UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

Análise Local e Semi-Local do Método de Newton do Ponto de Vista do Princípio Majorante

por

Luis Enrique Zelaya De los Santos

Orientador: Dr. Orizon Pereira Ferreira

Dissertação de Mestrado em Matemática
Goiânia - Goiás
2007

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

Análise local e semi-local do método de Newton do ponto de vista do princípio majorante

por

Luis Enrique Zelaya De los Santos

Área de Concentração : **Matemática Aplicada**
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
2007

Dados Internacionais de Catalogação-na-Publicação (CIP)
(GPT/BC/UFG)

Zelaya De los Santos, Luis Enrique.

Z49a local e semi-local do método de Newton do ponto de vista do princípio majorante / Luis Enrique Zelaya De los Santos. - 2007.

64f.

Orientador: Prof. Orizon Pereira Ferreira.

Dissertação (Mestrado) - Universidade Federal de Goiás. Instituto de Matemática e Estatística, 2007.

Bibliografia: f. 62-64.

1. Matemática aplicada I. Ferreira, Orizon Pereira II. Universidade Federal de Goiás. **Instituto de Matemática e Estatística**. III. Título.

CDU: 519.6

Agradecimentos

À Deus, pela proteção durante este percurso da minha carreira estudantil, por ter me dado a oportunidade e a capacidade, pois sem Deus nada disso seria possível. À ele toda honra e toda glória.

Ao Prof. Dr. Orizon Pereira Ferreira, pela amizade, paciência e dedicação que foram indispensáveis para a concretização deste trabalho.

À minha esposa Magdalena, pela compreensão e a grande resposabilidade que teve durante estes dois anos.

À meus filhos: Lady, Yohann e Luis, pelo amor demonstrado durante esses anos de aussencia.

À Armando, por tuda ajuda que me ofereceu e a confiança que ele teve na minha pessoa.

À familiares e amigos, que apoiaram-me em mais um degrau de minha vida. Em especial, aos colegas de mestrado, que ajudaram-me nos momentos de dificuldades, não vou esquecer nenhum de vocês.

À (CNPq), pela Bolsa de Estudos Concedida, sem a qual seria difícil a realização desta dissertação[1].

[1] "O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq - Brasil".

Sumário

Resumo

A convergência local e semi-local do método de Newton baseada no princípio majorante de Kantorovich é apresentada nesta dissertação. Esta análise torna claro a relação entre o operador não linear em consideração e a função escalar majorante, a qual relaxa a continuidade Lipschitz. Ela também permite obter a maior domínio de unicidade de solução e unificação de alguns resultados previamente sem relação. Além disso, para o caso local, obtemos o maior raio de convergência possível.

Abstract

A local and semi-local convergence analysis for Newton's method based on Kantorovich's majorant principle is presented in this dissertation. This analysis makes clear the relationship between the non-linear operator under consideration and the scalar majorant function, which relaxes the Lipschitz continuity. It also allows the achievement of the biggest domain of the uniqueness of solution and the unification of some results previously unrelated. Moreover, for the local case, the optimal ball of convergence is obtained.

Capítulo 1

Introdução

O método de Newton e suas variações são os mais eficientes métodos conhecidos para resolver sistemas não lineares, incluindo a procura de um minimizador local de uma função e muitas outras aplicações. Uma das mais importantes aplicações do método de Newton é que ele permite desenvolver algoritmos de tempo polinomial em programação convexa, ver Nesterov e Nemiroskii [17]. Além de suas aplicações práticas, o método de Newton é também uma poderosa ferramenta teórica tendo um amplo domínio de aplicações em matemática pura, ver Blum, Cucker, Shub e Smale [2], Krantz e Parks [11], Moser [15], Nash [16]. Para uma atual perspectiva histórica do método de Newton e suas aplicações em otimização, ver Polyak [20].

A análise de convergência clássica (convergência local) do método de Newton requer que o ponto inicial esteja "suficientemente perto" da solução. Um problema com esta análise é que a solucão deve ser conhecida ou dada *a priori*. Mas, isto tem vantagem de dar o raio ótimo da bola de convergência, veja Rall [21], Traub e Wozniakowski [24] e Wang [27]. Por outro lado, o teorema de Kantorovich sobre o método de Newton garante convergência a solução usando condições semi-locais. Esta análise não requer *a priori* a existência de uma solução, em vez disto garante a existência dela e sua unicidade em alguma região, veja Kantorovich e Akilov [10].

Entre as hipóteses usuais para a análise de convergência do método de Newton, a continuidade Lipschitz da pimeira derivada do operador não-linear em questão é crítica.

Em outras palavras, manter o controle da derivada é um dos principais pontos na análise de convergência do método de Newton, veja e.g. Dennis e Schnabel [3], Ortega e Rheimboldt [?], Ortega [18], Ostrowski [19], Rall [21], e Traub e Wozniakowski [24]. Nos últimos anos, vários trabalhos relacionados com a convergência do método de Newton enfraquecem a hipótese da continuidade Lipschitz. Além de melhorar a convergência teórica, essa nova modificacão da condição Lipschitz permite unificar vários resultados anteriormente sem relação, trabalhos relacionados com este assunto inclui Alvarez, Botle e Munier [1], Ferreira e Svaiter [6], Wang [26] e Wang [27].

Nesta dissertação apresentaremos uma análise de convergência local e semi-local do método de Newton baseada no princípio majorante de Kantorovich, introducido por L.V. Kantorovich [9], veja também Ferreira e Svaiter [6]. Em nossa análise a condição Lipschitz clássica é relaxada usando uma função majorante. Nesta análise veremos claramente a relação da função majorante, a qual relaxa a continuidade Lipschitz, com o operador não-linear em consideração. Também, é possível obter para o caso local o maior raio possível para a unicidade de solução e também o raio ótimo de convergência para o método com respeito à função majorante. Para o caso semi-local, além da existência e unicidade de solução, mostraremos que para garantir taxa de convergência Q-quadrática não é necessário a existência de uma segunda raiz para a função majorante, necessitamos apenas que a função seja definida até sua primeira raiz. Obteremos também, neste caso, uma estimativa desta taxa baseada na derivada direcional *da derivada* da função majorante. Finalmente, vários resultados sobre o Método de Newton previamente sem relacão serão unificados.

Com o objetivo de tornar este trabalho mais completo, destinamos o capítulo II a uma breve revisão de análise convexa, funções analíticas, espaços de Banach, derivada de Fréchet e funções auto-concordantes. No capítulo III apresentaremos a definição formal do método de Newton e vários exemplos com objetivo de entender o seu comportamento. A parte central deste trabalho encontra-se nos capítulos IV e V. No capítulo IV estudaremos o comportamento local do método de Newton e no V estudaremos o seu comportamento semi-local (Teorema de Kantorovich), ambos do ponto de vista do princípio majorante.

Capítulo 2

Notação e resultados auxiliares

2.1 Análise Convexa

Os seguintes resultados elementares de análise convexa serão necessários para o desenvolvimento de nosso trabalho. Utilizamos nesta seção as referências bibliográficas de Izmailov e Solodov [8] e Lima [14].

Definição 2.1. *Um conjunto $D \subset \mathbb{R}^n$ é dito convexo se*

$$\alpha x + (1-\alpha)\overline{x} \in D, \qquad \forall\, x,\, \overline{x} \in D,\ \alpha \in [0,1].$$

O ponto $\alpha x + (1-\alpha)\overline{x}$, onde $\alpha \in [0,1]$, se chama a combinação convexa de x e $\overline{x}$ (com parâmetro α).

Definição 2.2. *Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Uma função $\varphi : D \to \mathbb{R}$ é chamada convexa se*

$$\varphi(\alpha x + (1-\alpha)y) \leq \alpha\varphi(x) + (1-\alpha)\varphi(y), \qquad \forall\, x, y \in D, \quad \forall\, \alpha \in [0,1].$$

Definição 2.3. *Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Uma função $\varphi : D \to \mathbb{R}$ é estritamente convexa se*

$$\varphi(\alpha x + (1-\alpha)y) < \alpha\varphi(x) + (1-\alpha)\varphi(y), \qquad \forall\, x, y \in D,\, x \neq y, \quad \forall\, \alpha \in (0,1).$$

Proposição 2.4. *Seja $I \subset \mathbb{R}$ um intervalo e $\varphi : I \to \mathbb{R}$ uma função convexa. Então dados $a < b < c$ com $a, b, c \in I$, temos*

$$\frac{\varphi(b) - \varphi(a)}{b - a} \leq \frac{\varphi(c) - \varphi(a)}{c - a} \leq \frac{\varphi(c) - \varphi(b)}{c - b}, \tag{2.1}$$

em particular, a função $s : I - \{d\} \to \mathbb{R}$ definida por

$$s(x) = \frac{\varphi(x) - \varphi(d)}{x - d}, \quad d \in int(I)$$

é não-decrescente. Se φ é estritamente convexa as desigualdades acima são estritas e a função s é crescente.

Demonstração. Notemos que

$$b = \frac{c - b}{c - a}\, a + \frac{b - a}{c - a}\, c;$$

Como $a < b < c$ então $(c - b)/(c - a) < 1$ e $(b - a)/(c - a) < 1$. Assim, usando este fato, a igualdade acima e a convexidade de φ obtemos

$$\varphi(b) \leq \frac{c - b}{c - a}\, \varphi(a) + \frac{b - a}{c - a}\, \varphi(c).$$

Com simples manipulações algébricas na inequação acima podemos concluir que

$$\varphi(b) - \varphi(a) \leq \left(\frac{c - b}{c - a} - 1 \right) \varphi(a) + \frac{b - a}{c - a}\, \varphi(c) = (\varphi(c) - \varphi(a))\, \frac{b - a}{c - a},$$

ou equivalentemente

$$\frac{\varphi(b) - \varphi(a)}{b - a} \leq \frac{\varphi(c) - \varphi(a)}{c - a}.$$

o que prova a primeira desigualdade. A segunda desigualdade de (2.1) é feita de modo análogo. É imediato concluír apartir de (2.1) que s é não-decrescente.

As demais afirmações seguem da convexidade estrita de φ e de desigualdades análogas.

$\square$

Corolário 2.5. *Se $\varphi : I \to \mathbb{R}$ é convexa no intervalo $I \subset \mathbb{R}$; então existem as derivadas laterais $\varphi_+(d)$ e $\varphi_-(d)$, para todo $d \in int(I)$.*

Demonstração. Como $d \in Int(I)$ existe $a \in I$ tal que $a < d$, portanto da Proposição 2.4, segue que

$$s(a) = \frac{\varphi(a) - \varphi(d)}{a - d} \leq \frac{\varphi(x) - \varphi(d)}{x - d} = s(x), \qquad \forall\, x > d,$$

isto é, s é limitada inferiormente. Devido a monotonicidade da função s existe o limite

$$\varphi_+(d) = \lim_{x \to d^+} s(x) = \lim_{x \to d^+} \frac{\varphi(x) - \varphi(d)}{x - d} = \inf_{x > d} \frac{\varphi(x) - \varphi(d)}{x - d}.$$

Analogamente prova-se que existe o limite $\varphi_-(d)$. □

Corolário 2.6. *Se $\varphi : I \to \mathbb{R}$ é convexa no intervalo I, então vale a seguinte desigualdade*

$$\varphi(x) \geq \varphi(d) + \varphi'_-(d)(x - d), \qquad \forall\, d \in int(I), \quad \forall\, x \in I.$$

Demonstração. Pelo Corolário 2.5 e Proposição 2.4 temos

$$\varphi'_-(d) = \sup_{x < d} \frac{\varphi(x) - \varphi(d)}{x - d} \leq \inf_{x > d} \frac{\varphi(x) - \varphi(d)}{x - d}.$$

Portanto, da segunda desigualdade na inequação acima segue-se que

$$\varphi(x) \geq \varphi(d) + \varphi'_-(d)(x - d), \qquad \forall\, x > d,$$

e da primeira que

$$\varphi(x) \geq \varphi(d) + \varphi'_-(d)(x - d), \qquad \forall\, x < d.$$

Assim, das duas últimas desigualdades, obtemos o resultado. □

Corolário 2.7. *Sejam $\varphi : I \to \mathbb{R}$ uma função diferenciável no intervalo aberto I. Então φ é convexa se, e somente se,*

$$\varphi(x) \geq \varphi(a) + \varphi'(a)(x - a), \qquad \forall\, x \in I, \quad \forall\, a \in int(I). \tag{2.2}$$

Demonstração. Se φ é diferenciavel então $\varphi' = \varphi'_-$ e o resultado segue do Corolário 2.6. A volta é imediato. □

Corolário 2.8. *Seja* $\varphi : [0,\ R) \to \mathbb{R}$ *uma função convexa, então* $t \longmapsto \varphi(t+\alpha) - \varphi(t)$ *é não decrescente para cada* $\alpha \geq 0$.

