_*]$ está provada.

Agora, suponha que $\ell = Lb < 1/2$, isto é, $f'(t_*) = Lt_* - 1 < 0$. Faremos a prova da unicidade na bola $B(x_0, t_{**})$ por contradição. Portanto suponha que exista y_* tal que

$$t_* < \|y_* - x_0\| < t_{**} \qquad \text{e} \qquad F(y_*) = 0.$$

Note que, usando o Lema 4.6 com $x = x_0$, $y = y_*$, $t = t_*$ e $v = \|y_* - x_0\|$, nós obtemos que

$$\|F'(x_0)^{-1}E(x_0, y_*)\| = \frac{L}{2}\|y_* - x_0\|^2. \tag{4.17}$$

Agora, usando a expressão do $E(x_0, y_*)$ e a hipótese **h2** do Teorema 4.1, temos

$$\begin{aligned} \|F'(x_0)^{-1}E(x_0, y_*)\| &= \|y_* - x_0 + F'(x_0)^{-1}F(x_0)\| \\ &\geqslant \|y_* - x_0\| - \|F'(x_0)^{-1}F(x_0)\| \\ &\geqslant \|y_* - x_0\| - b. \end{aligned}$$

Finalmente, combinando a desigualdade anterior e a equação (4.17), obtemos

$$\frac{L}{2}\|y_* - x_0\|^2 \geqslant \|y_* - x_0\| - b\,,$$

ou equivalentemente

$$f(\|y_* - x_0\|) = \frac{L}{2}\|y_* - x_0\|^2 - \|y_* - x_0\| + b \geqslant 0,$$

isto é um absurdo, pois $f(\tau) < 0$ para todo $t_* < \tau < t_{**}$ e $t_* < \|y_* - x_0\| < t_{**}$. Portanto, assumindo a hipótese $\ell = Lb < 1/2$, temos a unicidade de solução da equação $F(x) = 0$ na bola $B(x_0, t_{**})$. $\square$

Combinado os resultados obtidos na Proposição 4.9, Corolário 4.11 e Corolário 4.12, demonstramos todos os resultados do teorema de Kantorovich para a seqüência $\{x_k\}$.

4.3 Considerações Finais

Para finalizar queremos destacar que existem várias versões do teorema de Kantorovich. Na versão que estudamos, fizemos a hipótese de que o gradiente do operador $F'(x_0)^{-1}F(x)$ é Lipschitz. O operador em questão, tem os mesmos zeros do operador F. Observamos que o enunciado do teorema para o operador $F'(x_0)^{-1}F(x)$ é mais simples. Resultando que o polinômio quadrático (função majorante) depende de menos constantes. No livro de Dennis e Schnabel[4] podemos encontrar o enunciado do teorema de Kantorovich para o operador F (Caso clássico). A demonstração do teorema de Kantorovich para o caso clássico, é feita por indução. Poderíamos ter procedido assim, mas preferimos mostrar o teorema via a existência de um conjunto invariante.

Na análise de convergência local do método de Newton usamos, em certo sentido, uma técnica semelhante à técnica usada para o caso semi-local, definindo um conjunto invariante pela iteração de Newton. Um teorema similar para o caso local, que está em Dennis e Schnabel [4], faz a prova por indução sobre o índice da seqüência de Newton.

No enunciado do Teorema 3.5 supomos que o gradiente do operador F fosse Hölder contínuo. Isso nos motiva buscar um resultado sobre o teorema de Kantorovich, enunciado como no Teorema 4.1, supondo que o gradiente do operador $F'(x_0)^{-1}F(x)$ é Hölder contínuo ao invés de supor que ele seja Lipschtiz.

Referências Bibliográficas

[1] Appel,J., De Pascale, E., Lysenko, J.V., and Zabrejko, P.P., *New results on Newton-Kantorovich approximations with applications to nonlinear integral equations*, Numer. funct. Anal. and Optimiz. , **18**(1 e 2), (1997), 1–17.

[2] Bennet, A.A., *Newton's Method in General Analysis*, Proc. Nat. Acad. Sci. USA **2**(10), (1916), pp 592–598.0,

[3] Chen, D. *Ostrowski-Kantorovich Theorem and S-order of Convergence of Halley Method in Banach Spaces*, Comment. Math. Univ. Carolin., **34**(1), (1993), pp 153–163.

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

[5] Ostrowski, A. M., *Solution of Equations and Sistems of Equations*, Academic Press, New York (1996) .

[6] Ferreira, O.P. and Svaiter, B. F., *Kantorovich's Theorem on Newton's method in Riemannian Manifolds* Journal of Complexity, **18**, (2002), 304–329.

[7] Ferreira, O.P. and Svaiter, B. F., *Kantorovich's Majorants Principle for Newton's Method*, to appear in Optimization Methods and Software, (2006).

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

[9] Gerlach, J. *Accelerated Convergence in Newton's Method* , SIAM Review, **36**(2), (1994), pp 272–276.

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

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

[12] Kantorovich, L.V., and Akilov, G.P., *Functional analysis in normed spaces*, Oxford, Pergamon (1964).

[13] Krantz, S. G., Parks ,H. R., *The implicit function theorem : history, theory, and applications*, Boston, Birkhäuser, (2002).

[14] Mathews, J. H., *Bibliography for Newton's Method*, math.fullerton.edu/mathews/newtonsmethod/Newton'sMethodBib/Links/Newton'sMethodBib_lnk_3.html.

[15] Moser, J., *A new techniques for the construction of solutions of nonlinear differential equations*, Proc. Nat. Acad. sci. USA, 47 (1961), 1824-1831.

[16] Nash, J., *The embedding problem for Riemannian manifolds*, Ann. of Math. 63 (1956), 20–63.

[17] Ortega, J. M. and Rheimboldt, W. C., *Iteractive Soluction of Nonlinear Equations in Several Variables* , New York-London, Academic Press, (1970).

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

[19] Potra, Florian A., *The Kantorovich Theorem and Interior Point Methods*, Mathematical Programming, **102** (1), (2005), pp 47–70.

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

[21] Wang, X., *Convergence of Newton's Method and Inverse Function Theorem in Banach Space*, Math. Comp. **68**(225), (1999), pp 169–186.