Demonstração. Definamos $\phi(t) = \varphi(t+\alpha) - \varphi(t)$. Devemos provar que $\phi(u) \leq \phi(v)$ para $u, v \in [0,\ R)$, com $u \leq v$ e $\alpha \geq 0$. Se $\alpha = 0$, o resultado é evidente. Suponhamos que $\alpha > 0$. Já que $u \leq u + \alpha \leq v + \alpha$ segue-se da Proposição 2.4 que

$$\frac{\varphi(u+\alpha) - \varphi(u)}{(u+\alpha-u)} \leq \frac{\varphi(v+\alpha) - \varphi(u)}{(v+\alpha-u)}. \tag{2.3}$$

Usando novamente a Proposição 2.4 e o fato $u \leq v \leq v + \alpha$, obtemos a desigualdade

$$\frac{\varphi(v+\alpha) - \varphi(u)}{(v+\alpha-u)} \leq \frac{\varphi(v+\alpha) - \varphi(v)}{(v+\alpha-v)}. \tag{2.4}$$

Combinando (2.3) e (2.4) obtemos o resultado desejado. □

Com o objetivo de uniformizar a notação utilizada no desenvolvimento de nosso trabalho enunciaremos duas proposições, e suas respectivas demonstrações, as quais são conseqüência direta da Proposição 2.4 e seus correspondentes corolários.

Proposição 2.9. *Sejam* $R > 0$ *e* $\varphi : [0,\ R) \to \mathbb{R}$ *uma função convexa.*

1) Para todo $t \in (0,\ R)$ *e* $0 \leq \tau \leq 1$, *vale a desigualdade* $(\varphi(t) - \varphi(\tau t))/t \leq \varphi'_{-}(t)(1-\tau)$;

2) Se $u, v \in [0,\ R)$, $u < v$ *e* $0 \leq \tau \leq 1$, *então* $(\varphi(u) - \varphi(\tau u)) \leq (\varphi(v) - \varphi(\tau v))(u/v)$.

Demonstração. Para provar o item 1, façamos $t = d$ e $x = \tau t$ no Corolário 2.6. Se $\tau = 1$ o item 2 é óbvio. Suponhamos agora que $0 \leq \tau < 1$. Devido a que, $u < v$ e $0 \leq \tau < 1$, temos que $\tau u < u < v$. Daquí, e da Proposição 2.4, obtemos

$$\frac{\varphi(u) - \varphi(\tau u)}{u - \tau u} \leq \frac{\varphi(v) - \varphi(\tau u)}{v - \tau u}. \tag{2.5}$$

De maneira semelhante obtemos a desigualdade

$$\frac{\varphi(v) - \varphi(\tau u)}{v - \tau u} \leq \frac{\varphi(v) - \varphi(\tau v)}{v - \tau v}, \tag{2.6}$$

onde a última inequação resulta usando a Proposição 2.4 e o fato que $\tau u < \tau v < v$. Combinando (2.5) e (2.6), concluímos que vale o item 2. □

Proposição 2.10. *Sejam $I \subset \mathbb{R}$ um intervalo e $\varphi : I \to \mathbb{R}$ uma função convexa. Se $u, v, w \in I$, $u < w$ e $u \leq v \leq w$, então*

$$\varphi(v) - \varphi(u) \leq [\varphi(w) - \varphi(u)]\,\frac{v-u}{w-u}.$$

Demonstração. Segue-se diretamente da Proposição 2.4. □

2.2 Espaço de Banach

Nesta subseção definiremos espaço de Banach, convergência de seqüências, derivada de operadores. Establecemos a norma de um operador o Lema de Banach, e três lemas importantes que serão de utilidade no nosso trabalho.

Definição 2.11. *Sejam $(X, ||.||)$ um espaço vetorial normado e $\{x_n\} \subset X$ 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

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

Definição 2.12. *Um espaço vetorial normado X é dito de Banach se toda seqüência de Cauchy em X é convergente.*

Lema 2.13. *Sejam X um espaço de Banach e $\{t_k\}$ uma seqüência de números reais monótona crescente e convergente para t_*. Consideremos uma seqüência $\{x_k\} \subset X$ que satisfaça a condição*

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

Então a seqüência $\{x_k\}$ é de Cauchy. Em particular, $\{x_k\}$ converge, digamos para $x_* \in X$. 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

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

Como $\{t_k\}$ é uma seqüência convergente para t_*, segue-se que $t_* - t_k$ pode ser tomado arbitrariamente pequeno para k suficientemente 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. □

Definição 2.14. *Sejam Z, W espaços de Banach e $\mathcal{L}(Z, W)$ o espaço dos operadores lineares de Z em W e $T \in \mathcal{L}(Z, W)$. Definimos a norma de operadores $\|.\|$ como sendo o número*

$$\|T\| := \sup\{\|Tu\|;\ \|u\| \leq 1\}.$$

Note que valem as seguintes desigualdades:

i) $\|Tu\| \leq \|T\|\|u\|$ *para todo* $T \in \mathcal{L}(Z, W)$ *e* $u \in Z$.

ii) $\|S\,T\| \leq \|S\|\|T\|$ *e* $\|S + T\| \leq \|S\| + \|T\|$ *para todo* $S, T \in \mathcal{L}(Z, W)$.

Podemos mostrar que com a norma definda acima o espaço $\mathcal{L}(Z, W)$ é um espaço de Banach, veja Kolmogorov e Fomin [12], pág. 215.

Definição 2.15. *Sejam $\Omega \subset X$ aberto e $G : \Omega \to W$ um operador não linear qualquer. Um operador $\mathrm{T} \in \mathcal{L}(X, W)$ é a derivada de Fréchet de G em $x_0 \in \Omega$ se para todo $w \in X$ tal que $(w + x_0) \in \Omega$ vale a igualdade*

$$\lim_{\|w\| \to 0} \frac{1}{\|w\|} \|G(x_0 + w) - G(x_0) - Tw\| = 0.$$

Dizemos que F é Fréchet derivável em Ω se F for Fréchet derivável em todo ponto $x \in \Omega$. Denotaremos a derivada de Fréchet de um operador F em x por $F'(x)$ e assim $F'(x) \in \mathcal{L}(X, Y)$, ou seja, F' é uma aplicação linear de Ω no espaço dos operadores lineares $\mathcal{L}(X, Y)$.

Se o operador F' é diferenciável em Ω, a sua derivada é chamada de segunda derivada e será denotada por F''. Assim, $F''(x)$ é um elemento do espaço $\mathcal{L}(X, \mathcal{L}(X, Y))$ dos operadores lineares que levam X em $\mathcal{L}(X, Y)$. O espaço $\mathcal{L}(X, \mathcal{L}(X, Y))$, com a norma de operadores da Definição 2.14, é novamente um espaço de Banach e é naturalmente identificado com o espaço dos operadores bilineares $\mathcal{L}_2(X, Y)$ de X em Y. Para simplificar a notação denotaremos o espaço $\mathcal{L}(X, \mathcal{L}(X, Y))$ por $\mathcal{L}_2(X, Y)$.

Se introduzem naturalmente os conceitos de terceira, quarta e, de um modo geral, de n-ésima derivada de uma aplicação F de X em Y, definindo-se a n-ésima derivada como a derivada da derivada de ordem $n-1$. A n-ésima derivada assim definida, denotada por $F^{(n)}(x)$, será obviamente, um elemento do espaço $\mathcal{L}_n(X, Y)$ dos operadores n-lineares de X em Y. Novamente identificamos o espaço $\mathcal{L}(X, \mathcal{L}(X, \ldots, \mathcal{L}(X, Y), \ldots))$ com o espaço dos operadores n-lineares $\mathcal{L}_n(X, Y)$. Para mais detalhes veja Kolmogorov e Fomin [12], pág. 478.

Utilizando as notações acima temos a seguinte observação:

Observação 2.16. *Sejam X, Y espaços de Banach e F é uma função diferenciável em $\Omega \subset X$, então $F'(x) \in \mathcal{L}(X, Y)$ e temos que*

$$\|F'(x)\| := \sup\{\|F'(x)u\|;\ \|u\| \leq 1\}.$$

Agora, se F' é uma função diferenciável em Ω, então $F''(x) \in \mathcal{L}_2(X, Y)$ e

$$\|F''(x)\| := \sup\{\|F''(x)(u, v)\|;\ \|u\| \leq 1, \|v\| \leq 1\}.$$

Por outro lado, se $\bar{x} \in \Omega$, temos das relações acima que $F'(\bar{x})^{-1}F''(x) \in \mathcal{L}_2(X, X)$ e

$$\|F'(\bar{x})^{-1}F''(x)\| := \sup\{\|F'(\bar{x})^{-1}F''(x)(u, v)\|;\ \|u\| \leq 1, \|v\| \leq 1\}.$$

De modo semelhante, se F é uma função n-vezes diferenciável em Ω, introduzimos a *norma do operador n-linear* $F'(\bar{x})^{-1}F^{(n)}(x) \in \mathcal{L}_n(X,X)$ *como sendo*

$$\|F'(\bar{x})^{-1}F^{(n)}(x)\| := \sup\{\|F'(\bar{x})^{-1}F^{(n)}(x)(u_1,\dots u_n)\|;\ \|u_1\| \leq 1,\dots \|u_n\| \leq 1\}.$$

As seguintes notações e resultados são usados até o fim de nosso trabalho. Sejam X, Y espaços de Banach. As bolas aberta e fechada em x serão denotadas, respectivamente por

$$B(x,\delta) = \{y \in X;\ \|x-y\| < \delta\} \quad \text{e} \quad B[x,\delta] = \{y \in X;\ \|x-y\| \leqslant \delta\}.$$

Sejam $\Omega \subseteq X$ e $F : \Omega \to Y$ uma função. Para $x \in int(\Omega)$ denotaremos por $F'(x)$ sua derivada de Fréchet.

Lema 2.17 (Lema de Banach)**.** *Sejam T um operador linear e I o operador identidade em um espaco de Banach X. Se $\|T-I\| < 1$, então T é não-singular e satisfaz a desigualdade*

$$\|T^{-1}\| \leq \frac{1}{1-\|T-I\|}.$$

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

$$\left\|(I-E)^{-1}\right\| \leq \frac{1}{1-\|E\|}. \tag{2.7}$$

Para isso, consideremos as sequências $\{S_k\}$ e $\{t_k\}$ definidas respectivamente por:

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

Agora notemos que

$$\|S_{k+1} - S_k\| = \left\|(I + E + E^2 + ... + E^{k+1}) - (I + E + E^2 + ... + E^k)\right\| \leq \|E\|^{k+1}.$$

Observemos também que

$$S_k(I-E) = (I + E + E^2 + ... + E^k)(I-E) = I - E^{k+1}. \tag{2.8}$$

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

$$\left\|I - (I-E^k)\right\| = \left\|E^k\right\| \leq \|E\|^k \quad e \qquad \lim_{k\to\infty} \|E\|^k = 0.$$

Assim, pela equação (2.8) temos que

$$\lim_{k\to\infty} S_k = (I - E)^{-1}.$$

Notemos ainda que

$$\left\|(I-E)^{-1}\right\| = \left\|\lim_{k\to\infty} S_k\right\| \leq \lim_{k\to\infty}(1 + \|E\| + \|E\|^2 + \cdots + \|E\|^k) = \lim_{k\to\infty} t_k = \frac{1}{1-\|E\|}.$$

Isto prova (2.7). Agora para concluir a prova tomemos $E = I - T$ e observando a hipótese $\|T - I\| < 1$, temos que $(I - E) = T$ é inversível e vale a estimativa para a norma da inversa T^{-1}. □

Lema 2.18. *SejamX um espaço de Banach, $C \subseteq X$ e $F : C \to Y$ uma função Fréchet derivável no int(C). Tomemos $\overline{x}$ no int(C) com $F'(\overline{x})$ não singular. Suponhamos que existam um $R > 0$ e uma função diferenciável f: $[0, R) \to \mathbb{R}$ tal que $B(\overline{x}, R) \subseteq C$, $f'(0) = -1$, f' é crescente e*

$$\|F'(\overline{x})^{-1}\left[F'(x) - F'(\overline{x})\right]\| \leq f'(\|x-\overline{x}\|) - f'(0), \qquad \forall\, x \in B(\overline{x}, R). \tag{2.9}$$

Se $x \in B(\overline{x}, R)$ e $f'(\|x - \overline{x}\|) < 0$, então $F'(x)$ é não singular e

$$\|F'(x)^{-1}F'(\overline{x})\| \leq \frac{1}{|f'(\|x-\overline{x}\|)|}.$$

Demonstração. Seja $x \in B(\overline{x}, R)$ tal que $f'(\|x - \overline{x}\|) < 0$. Usando (2.9) temos que

$$\|F'(\overline{x})^{-1}F'(x) - I\| = \|F'(\overline{x})^{-1}[F'(x) - F'(\overline{x})]\| \leq f'(\|x-\overline{x}\|) - f'(0) < -f'(0) = 1.$$

Deste modo, o Lema 2.17 e a última equação implicam que $F'(\overline{x})^{-1}F'(x)$ é não-singular, conseqüentemente $F'(x)$ também é não-singular. Além disso,

$$\|F'(x)^{-1}F'(\overline{x})\| \leq \frac{1}{1-\|F'(\overline{x})^{-1}F'(x) - I\|} \leq \frac{1}{1-(f'(\|x-\overline{x}\|) - f'(0))} = \frac{1}{|f'(\|x-\overline{x}\|)|},$$

onde foi usado que $f'(0) = -1$ e $f'(\|x - \overline{x}\|) < 0$. □

2.3 Funções Analíticas

Nesta seção abordaremos os conceitos de funções analíticas en $\mathbb{R}$ e funções analíticas em espaços de Banach as quais serão necessárias nos capítulos IV e V desta dissertação, quando trataremos o Teorema de Smale. O conteúdo desta seção foi baseado em Kolmogorov e Fomin [12] e Lima [14].

2.3.1 Funções analíticas em $\mathbb{R}$

Definição 2.19. *Uma função $f : I \to \mathbb{R}$, definida num intervalo aberto I, chama-se analítica quando, para cada $a \in I$, existe um $\epsilon > 0$ tal que a série de Taylor*

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}\, h^n,$$

converge para $f(a+h)$ desde que $|h| < \epsilon$.

Observação 2.20. *A fim de que a série de Taylor $\sum_{n=0}^{\infty}(f^{(n)}(a)/n!)h^n$ convirja para $f(a+h)$ é necessário e suficiente que $lim_{n\to\infty} r_n(h) = 0$, onde $r_n(h) = (f^{(n)}(a+\theta_n\, h)/n!)h^n$ com $0 < \theta_n < 1$.*

Observação 2.21. *$f : I \to \mathbb{R}$ é analítica, a sua derivada é também uma função analítica. Com efeito, se f é representada em uma vizinhança de $a \in I$ por uma série*

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}\, h^n,$$

a sua derivada será representada na mesma vizinhança pela série

$$\sum_{n=0}^{\infty} \frac{f^{(n+1)}(a)}{n!}\, h^n.$$

Decorre daí que f é na verdade de classe C^{∞} e as suas derivadas sucessivas f', f'', $\ldots$ são também funções analíticas.

Lema 2.22. *Se $0 \le t < 1$, então $\sum_{i=0}^{\infty}(i+2)(i+1)\, t^i = 2/(1-t)^3$.*

Demonstração. Consideremos a funcão $g : (-1, 1) \to \mathbb{R}$, definida por $g(t) = (1 - t)^{-1}$. É fácil mostrar que ela é analítica em $(-1, 1)$ e que

$$g'(t) = (1 - t)^{-2}, \quad g''(t) = 2(1 - t)^{-3}, \ldots, \quad g^{(i)}(t) = i!(1 - t)^{-(i+1)}. \tag{2.10}$$

Pela definição 2.19 podemos escrever g da seguinte maneira

$$g(t) = \sum_{i=0}^{\infty} \frac{g^{(i)}(0)}{i!} t^i.$$

Agora combinando (2.10) e a igualdade acima, obtemos que $g(t) = \sum_{i=0}^{\infty} t^i$, a qual derivando duas vezes resulta que $g''(t) = \sum_{i=0}^{\infty} (i + 2)(i + 1) t^i$. Novamente da segunda equação em (2.10) e da última equação segue-se o resultado. □

2.3.2 Funções analíticas em espaços de Banach

Teorema 2.23. *Sejam X, Y espaços de Banach e F uma função de Xem Y definida num aberto $\Omega \subset X$ e tal que $F^{(n)}(x)$ está definida e é uniformemente contínua em Ω. Então, se cumpre a relação*

$$F(x + h) = F(x) + F'(x).h + 1/2!\, F''(x)h^2 + \ldots + 1/n!\, F^{(n)}(x)h^n + r(x, h),$$

onde $\|r(x, h)\| = o(\|h\|^n)$, $h^2 := (h, h)$ e $h^n := (h, \ldots, h)$.

Demonstração. Veja Kolmogorov e Fomin [12], pag. 480. □

Definição 2.24. *Sejam X, Y espaços de Banach e $\Omega \subset X$. Dizemos que uma função $F : \Omega \to Y$ é analítica, se para cada ponto $a \in \Omega$ existe uma série de potências*

$$\sum_{n=0}^{\infty} \frac{F^{(n)}(a)}{n!} h^n,$$

com raio de convergência $\delta > 0$, isto é,

$$F(x) = \sum_{n=0}^{\infty} \frac{F^{(n)}(a)}{n!} (x - a)^n, \qquad \|x - a\| < \delta.$$

2.4 Funções a-auto-concordantes

Nesta subseção enunciaremos alguns conceitos relacionados a funções auto-concordantes, os quais serão utilizados nos capitulos IV e V. Estes conceitos foram extraídos de Nesterov e Nemirovskii [17] e Alvarez, Botle e Munier [1].

Definição 2.25. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto convexo. Uma função $g : \Omega \to \mathbb{R}$ é chamada a-auto-concordante com o parâmetro $a > 0$, se $g \in C^3(\Omega)$, i.e., três vezes continuamente diferenciável em Ω, é uma função convexa em Ω e satisfaz a seguinte inequação*

$$|g'''(x)[h,h,h]| \leqslant 2a^{-1/2}(g''(x)[h,h])^{3/2}, \qquad \forall\, x \in \Omega,\ \forall\, h \in \mathbb{R}^n. \tag{2.11}$$

Tomemos $\bar{x} \in \Omega$ tal que $g''(\bar{x})$ é não-singular. Defina o espaço de Banach $X := (\mathbb{R}^n, \langle .,. \rangle_{\bar{x}})$ como sendo o espaço Euclideano $\mathbb{R}^n$ com um novo produto interno e a norma associada definida, respectivamente, por

$$\langle u,v \rangle_{\bar{x}} := a^{-1}\langle g''(\bar{x})u, v\rangle, \qquad \|u\|_{\bar{x}} := \sqrt{\langle u,u \rangle_{\bar{x}}}. \tag{2.12}$$

Assim, no espaço de Banach X, a bola aberta e fechada de raio $r > 0$ centrada no ponto $\bar{x}$ (elipsóide de Dikin de raio $r > 0$ centrado em $\bar{x}$) são definidas, respectivamente, como

$$W_r(\bar{x}) := \{x \in \mathbb{R}^n : \|x - \bar{x}\|_{\bar{x}} < r\}, \qquad W_r[\bar{x}] := \{x \in \mathbb{R}^n : \|x - \bar{x}\|_{\bar{x}} \leq r\}.$$

Proposição 2.26. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto convexo e seja $g : \Omega \to \mathbb{R}$ uma função a-auto-concordante. Então,*

$$|g'''(x)[h_1,h_2,h_3]| \leqslant 2a^{-1/2}\Pi_{i=1}^3(g''(x)[h_i,h_i])^{1/2}, \qquad \forall\, x \in \Omega,\ \forall\, h_1,h_2,h_3 \in \mathbb{R}^n.$$

Demonstração. Ver Proposição 9.1.1, Apêndice 1, pp.361 de Nesterov e Nemirovskii [17]. □

Proposição 2.27. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto convexo e seja $g : \Omega \to \mathbb{R}$ uma função a-auto-concordante. Então vale*

$$g''(x)[h,h] \leq \frac{1}{(1 - \|x - \bar{x}\|_{\bar{x}})^2} g''(\bar{x})[h,h], \qquad \forall\, x \in W_1(\bar{x}) \cap \Omega, \quad \forall\, h \in \mathbb{R}^n.$$

Demonstração. Fixemos $x \in W_1(\bar{x})$. Denotemos $e = x - \bar{x}$, $y(t) = \bar{x} + te$ e

$$I = \{t \geq 0 : \|y(t) - \bar{x}\|_{\bar{x}} < 1\}.$$

Fixemos $\bar{t} \in I$ tal que $y(t) \in \Omega$ e $h \in \mathbb{R}^n$. Notemos que $\bar{t} \leq 1$. Seja $J = [0, \bar{t}\,]$ e definamos as funções $\psi : J \to \mathbb{R}$ e $\phi : J \to \mathbb{R}$ como

$$\psi(t) := g''(y(t))[e, e], \qquad \phi(t) := g''(y(t))[h, h].$$

Como $y'(t) = e$, derivando ψ e usando a Proposição 2.26 obtemos a desigualdade

$$|\psi'(t)| = |g'''(y(t))\,[y'(t), h, h]| = |g'''(y(t))[e, e, e]| \leq 2a^{-1/2} g''(y(t))[e, e]^{3/2}.$$

Deste modo, combinando a última desigualdade com a definição de ψ concluímos que

$$|\psi'(t)| \leq 2a^{-1/2}(\psi(t))^{3/2}. \tag{2.13}$$

Agora, derivando ϕ e novamente usando que $y'(t) = e$ e a Proposição 2.26 temos

$$\begin{aligned}|\phi'(t)| = |g'''(y(t))\,[y'(t), h, h]| &= |g'''(y(t))[e, h, h]| \\ &\leq 2a^{-1/2}(g''(y(t))[h, h])(g''(y(t))[e, e])^{1/2},\end{aligned}$$

e assim usando a definição das funções ψ e ϕ obtemos a seguinte desigualdade

$$|\phi'(t)| \leq 2a^{-1/2}\phi(t)\psi(t)^{1/2}. \tag{2.14}$$

Da relação em (2.13), temos duas possibilidades:

a) $\psi(t) = 0$, para todo $t \in J$;

b) $\psi(t) > 0$, para todo $t \in J$.

Para o caso **a**, usando (2.14) concluimos que $\phi(t) = \phi(0)$.

Vamos estudar o caso **b**. Primeiro notemos que (2.13) e equivalente a seguinte inequação

$$|(\psi^{-1/2}(t))'| \leq a^{-1/2} \qquad \forall\, t \in J.$$

Esta inequação implica, evidentemente, que $-a^{-1/2} \leq (\psi^{-1/2}(t))'$, a qual após ser integrada implica que

$$\psi^{-1/2}(t) \geq \psi^{-1/2}(0) - ta^{-1/2}, \qquad \forall\, t \in J.$$

No último caso, em virtude da definição de ψ, do vetor e e da equação (2.12) temos

$$\psi(0) = g''(y(0))[e,\, e] = a\|e\|_{\bar{x}}^2 = a\|x - \bar{x}\|_{\bar{x}}^2.$$

Assim, sustituindo o valor $\psi(0)$ na última desigualdade e após algunas manipulações algébricas obtemos

$$\psi^{1/2}(t) \leq \frac{a^{1/2}\|x - \bar{x}\|_{\bar{x}}}{1 - t\|x - \bar{x}\|_{\bar{x}}}, \qquad \forall\, t \in J.$$

Conseqüentemente, usando a desigualdade acima juntamente com (2.14), resulta que

$$|\phi'(t)| \leq \frac{2\|x - \bar{x}\|_{\bar{x}}\, \phi(t)}{1 - t\|x - \bar{x}\|_{\bar{x}}}, \qquad \forall\, t \in J. \tag{2.15}$$

Da desigualdade acima, temos duas possibilidades:

c) $\phi(t) = 0$, para todo $t \in J$;

d) $\phi(t) > 0$, para todo $t \in J$.

Para o caso **c** temos que $\phi \equiv 0$ em J. Para o caso **d**, segue de (2.15), após integração, em t, que

$$\left| \ln \frac{\phi(t)}{\phi(0)} \right| \leq 2 \ln \frac{1}{1 - t\|x - \bar{x}\|_{\bar{x}}},$$

Da última desigualdade obtemos $\phi(t) \leq \phi(0)/(1 - t\,\|x - \bar{x}\|_{\bar{x}})^2$, a qual, pela definição de ϕ implica que

$$g''(y(t))[h,\, h] \leq \frac{g''(y(0))[h,\, h]}{(1 - t\|x - \bar{x}\|_{\bar{x}})^2}, \qquad \forall\, t \in J.$$

Se $y(t) = x$, então $t = 1$ e daí segue-se o resultado desejado. $\square$

O seguinte resultado é uma combinação das duas proposições acima, ele apareceu pela primeira vez em Alvarez, Botle e Munier [1], Lema 5.1.

Lema 2.28. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto convexo e seja $g : \Omega \to \mathbb{R}$ uma função a-auto-concordante. Assumamos que $W_1(\bar{x}) \subset \Omega$. Então vale a desigualdade*

$$\|g''(\bar{x})^{-1}g'''(x)\|_{\bar{x}} \leq \frac{2}{(1 - \|x - \bar{x}\|_{\bar{x}})^3}, \qquad \forall\, x \in W_1(\bar{x}).$$

Demonstração. Sejam $x \in W_1(\bar{x})$ e $h_1,\ h_2,\ h_3 \in \mathbb{R}^n$. Da definição de produto interno associado a $\bar{x}$ temos

$$\begin{aligned}\langle g''(\bar{x})^{-1}g'''(x)h_1h_2, h_3\rangle_{\bar{x}} &= a^{-1}\langle g''(\bar{x})\left(g''(\bar{x})^{-1}g'''(x)\right)h_1h_2, h_3\rangle \\ &= a^{-1}g'''(x)[h_1, h_2, h_3].\end{aligned}$$

Tomando valor absoluto na última igualdade e da Proposição 2.26 segue que

$$|g''(\bar{x})^{-1}g'''(x)h_1, h_2, h_3\rangle_{\bar{x}}| \leq 2a^{-3/2}\Pi_{i=1}^3(g''(x)[h_i, h_i])^{1/2}.$$

Já que $\|g''(\bar{x})^{-1}g'''(x)\|_{\bar{x}} := \sup\{|\langle g''(\bar{x})^{-1}g'''(x)h_1h_2, h_3\rangle_{\bar{x}}| : \|h_i\|_{\bar{x}} \leqslant 1,\ i = 1, 2, 3\}$, temos da última inequação que

$$\|g''(\bar{x})^{-1}g'''(x)\|_{\bar{x}} \leq 2a^{-3/2}\sup\left\{\Pi_{i=1}^3(g''(x)[h_i, h_i])^{1/2} : \|h_i\|_{\bar{x}} \leq 1\right\}. \tag{2.16}$$

Assim, segue de (2.16) e da Proposição 2.27 que

$$\begin{aligned}\|g''(\bar{x})^{-1}g'''(x)\|_{\bar{x}} &\leq \frac{2a^{-3/2}}{(1 - \|x - \bar{x}\|_{\bar{x}})^3}\sup\left\{\Pi_{i=1}^3(g''(\bar{x})[h_i, h_i])^{1/2} : \|h_i\|_{\bar{x}} \leq 1\right\} \\ &= \frac{2}{(1 - \|x - \bar{x}\|_{\bar{x}})^3}\sup\left\{\Pi_{i=1}^3 a^{-1/2}(g''(\bar{x})[h_i, h_i])^{1/2} : \|h_i\|_{\bar{x}} \leq 1\right\} \\ &= \frac{2}{(1 - \|x - \bar{x}\|_{\bar{x}})^3}\sup\left\{\Pi_{i=1}^3\|h_i\|_{\bar{x}} : \|h_i\|_{\bar{x}} \leq 1\right\}.\end{aligned}$$

Portanto, como $\sup\{\Pi_{i=1}^3\|h_i\|_{\bar{x}} : \|h_i\|_{\bar{x}} \leq 1,\ i = 1, 2, 3\} \leq 1$ o resultado segue-se. $\square$

Capítulo 3

O método de Newton

3.1 Introdução

Nosso objetivo neste capítulo é apresentar a definição do método de Newton e analisar seu comportamento em alguns casos específicos. Como veremos adiante uma condição para que o método convirja é que o ponto inicial deve ser tomado "perto"da solução do problema em consideração. Devido a este fato podemos dizer que o método de Newton é localmente convergente.

A idéia básica do Método de Newton para encontrar um zero de uma função diferenciável não-linear é bastante simples. O método consiste em aproximar o zero da função não-linear por zeros de aproximações lineares sucessivas em uma vizinhança apropriada. Mais precisamente, queremos resolver a equação

$$F(x) = 0, \qquad x \in \Omega,$$

onde $\Omega \subset X$ é um conjunto aberto e $F : \Omega \to Y$ é uma função continuamente diferenciável não-linear. Começamos com um ponto inicial $x_0 \in \Omega$ construímos a aproximação linear de F, isto é,

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

Então ao invés de resolver a equação não-linear $F(x) = 0$, resolvemos a equação linear

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

Se a derivada $F'(x_0)$ for não-singular a equação linear acima tem uma única solução, digamos

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

Note que x_1 pode ou não pertencer a Ω, ou ainda, $F'(x_1)$ pode ser singular. Para assegurar que o procedimento possa ser repetido infinitamente, precisamos de hipóteses adequadas sobre F e o ponto inicial x_0. Suponhamos, por um momento, que o processo possa ser repetido indefinidamente, isto é, $x_1 \in \Omega$ e $F'(x_1)$ é não-singular e assim sucesivamente. Portanto, podemos definir uma seqüência $\{x_k\}$ contida em Ω dada por

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

ainda assim não temos garantia de convergência da sequência $\{x_k\}$. Novamente, necessitamos de hipóteses adequadas sobre F e o ponto inicial x_0 para garantir convergência quadrática da seqüência $\{x_k\}$ para um zero de F.

Observação 3.1. *Suponhamos que $\Omega \subset \mathbb{R}^n$ seja um conjunto aberto e $F : \Omega \to \mathbb{R}^n$ é uma função continuamente diferenciável em Ω e contínua em $\bar{\Omega}$, o fecho de Ω. Se $\{x_k\}$ está bem definida ($x_k \in \Omega$ e $F'(x_k) \neq 0 \; \forall \; k$), converge para algum $x_* \in \bar{\Omega}$ e $\{F'(x_k)\}$ é limitada, então $F(x_*) = 0$. De fato, de* (3.1) *temos*

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

Tomando limite nesta equação, segue imediatamente que $F(x_*) = 0$, pois $\{x_k\} \subset \Omega$, $\lim_{k\to\infty} x_k = x_*$, F é continua em $\bar{\Omega}$ e $\{F'(x_k)\}$ é limitada.

Exemplo 3.2. *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, independente da escolha do ponto inicial, que a seqüência de Newton para resolver $f(x) = 0$ está bem definida e converge para a único zero da função, $x_* = 0$. Notemos que, neste caso, a taxa de convergência da sequência $\{x_k\}$ é linear e que $f'(x_*) = 0$.

Veremos que a condição da derivada não se anular na solução é suficiente para garantir convergência quadrática.

Exemplo 3.3. *Consideremos a função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x/\sqrt{1+x^2}$. Notemos que $x = 0$ é o único zero de f e que*

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

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

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

Assim:

i) se x_0 for tomado tal que $|x_0| = 1$, a seqüência de Newton irá oscilar nos valores 1 e -1, e portanto não convergirá;

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

iii) se x_0 for tomado tal que $|x_0| > 1$, a seqüência de Newton não convergirá.

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

Exemplo 3.4. *Consideremos a função $f : (0, +\infty) \to \mathbb{R}$ dada por $f(x) = 1 - 1/x$. Notemos que $x_* = 1$ é o único zero de f e que $f'(x) = 1/x^2 \neq 0$ para todo $x \in (0, +\infty)$. A seqüência de Newton para encontrar a solução de $f(x) = 0$ é dada por*

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

Notemos que:

i) *se* $x_0 \geq 2$ *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* $x_* = 1$.

Logo, para este caso, temos que a boa definição da seqüência e sua convergência é condicionada ao ponto inicial ser tomado próximo da solução.

Exemplo 3.5. *Consideremos a função* $f : \mathbb{R} \to \mathbb{R}$ *dada por* $f(x) = x(x-1)(x+1)$. *Notemos que os zeros de* f *são* 0, 1 *e* -1. *Observemos que* $f'(x) = 3x^2 - 1$ *e a seqüência de Newton para encontrar uma solução de* $f(x) = 0$ *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 em uma iteracão);*

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

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, tomando $|x_0| = 1/2$ *temos que a seqüência de Newton converge em um único passo!. Agora tomando* x_0 *como sendo a raíz real do polinômio* $q(x) = 2\sqrt{3}x^3 + 3x^2 - 1$, *teremos* $x_1 = -1/\sqrt{3}$ *que é um ponto singular da derivada. Isto mostra que a seqüência de Newton pode não estar bem definida com essa escolha de* x_0*.*

Como podemos notar nos exemplos acima, para assegurar a boa definição da seqüência de Newton ($x_k \in \Omega$ e $F'(x_k) \neq 0, \quad \forall k$) e a sua convergência a escolha do ponto inicial deve ser tomado "próximo"da solução. Nos capítulos IV e V vamos estudar critérios para a escolha deste ponto.

Capítulo 4

Análise local do método de Newton

4.1 Introdução

Como vimos no capítulo anterior, o método de Newton pode não convergir ou pode falhar em gerar uma seqüência infinita quando um ponto singular da derivada é encontrado. Para garantir convergência quadrática do método de Newton para a solução da equação não linear, algumas condições devem ser impostas. Por exemplo, o análise de convergência clássica requer que o ponto inicial seja tomado "suficientemente próximo"da solução e a derivada do operador não-linear em consideração seja não-singular nesta solução. Além disso, a continuidade Lipschitz da derivada é assumida, ver e.g., Dennis e Schnabe [3] e Traub e Wozniakowski [24]. Um desvantagen do análise é que a solução deve ser conhecida ou dada *a priori*. Por outro lado, isto tem vantagem para dar o raio ótimo da bola de convergência com respeito à constante Lipschitz, veja Rall [21], Traub e Wozniakowski [24] e Wang [27].

Nosso objetivo neste capítulo é apresentar uma análise de convergência local do método de Newton baseado no princípio majorante, introduzido por L. V. Kantorovich em [9], veja também Ferreira [4], Ferreira e Svaiter [6], Wayne [27]. Neste análise a condição Lipschitz clássica é relaxada por meio de uma função majorante. Na análise apresentada podemos ver claramente a relação da função majorante, a qual relaxa a continuidade Lipschitz, com o operador não-linear em consideração. Também, isto nos permite obter o maior raio para a unicidade da solução e também o raio ótimo de convergência para o método com respeito à função majorante. Além disso, vários resultados locais sobre o método de Newton ainda não relacionados serão unificados.

4.1.1 Análise local para o método de Newton

Neste capítulo provaremos e explicaremos o teorema local para o método de Newton. Para isto, primeiro demonstraremos os resultados para a função escalar majorante. Em seguida, boa definição, convergência, taxa de convergência, unicidade da solução e a bola ótima de convergência serão establecidos. O conteúdo deste capítulo foi baseado em Ferreira [4]. A afirmação do teorema é:

Teorema 4.1. *Sejam X, Y espaços de Banach, $\Omega \subseteq X$ aberto e $F : \Omega \to Y$ continuamente diferenciável. Sejam $R > 0$, $x_* \in \Omega$ e $\kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}$. Suponhamos que $F(x_*) = 0$, $F'(x_*)$ não-singular e que exista uma função continuamente diferenciável $f : [0,\ R) \to \mathbb{R}$ tais que*

$$\left\|F'(x_*)^{-1}\left[F'(x) - F'(x_* + \tau(x - x_*))\right]\right\| \leq f'\left(\|x - x_*\|\right) - f'\left(\tau\|x - x_*\|\right), \qquad (4.1)$$

para $\tau \in [0, 1]$, $x \in B(x_*, \kappa)$, $\|x - x_*\| < R$ e

h1) *$f(0) = 0$ e $f'(0) = -1$;*

h2) *f' é convexa e crescente.*

Sejam $\nu := \sup\{t \in [0, R) : f'(t) < 0\}$, $\rho := \sup\{t \in (0, \nu) : f(t)/(tf'(t)) - 1 < 1\}$ e

$$r := \min\{\kappa,\ \rho\}.$$

Então, as seqüências com ponto inicial $x_0 \in B(x_*, r)/\{x_*\}$ e $t_0 = \|x_* - x_0\|$, respectivamente,

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad t_{k+1} = |t_k - f(t_k)/f'(t_k)|, \qquad k = 0, 1, \ldots, \qquad (4.2)$$

estão bem definidas, $\{t_k\}$ *é decrescente, está contida em* $(0, r)$ *e converge para* 0, $\{x_k\}$ *está contida em* $B(x_*, r)$*, converge para o ponto* x_* *e valem as desigualdades*

$$\|x_* - x_{k+1}\| \leq \left[t_{k+1}/t_k^2\right] \|x_k - x_*\|^2, \quad t_{k+1}/t_k^2 \leq D_- f'(t_0)/(2|f'(t_0)|), \quad k = 0, 1, \ldots, \tag{4.3}$$

e $\{t_{k+1}/t_k^2\}$ *é não crescente. Além disso,* x_* *é o único zero de* F *em* $B(x_*, \sigma)$*, onde* $0 < \sigma := \sup\{0 < t < \kappa : f(t) < 0\}$. *Se adicionalmente* $f(\rho)/(\rho f'(\rho)) - 1 = 1$ *e* $\rho < \kappa$*, então* $r = \rho$ *é o melhor raio de convergência possível.*

Observação 4.2. *Combinando as inequacões em* (4.3)*, obtemos que a seqüência* $\{x_k\}$ *converge Q-quadráticamente para* x_**. Além disso, já que* $\{t_{k+1}/t_k^2\}$ *é não crescente temos que* $t_{k+1}/t_k^2 < t_1/t_0^2$*, para* $k = 0, 1, \ldots$*. Assim, a primeira inequação em* (4.3) *implica* $\|x_* - x_{k+1}\| \leq [t_1/t_0^2] \|x_k - x_*\|^2$*, para* $k = 0, 1, \ldots$*. Portanto,*

$$\|x_* - x_k\| \leq t_0 \left(t_1/t_0\right)^{2^k - 1}, \qquad k = 0, 1, \ldots .$$

onde $t_1/t_0 < 1$*, porque* $\{t_k\}$ *é decrescente.*

Para provar o Teorema 4.1 primeiro necessitaremos de alguns resultados. De agora em adiante, aceitaremos que todas as hipótese do teorema são válidas.

4.1.2 A função majorante

Nosso primeiro objetivo é provar que a constante κ associada a Ω e as constantes ν, ρ e σ associadas à função majorante f são positivas. Também, provaremos todas as afirmações no Teorema 4.1 envolvendo a seqüência $\{t_k\}$. Iniciaremos provando que κ, ν e σ são positivas.

Proposição 4.3. *As constantes* κ, ν *e* σ *são positivas e* $t - f(t)/f'(t) < 0$*, para* $t \in (0, \nu)$.

Demonstração. Desde que Ω é aberto e $x_* \in \Omega$ é imediato concluir que $\kappa > 0$. Como $f'(0) = -1$, existe $\delta > 0$ tal que $f'(t) < 0$ para todo $t \in (0, \delta)$. Assim, $\nu > 0$. Agora, devido que $f(0) = 0$ e $f'(0) = -1$, existe $\delta > 0$ tal que $f(t) < 0$ para todo $t \in (0, \delta)$. Daquí $\sigma > 0$.

Devido a f' ser crescente, temos que f é estritamente convexa. Assim, o Corolário 2.7 implica que

$$f(0) > f(t) - tf'(t), \qquad \forall\, t \in (0,\, R).$$

Como $f(0) = 0$ e $f'(t) < 0$ para todo $t \in (0,\, \nu)$, a inequação desejada segue da inequação acima. $\square$

De **h2** no Teorema 4.1 e da definição de ν, temos que $f'(t) < 0$ para todo $t \in [0,\, \nu)$. Portanto, a aplicação iteração de Newton associada à função majorante f é bem definida em $[0,\, \nu)$. Denotaremos esta por n_f,

$$\begin{array}{rcl} n_f : [0,\, \nu) & \to & (-\infty,\, 0] \\ t & \mapsto & t - f(t)/f'(t). \end{array} \tag{4.4}$$

Proposição 4.4. *A aplicação* $(0,\, \nu) \ni t \mapsto |n_f(t)|/t^2$ *é não decrescente e satisfaz a estimativa*

$$\frac{|n_f(t)|}{t^2} \le \frac{D_- f'(t)}{2|f'(t)|}, \qquad \forall\, t \in (0,\, \nu).$$

Demonstração. Desde que $f(0) = 0$, $f' < 0$ e é crescente em $[0, \nu)$, depois de simples manipulações temos

$$\frac{|n_f(t)|}{t^2} = \frac{tf'(t) - f(t)}{t^2|f'(t)|} = \frac{1}{|f'(t)|}\int_0^1 \frac{f'(t) - f'(\tau t)}{t}\, d\tau, \qquad \forall\, t \in (0,\, \nu). \tag{4.5}$$

De **h2** temos que f' é convexa, e em virtude da Proposição 2.9 item 2, afirmamos que a aplicação $(0,\, \nu) \ni t \mapsto (f'(t) - f'(\tau t))/t$ é positiva e não-decrescente, para $\tau \in [0, 1]$, bem como sua integral. Por outro lado, a aplicação $[0,\, \nu) \ni t \mapsto 1/|f'(t)|$ é positiva e crescente, já que $f' < 0$ e crescente em $[0,\, \nu)$. Assim, o lado direito de (4.5) é não-decrescente, portanto a primeira afirmação está provada.

Usando a Proposição 2.9, item 1, temos que a desigualdade $(f'(t) - f'(\tau t))/t \le D_- f'(t)\,(1 - \tau)$, vale para qualquer $0 \le \tau \le 1$. Daí substituindo a última expressão em (4.5) e avaliando a integral, resultará a inequação desejada. $\square$

Proposição 4.5. *A constante ρ é positiva. Como uma conseqüência, $|n_f(t)| < t$ para todo $t \in (0,\, \rho)$.*

Demonstração. Usando a definição (4.4), a Proposição 4.3 e algumas simples manipulações algébricas resulta

$$0 < \frac{f(t)}{tf'(t)} - 1 = \frac{|n_f(t)|}{t}, \qquad \forall\, t \in (0,\, \nu). \tag{4.6}$$

Agora, $\lim_{t\to 0} |n_f(t)|/t = \lim_{t\to 0}(|n_f(t)|/t^2)\, t = 0$ porque a Proposição 4.4 implica que $|n_f(t)|/t^2$ é limitada perto de zero. Desta maneira, da definição de límite de uma função e de (4.6), existe $\delta > 0$ tal que $0 < f(t)/(tf'(t)) - 1 < 1$, para todo $t \in (0, \delta)$. Daquí ρ é positiva. Além disso, pela definição de ρ resulta que $f(t)/(tf'(t)) - 1 < 1$, para $t \in (0, \rho)$. Assim, usando (4.6) segue-se a segunda parte. $\square$

A definição de $\{t_k\}$ no Teorema 4.1 é equivalente à seguinte

$$t_0 = \|x_0 - x_*\|\,, \quad t_{k+1} = |n_f(t_k)|, \qquad k = 0, 1, 2, \ldots\,. \tag{4.7}$$

Corolário 4.6. *A sequência $\{t_k\}$ é bem definida, é decrescente e está contida em $(0, \rho)$. Além disso, $\{t_{k+1}/t_k^2\}$ é não crescente, $\{t_k\}$ converge para 0 e satisfaz a estimativa*

$$t_{k+1}/t_k^2 \le [D_- f'(t_0)/(2|f'(t_0)|)], \qquad k = 0, 1, \ldots\,.$$

Demonstração. Das Proposições 4.3, 4.5 e das equações (4.4) e (4.7) é fácil concluir que $\{t_k\}$ está bem definida, é decrescente e está contída em $(0, \rho)$.

A definição de $\{t_k\}$ na equação (4.7) implica que

$$t_{k+1}/t_k^2 = |n_f(t_k)|/t_k^2, \quad k = 0, 1, \ldots\,. \tag{4.8}$$

Já que $\{t_k\}$ é decrescente, a primeira parte da Proposição 4.4 implica que $\{|n_f(t_k)|/t_k^2\}$ é não crescente. Deste modo, de (4.8) concluimos que $\{t_{k+1}/t_k^2\}$ é também não crescente. De novo, devido que $\{|n_f(t_k)|/t_k^2\}$ é não crescente temos

$$|n_f(t_k)|/t_k^2 \le |n_f(t_0)|/t_0^2, \qquad k = 0, 1, \ldots\,. \tag{4.9}$$

Como $\{t_k\}$ é decrescente $t_k < t_0$, para $k = 0, 1, \ldots$, que combinado com (4.8) e (4.9) implica na desigualdade

$$t_{k+1} \leq [|n_f(t_0)|/t_0]t_k, \qquad k = 0, 1, \ldots.$$

Em vista de $t_0 = \|x_* - x_0\| < r \leq \rho$, a Proposição 4.5 implica que $|n_f(t_0)|/t_0 < 1$. Assim, a inequação acima implica que $\{t_k\}$ converge para 0. Concluíndo, notemos que a segunda parte da Proposição 4.4 implica na desigualdade $|n_f(t_0)|/t_0^2 \leq D_-f'(t_0)/(2|f'(t_0)|)$ que junto com (4.8) e (4.9) resulta a inequação desejada. □

4.1.3 Relação entre a função majorante e o operador não-linear

Nesta subseção apresentamos as principais relações entre a função majorante f com o operador não-linear F em consideração.

Lema 4.7. *Seja $x \in \Omega$. Se $\|x - x_*\| < min\{\nu, \kappa\}$, então $F'(x)$ é não singular e*

$$\|F'(x)^{-1}F'(x_*)\| \leqslant 1/|f'(\|x - x_*\|)|.$$

Em particular, F' é não-singular em $B(x_*, r)$.

Demonstração. Façamos $x_* = \overline{x}$ no Lema 2.18, pois $\|x - x_*\| < min\{\nu, \kappa\}$ e $f'(t) < 0$ para $t < \nu$. □

A iteração de Newton produz pontos que são os zero da linearização de F, tais pontos, também são a primeira-ordem da expansão de Taylor para F. Assim, estudaremos o erro na linearização de F em pontos de Ω.

$$E_F(x, y) := F(y) - [F(x) + F'(x)(y - x)], \qquad y,\, x \in \Omega. \tag{4.10}$$

Limitaremos este erro pelo erro na linearização da função majorante f

$$e_f(t, u) := f(u) - [f(t) + f'(t)(u - t)], \qquad t,\, u \in [0, R). \tag{4.11}$$

Lema 4.8. *Se* $\|x_* - x\| < \kappa$, *então vale a desigualdade*

$$\|F'(x_*)^{-1}E_F(x, x_*)\| \le e_f(\|x - x_*\|, 0).$$

Demonstração. Desde que $B(x_*, \kappa)$ é convexa, obtemos que $x_* + (1-u)(x - x_*) \in B(x_*, \kappa)$, para $0 \le u \le 1$. Deste modo, como F é continuamente diferenciável em Ω, a definição de E_F e algumas simples manipulações resulta na estimativa

$$\|F'(x_*)^{-1}E_F(x, x_*)\| \le \int_0^1 \left\|F'(x_*)^{-1}[F'(x) - F'(x_* + (1-u)(x - x_*)]\right\| \, \|x_* - x\| \; du.$$

Da última equação com a hipótese (4.1) obtemos a desigualdade

$$\|F'(x_*)^{-1}E_F(x, x_*)\| \le \int_0^1 \left[f'\left(\|x - x_*\|\right) - f'\left((1-u)\|x - x_*\|\right)\right] \|x - x_*\| \; du.$$

Avaliando a última integral e usando definição de e_f, segue-se a afirmação. □

O Lema 4.7 garante a não-singularidade de F' em $B(x_*, r)$ e assim uma boa definição da aplicação iteração de Newton. Chamaremos de N_F, a aplicação iteração de Newton para F em tal região, isto é

$$\begin{array}{rcl} N_F : B(x_*, r) & \to & X \\ x & \mapsto & x - F'(x)^{-1}F(x). \end{array} \tag{4.12}$$

Podemos aplicar a iteração de Newton para qualquer $x \in B(x_*, r)$ para obter $N_F(x)$ o qual pode não pertencer a $B(x_*, r)$, ou até mesmo não pertencer ao domínio de F. Assim, a Proposição 4.7 garante a boa definição de uma única iteração. Para assegurarmos que a iteração de Newton pode ser repetida infinitamente necessitaremos de algums resultados adicionais.

Lema 4.9. *Tomemos* $0 < t < r$. *Se* $\|x - x_*\| \le t$, *então*

$$\|N_F(x) - x_*\| \le \frac{|n_f(t)|}{t^2}\|x - x_*\|^2.$$

Demonstração. A inequação é trivial para $x = x_*$, desde que $F(x_*) = 0$. Agora assumamos que $0 < \|x - x_*\| \le t$. O Lema 4.7 implica que $F'(x)$ é não-singular. Deste modo, como $F(x_*) = 0$ resulta que

$$x_* - N_F(x) = -F'(x)^{-1}\left[F(x_*) - F(x) - F'(x)(x_* - x)\right] = -F'(x)^{-1}E_F(x, x_*).$$

Combinando a equação acima, com os Lemas 4.7 e 4.8, obtemos a estimativa

$$\|x_* - N_F(x)\| \leq \| - F'(x)^{-1}F'(x_*)\|\|F'(x_*)^{-1}E_F(x,x_*)\| \leq \frac{e_f(\|x-x_*\|,0)}{|f'(\|x-x_*\|)|}. \qquad (4.13)$$

Por outro lado, observando que $f(0)=0$, as definicões de e_f e n_f implicam em

$$e_f(\|x-x_*\|,0)/|f'(\|x-x_*\|)| = -n_f(\|x-x_*\|) = |n_f(\|x-x_*\|)|.$$

Como $\|x-x_*\| \leq t$, a Proposição 4.4 resulta em $|n_f(\|x-x_*\|)|/\|x-x_*\|^2 \leq |n_f(t)|/t^2$. Assim, a última equação ficará

$$\frac{e_f(\|x-x_*\|,0)}{|f'(\|x-x_*\|)|} \leq \frac{|n_f(t)|}{t^2}\|x-x_*\|^2.$$

Daquí, e de (4.13) segue-se a afirmação. □

Corolário 4.10. *Se $0<t<r$, então $N_F(B[x_*,t]) \subset B[x_*,|n_f(t)|]$. Além disso,*

$$N_F(B(x_*,r)) \subset B(x_*,r).$$

Demonstração. Tomemos $x \in B[x_*,t]$. Como $\|x-x_*\|/t \leq 1$, o Lema 4.9 implica que

$$\|N_F(x) - x_*\| \leq |n_f(t)|,$$

e segue-se a primeira inclusão. Já que $r \leq \rho$, a Proposição 4.5 implica que $|n_f(t)| < t$. Assim, a última inclusão é consequência imediata da primeira. □

4.1.4 Unicidade e bola ótima de convergência

Nesta subseção mostraremos a unicidade da solução e a bola ótima de convergência.

Lema 4.11. *Tomemos $0<t<\kappa$. Se $f(t)<0$, i. e., 0 é o único zero de f em $[0,t]$, então x_* é o único zero de F em $B[x_*,t]$. Como uma conseqüência x_* é o único zero de F em $B(x_*,\sigma)$.*

Demonstração. Assumamos que $F(y) = 0$ e $\|y - x_*\| \leq t$. Já que $F(x_*) = 0$ e $F(y) = 0$ temos que

$$y - x_* = -\int_0^1 F'(x_*)^{-1} \left[F'(x_* + u(y - x_*)) - F'(x_*)\right](y - x_*)du.$$

Usando (4.1) com $x = x_* + u(y - x_*)$ e $\tau = 0$ é fácil concluir da última equação que

$$\|y - x_*\| \leq \int_0^1 \left[f'(u\|y - x_*\|) - f'(0)\right] \|y - x_*\| du = f(\|y - x_*\|) - f(0) - f'(0)\|y - x_*\|.$$

Levando em conta que $f(0) = 0$ e $f'(0) = -1$, a última inequação fica

$$f(\|y - x_*\|) \geq 0.$$

Agora, como f é estritamente convexa e $f(t) < 0$ devemos ter $f < 0$ em $(0, t]$, i.e., 0 é o único zero de f em $[0, t]$. Daquí, e da inequação acima concluimos que $\|y - x_*\| = 0$, i.e., $y = x_*$. Assim, x_* é o único zero de F em $B[x_*, t]$. A segunda parte segue da definição da constante σ. □

Observação 4.12. *Notemos que, no lema acima usamos que a condição* (4.1) *é satisfeita apenas para* $\tau = 0$.

Lema 4.13. *Se* $f(\rho)/(\rho f'(\rho)) - 1 = 1$ *e* $\rho < \kappa$, *então* $r = \rho$ *é o melhor possível.*

Demonstração. Assumamos que $\rho < \kappa$. Definamos a função $h : (-\kappa, \kappa) \to \mathbb{R}$ por

$$h(t) = -f(-t), \qquad \text{para } t \in (-\kappa, 0], \qquad \text{e} \qquad h(t) = f(t), \qquad \text{para } t \in [0, \kappa). \tag{4.14}$$

É simples mostrar que $h(0) = 0$, $h'(0) \neq 0$ (de fato $h'(0) = f'(0) = -1$) e

$$\left|h'(0)^{-1}\left[h'(t) - h'(\tau t)\right]\right| \leq f'(|t|) - f'(\tau |t|),$$

para $\tau \in [0, 1]$ e $t \in (-\kappa, \kappa)$. Assim, $F = h$ satisfaz todas as hipóteses do Teorema 4.1. Deste modo, como $\rho < \kappa$, é suficiente mostrar que a aplicação do método de Newton para resolver $h(t) = 0$, com ponto inicial $t_0 = -\rho$ não converge. Já que $f(\rho)/(\rho f'(\rho)) - 1 = 1$, e da definição de h em (4.14) resulta que

$$t_1 = -\rho - h(-\rho)/h'(-\rho) = -\rho + f(\rho)/f'(\rho) = [f(\rho)/(\rho f'(\rho)) - 1]\rho = \rho.$$

De novo, da definição de h em (4.14) e supondo que $f(\rho)/(\rho f'(\rho)) - 1 = 1$ obtemos as igualdades

$$t_2 = \rho - h(\rho)/h'(\rho) = \rho - f(\rho)/f'(\rho) = -[f(\rho)/(\rho f'(\rho)) - 1]\rho = -\rho,$$

Portanto, o método de Newton para resolver $h(t) = 0$, com ponto inicial $t_0 = -\rho$, produz o ciclo

$$t_0 = -\rho, \quad t_1 = \rho, \quad t_2 = -\rho, \ \ldots .$$

Em particular este não converge, o qual prova o lema. □

4.1.5 Prova do Teorema 4.1

Primeramente notemos que a primeira equação em (4.2) junto com (4.12) implica que a seqüência $\{x_k\}$ satisfaz a relação

$$x_{k+1} = N_F(x_k), \qquad k = 0, 1, \ldots , \tag{4.15}$$

a qual é de fato uma definição equivalente desta seqüência.

Demonstração. Todas as afirmações envolvendo $\{t_k\}$ são provadas em Corolário 4.6.

Desde que $x_0 \in B(x_*, r)$ e $r \le \nu$, usando (4.15), a inclusão $N_F(B(x_*, r)) \subset B(x_*, r)$ do Corolário 4.10 e o Lema 4.7, concluímos que $\{x_k\}$ está bem definida e permanece em $B(x_*, r)$.

Para povar a convergência da seqüência $\{x_k\}$, primeiro devemos provar por indução que $\{t_k\}$ é uma seqüência majorante para $\{x_k\}$, i.e.,

$$\|x_* - x_k\| \le t_k, \qquad k = 0, 1, \ldots . \tag{4.16}$$

Como $t_0 = \|x_* - x_0\|$, a inequação acima vale para $k = 0$. Agora, suponhamos que $\|x_* - x_k\| \leq t_k$ para algum $k > 0$. Do Corolário 4.6 temos que $\{t_k\} \subset (0, r)$. Deste modo, usando o Corolário 4.10, as equações (4.7) e (4.15) concluímos que

$$\|x_* - x_{k+1}\| = \|x_* - N_F(x_k)\| \leq |n_f(t_k)| = t_{k+1},$$

o que prova que (4.16) vale. Como $\{t_k\}$ converge para 0, segue de (4.16) que $\{x_k\}$ converge para x_*.

Para provar a primeira inequação em (4.3), notemos que de (4.16) e do Lema 4.9 temos que

$$\|x_* - x_{k+1}\| = \|x_* - N_F(x_k)\| \leq \left[|n_f(t_k)|/t_k^2\right] \ \|x_* - x_k\|^2, \qquad k = 0, 1, \ldots.$$

Portanto, a desigualdade desejada segue de (4.7) (definição de $\{t_k\}$).

A unicidade é provada no Lema 4.11 e a última afirmação segue do Lema 4.13. □

4.2 Casos especiais

Nesta seção apresentaremos três casos especiais do Teorema 4.1, a saber, o teorema de convergência clássica sob condição Lipschitz, o teorema de Smale sob o método de Newton para funções analíticas e o teorema de Nesterov-Nemirovskii sob o método de Newton para funções auto-concordantes.

4.2.1 Resultado de Convergência sob a condição Lipschitz

O seguinte teorema foi publicado em Traub e Wozniakowski [24]

Teorema 4.14. *Sejam X, Y espaços de Banach, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função contínuamente diferenciável em Ω. Seja $x_* \in \Omega$ com $F'(x_*)$ não-singular. Suponhamos que $F(x_*) = 0$ e que exista uma constante $K > 0$ tal que*

$$\|F'(x_*)^{-1}\left[F'(x) - F'(y)\right]\| \leq K\|x - y\|, \qquad x,\, y \in \Omega. \tag{4.17}$$

Sejam $\kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}$ *e* $r := \min\{\kappa,\ 2/(3K)\}$. *Então, as seqüências com ponto inicial* $x_0 \in B(x_*, r)/\{x_*\}$ *e* $t_0 = \|x_* - x_0\|$, *respectivamente,*

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad t_{k+1} = \frac{Kt_k^2}{2(1-Kt_k)}, \qquad k = 0, 1, \ldots ,$$

estão bem definidas, $\{t_k\}$ *é decrescente, está contida em* $(0, r)$ *e converge para* 0, $\{x_k\}$ *está contida em* $B(x_*, r)$, *converge para o ponto* x_* *e valem as desigualdades*

$$\|x_* - x_{k+1}\| \le \frac{K}{2}\,\frac{1}{1-Kt_k}\,\|x_k - x_*\|^2 \le \frac{K}{2}\,\frac{1}{1-K\|x_0 - x_*\|}\,\|x_k - x_*\|^2, \qquad k = 0, 1, \ldots .$$

Além disso, o ponto x_* *é o único zero de* F *em* $B(x_*, 2/K)$ *e se* $2/(3K) < \kappa$, *então* $r = 2/(3K)$ *é o melhor raio de convergência possível.*

Demonstração. É imediato provar que F, x_* e $f : \mathbb{R} \to \mathbb{R}$ definida por $f(t) = Kt^2/2 - t$, satisfazem a desigualdade (4.1) e as condições **h1** e **h2** no Teorema 4.1. Neste caso, é fácil ver que as constantes ρ e ν definidas como no Teorema 4.1 satisfazem a relação

$$\rho = 2/(3K) \le \nu = 1/K,$$

como uma conseqüência $r := min\{\kappa, 2/(3K)\}$. Além disso, temos $f(\rho)/(\rho f'(\rho)) = 2$, $f(0) = f(2/K) = 0$ e $f(t) < 0$ para todo $t \in (0, 2/K)$. Também, notemos que a seqüência $\{t_k\}$ é equivalente a $t_{k+1} = |t_k - f(t_k)/f'(t_k)|$, para k= 0,1,2,... e satisfaz as relações

$$t_{k+1}/t_k^2 = \frac{1}{2/K - 2t_k} < \frac{1}{2/K - 2t_0} = \frac{K}{2}\,\frac{1}{1-K\|x_0 - x_*\|}, \qquad k = 0, 1, \ldots .$$

Assim, temos que F, f, r e x_* satisfazem todas as hipóteses do Teorema 4.1, tomando $x_0 \in B(x_*, r)/\{x_*\}$ e $t_0 = \|x_* - x_0\|$. Daí, todas as afirmações deste teorema seguem do Teorema 4.1 e da desigualdade acima. □

Observação 4.15. *Notemos que, como* $\|x_* - x_0\| < 2/(3K)$ *a última inequação no Teorema 4.14 implica que* $\|x_* - x_{k+1}\| \le [3K/2]\|x_k - x_*\|^2$, *para* $k = 0, 1, \ldots$ *Portanto, concluímos que*

$$\|x_* - x_k\| \le [2/(3K)]\left(3K\|x_0 - x_*\|/2\right)^{2^k}, \qquad k = 0, 1, \ldots ,$$

onde $3K\|x_0 - x_*\|/2 < 1$.

Observação 4.16. *Se a função* $F : \Omega \to Y$ *satisfaz a condição Lipschitziana clássica para derivadas, isto é,*

$$\|F'(x) - F'(y)\| \leq L\|x - y\|, \qquad x,\, y \in \Omega,$$

onde $L > 0$*, então também satisfaz a condição* (4.17) *com* $K = L\|F'(x_*)^{-1}\|$*. Neste caso, o raio de convergência para o método de Newton é* $r = 2/(3L\|F'(x_*)^{-1}\|)$*.*

4.2.2 Resultado de convergência sob condição de Smale

O seguinte resultado é o Corolário da Proposição 3 pp. 195 of Smale [23], ver também Proposição 1 pp. 157 e observação 1 pp. 158 de Blum, Cucker, Shub, e Smale [2].

Teorema 4.17. *Sejam* X*,* Y *espaços de Banach,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função analítica. Seja* $x_* \in \Omega$ *tal que* $F(x_*) = 0$ *e* $F'(x_*)$ *não-singular. Seja* $\kappa := \sup\{u > 0 : B(x_*, u) \subset \Omega\}$*,*

$$\gamma := \sup_{n>1} \left\| \frac{F'(x_*)^{-1}F^{(n)}(x_*)}{n!} \right\|^{1/(n-1)} \qquad e \qquad r := \min\left\{\kappa,\ \frac{5-\sqrt{17}}{4\gamma}\right\}. \tag{4.18}$$

Então, as seqüências com pontos iniciais $x_0 \in B(x_*, r)/\{x_*\}$ *e* $t_0 = \|x_* - x_0\|$*, respectivamente,*

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad t_{k+1} = \frac{\gamma\, t_k^2}{2[(1-\gamma t_k)^2 - 1]}, \qquad k = 0, 1, \ldots\,,$$

estão bem definidas, $\{t_k\}$ *é decrescente, está contida em* $(0, r)$ *e converge para* 0*,* $\{x_k\}$ *está contida em* $B(x_*, r)$*, converge para o ponto* x_* *e satisfaz as relações*

$$\|x_{k+1} - x_*\| \leq \frac{\gamma}{2(1-\gamma t_k)^2 - 1}\,\|x_k - x_*\|^2 \leq \frac{\gamma}{2(1-\gamma\|x_0 - x_*\|)^2 - 1}\,\|x_k - x_*\|^2, \quad k = 0, 1, \ldots.$$

Além disso, o ponto x_* *é o único zero de* F *em* $B(x_*, 1/(2\gamma))$ *e se* $(5 - \sqrt{17})/(4\gamma) < \kappa$*, então* $r = (5 - \sqrt{17})/(4\gamma)$ *é o melhor raio de convergência possível.*

Observação 4.18. *Notemos que, como* $\|x_* - x_0\| < (5-\sqrt{17})/(4\gamma)$ *a última inequação no Teorema 4.17 implica que* $\|x_* - x_{k+1}\| \leq [(4\gamma)/(5-\sqrt{17})]\|x_k - x_*\|^2$, *para* $k = 0, 1, \ldots ,$. *Portanto, concluímos que*

$$\|x_* - x_k\| \leq [(5-\sqrt{17})/(4\gamma)] \left((4\gamma)/(5-\sqrt{17})\|x_0 - x_*\|\right)^{2^k}, \qquad k = 0, 1, \ldots ,$$

onde $(4\gamma)/(5-\sqrt{17})\|x_0 - x_*\| < 1$.

Necessitaremos dos seguintes resultados para provar o teorema acima.

Lema 4.19. *Sejam* X, Y *espaços de Banach,* $\Omega \subseteq X$ *e* $F : \Omega \to Y$ *uma função analítica em* Ω. *Suponhamos que* $x_* \in \Omega$, $F'(x_*)$ *é não-singular e* $B(x_*, 1/\gamma) \subset \Omega$, *onde* γ *é definido em* (4.18). *Então para todo* $x \in B(x_*, 1/\gamma)$ *vale a desigualdade*

$$\|F'(x_*)^{-1}F''(x))\| \leqslant (2\gamma)/(1 - \gamma\|x - x_*\|)^3. \tag{4.19}$$

Demonstração. Seja $x \in \Omega$. Já que F é uma função analítica temos pela Observação 2.21 que F'' também é analítica e

$$F'(x_*)^{-1}F''(x) = \sum_{n=0}^{\infty} \frac{1}{n!} F'(x_*)^{-1}F^{(n+2)}(x_*)(x - x_*)^n.$$

Combinando (4.18) e a equação acima obtemos, após simples manipulaçẽs algébricas que

$$\|F'(x_*)^{-1}F''(x)\| \leqslant \gamma \sum_{n=0}^{\infty}(n+2)(n+1)(\gamma||x - x_*||)^n.$$

Por outro lado, como $B(x_*, 1/\gamma) \subset \Omega$ obtemos que $\gamma\|x - x_*\| < 1$. Assim, do Lema 2.22 concluímos que

$$\frac{2}{(1 - \gamma\|x - x_*\|)^3} = \sum_{n=0}^{\infty}(n+2)(n+1)(\gamma||x - x_*||)^n.$$

Portanto, combinando as duas equações acima obtemos o resultado desejado. $\square$

O seguinte resultado fornece uma condição mais fácil de verificar do que a condição (4.1), quando as funções em consideração são duas vezes contínuamente diferenciáveis.

Lema 4.20. *Seja X, Y espaços de Banach, $\Omega \subseteq X$ e $F : \Omega \to Y$ uma função contínua, duas vezes contínuamente diferenciável em Ω. Sejam $x_* \in \Omega$ com $F'(x_*)$ não-singular. Se existe $f : [0, R) \to \mathbb{R}$ uma função, duas vezes contínuamente diferenciável e satisfazendo a condição*

$$\|F'(x_*)^{-1}F''(x)\| \leqslant f''(\|x - x_*\|), \tag{4.20}$$

para todo $x \in \Omega$ tal que $\|x - x_*\| < R$. Então F e f satisfazem (4.1).

Demonstração. Tomando $\tau \in [0, 1]$ e $x \in \Omega$, tal que $x_* + \tau(x - x_*) \in \Omega$ e $\|x - x_*\| < R$, obtemos que

$$\|F'(x_*)^{-1}\left[F'(x) - F'(x_* + \tau(x - x_*))\right]\| \leq \int_\tau^1 \|F'(x_*)^{-1}F''(x_* + t(x - x_*))\|\,\|x - x_*\|dt.$$

Agora, como $\|x - x_*\| < R$ e f satisfaz em (4.20), obtemos da última inequação que

$$\|F'(x_*)^{-1}\left[F'(x) - F'(x_* + \tau(x - x_*))\right]\| \leq \int_\tau^1 f''(t\|x - x_*\|)\|x - x_*\|dt.$$

Avaliando a última integral segue a afirmação. □

[Prova do Teorema 4.17]. Assumamos que todas as hipóteses do Teorema 4.17 valem. Consideremos a função real $f : [0, 1/\gamma) \to \mathbb{R}$ definida por

$$f(t) = \frac{t}{1 - \gamma t} - 2t.$$

É simples mostrar que f é analítica e

$$f(0) = 0, \quad f'(t) = 1/(1-\gamma t)^2 - 2, \quad f'(0) = -1, \quad f''(t) = (2\gamma)/(1-\gamma t)^3, \quad f^{(n)}(0) = n!\,\gamma^{n-1},$$

para $n \geq 2$. Das quatro últimas equações é fácil concluir que f satisfaz **h1** e **h2**. Agora, combinando o Lema (4.20),com o Lema (4.19) e última equação implica que F e f satisfazem (4.1). Finalmente, notemos que neste caso, as constantes ν e ρ definidas como no Teorema 4.1 satisfazem as relações

$$\rho = \frac{5 - \sqrt{17}}{4\gamma} < \nu = \frac{\sqrt{2} - 1}{\sqrt{2}\gamma} < \frac{1}{\gamma},$$

assim, $r := min\{\kappa, (5-\sqrt{17})/(4\gamma)\}$. Além disso, $f(\rho)/(\rho f'(\rho)) = 2$, $f(0) = f(1/(2\gamma)) = 0$ e $f(t) < 0$ para todo $t \in (0, 1/(2\gamma))$. Também, notemos que a seqüência $\{t_k\}$ é equivalente a $t_{k+1} = |t_k - f(t_k)/f'(t_k)|$, para k=0,1,... e satisfaz as relações

$$t_{k+1}/t_k^2 = \frac{\gamma}{2(1-\gamma t_k)^2 - 1} < \frac{\gamma}{2(1-\gamma t_0)^2 - 1} = \frac{\gamma}{2(1-\gamma\|x_0 - x_*\|)^2 - 1}, \qquad k = 0, 1, \ldots.$$

Portanto, podemos afirmar que F, f, e x_* satisfazem todas as hipóteses do Teorema 4.1, tomando $x_0 \in B(x_*, r)/\{x_*\}$ e $t_0 = \|x_* - x_0\|$. Daí, as afirmações deste teorema seguem do Teorema 4.1 e da inequação acima. □

4.2.3 Resultado de convergência sob a condição de Nesterov-Nemirovskii

Teorema 4.21. *Sejam X um espaço de Banach, $\Omega \subseteq X$ um conjunto convexo e $g : \Omega \to \mathbb{R}$ uma função a-auto-concordante. Seja $x_* \in \Omega$ com $g''(x_*)$ não-singular. Suponhamos que $g'(x_*) = 0$. Sejam as constantes $\kappa := \sup\{u > 0 : W_u(x_*) \subset \Omega\}$ e*

$$r := \min\left\{\kappa,\ \frac{5-\sqrt{17}}{4}\right\}.$$

Então, as seqüências com pontos iniciais $x_0 \in B(x_*, r)/\{x_*\}$ e $t_0 = \|x_* - x_0\|$, respectivamente,

$$x_{k+1} = x_k - g''(x_k)^{-1}g'(x_k), \qquad t_{k+1} = \frac{t_k^2}{2(1-t_k)^2 - 1}, \qquad k = 0, 1, \ldots,$$

estão bem definidas, $\{t_k\}$ é decrescente, está contida em $(0, r)$ e converge para 0, $\{x_k\}$ está contida em $B(x_*, r)$, converge para o ponto x_* e valem as desigualdades

$$\|x_{k+1} - x_*\| \le \frac{1}{2(1-t_k)^2 - 1}\|x_k - x_*\|^2 \le \frac{1}{2(1-\|x_0 - x_*\|)^2 - 1}\|x_k - x_*\|^2, \qquad k = 0, 1, \ldots.$$

Observação 4.22. *Notemos que, como $\|x_* - x_0\| < (5-\sqrt{17})/4$ a última inequação no Teorema 4.21 implica que $\|x_* - x_{k+1}\| \le [4/(5-\sqrt{17})]\|x_k - x_*\|^2$, para $k = 0, 1, \ldots,$. Portanto, concluímos que*

$$\|x_* - x_k\| \le [(5-\sqrt{17})/4]\left(4/(5-\sqrt{17})\|x_0 - x_*\|\right)^{2^k}, \qquad k = 0, 1, \ldots,$$

onde $4/(5-\sqrt{17})\|x_0 - x_*\| < 1.$

Para provar o teorema serão necessários os resultados de funções auto-concordantes devido a Nesterov, Y. e Nemirovskii, A. [17] dadas na seção 2.4.
Assumamos que todas as hipóteses do Teorema 4.21 valem. Consideremos a função real $f : [0, 1) \to \mathbb{R}$ definida por

$$f(t) = \frac{t}{1-t} - 2t.$$

É fácil ver que f satisfaz as igualdades

$$f(0) = 0, \quad f'(t) = 1/(1-t)^2 - 2, \quad f'(0) = -1, \quad f''(t) = 2/(1-t)^3 > 0.$$

Das quatro últimas igualdades podemos concluir que f satisfaz **h1** e **h2**. Devido ao Lema 2.28, e a última igualdade obtemos a estimativa

$$\|g''(x_*)^{-1}g'''(x)\| \leq \frac{2}{(1-\|x-x_*\|)^3} = f''(\|x - x_*\|)$$

Fazendo $F = g'$ no Lema 4.20 e considerando a última inequação temos que g' e f satisfazem (4.1). Finalmente, notemos que neste caso, as constantes ν e ρ definidas como no Teorema 4.1 satisfazem as relações

$$\rho = \frac{5-\sqrt{17}}{4} < \nu = \frac{\sqrt{2}-1}{\sqrt{2}} < 1,$$

assim, $r := min\{\kappa, (5-\sqrt{17})/4\}$. Além disso, $f(\rho)/(\rho f'(\rho)) = 2$, $f(0) = f(1/2) = 0$ e $f(t) < 0$ para todo $t \in (0, 1/2)$. Também, notemos que a seqüência $\{t_k\}$ é equivalente a $t_{k+1} = |t_k - f(t_k)/f'(t_k)|$, para $k = 0, 1, 2, \ldots$ e satisfaz as relações

$$t_{k+1}/t_k^2 = \frac{1}{2(1-t_k)^2 - 1} < \frac{1}{2(1-t_0)^2 - 1} = \frac{1}{2(1-\|x_0 - x_*\|)^2 - 1}, \qquad k = 0, 1, \ldots.$$

Portanto, podemos afirmar que $F = g'$, f, e x_* satisfazem todas as hipóteses do Teorema 4.1, tomando $x_0 \in B(x_*, r)/\{x_*\}$ e $t_0 = \|x_* - x_0\|$. Daí, todas as afirmações deste teorema seguem do Teorema 4.1 e da desigualdade acima. □

Observação 4.23. *Nas afirmações do Teorema 4.21 não incluímos a otimalidade do raio de convergência porque, neste caso, sua função majorante* $f(t) = t/(1-t) - 2t$ *não é auto-concordante.*

Capítulo 5

Análise semi-local do método de Newton

5.1 Introdução

Nosso objetivo neste capítulo é desenvolver uma análise de convergência semi-local para o método de Newton, conhecida como teorema de Kantorovich. A análise de convergência apresentada será baseada no princípio majorante de Kantorovich. Diferentemente das outras, esta análise deixa claro a relação entre a função majorante e o operador não linear em consideração. Ela também garante existência e unicidade de solução e taxa de convergência Q-quadrática, onde para o último caso, não é necessário a existência de uma segunda raiz para a função majorante, necessita apenas que a função seja definida até sua primeira raiz. Finalmente, obteremos uma estimativa da taxa de convergência baseada na derivada direcional *da derivada* da função majorante. Por último, vários resultados sobre o Método de Newton aparentemente sem relação serão unificados.

Na análise apresentada, em vez de olhar só a seqüência gerada, nós identificaremos regiões onde o método de Newton é bem comportado, quando é comparado com o método de Newton aplicado a função majorante. Esta análise foi introduzida em Ferreira e Svaiter [5], para generalizar o teorema de Kantorovich sobre o método de Newton para variedades Riemannianas, e que tem sido usadas por Alvarez, Botle e Munier [1], Li e Wang [13] e Wang e Li [25] no mesmo contexto.

5.2 Teorema de Kantorovich

Nosso trabalho agora é em direção da explicação e prova do teorema de Kantorovich sobre o método de Newton. A primeira coisa que devemos fazer é provar que este teorema vale para uma função real majorante. Então, provaremos a boa definição e convergência do método de Newton. Além disso, mostraremos a unicidade da solução em uma adequada região e serão estabelecidas suas taxas de convergência. A referência bibliográfica deste capítulo foi baseada em Ferreira e Svaiter [6]. A afirmação do teorema é:

Teorema 5.1. *Sejam X, Y espaços de Banach, $C \subseteq X$ e $F : C \to Y$ uma função contínua, contínuamente diferenciável no int(C). Tomemos $x_0 \in \operatorname{int}(C)$ com $F'(x_0)$ não-singular. Suponhamos que existem $R > 0$ e uma função continuamente diferenciável $f : [0, R) \to \mathbb{R}$ tal que, $B(x_0, R) \subseteq C$,*

$$\|F'(x_0)^{-1} \left[F'(y) - F'(x)\right] \| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|), \tag{5.1}$$

para $x, y \in B(x_0, R)$, $\|x - x_0\| + \|y - x\| < R$,

$$\|F'(x_0)^{-1} F(x_0)\| \leq f(0)\,, \tag{5.2}$$

e

h1) *$f(0) > 0$, $f'(0) = -1$;*

h2) *f' é convexa e crescente;*

h3) *$f(t) = 0$ para algum $t \in (0, R)$.*

Então f tem uma menor raiz $t_* \in (0, R)$, as seqüências geradas pelo método de Newton para resolver $f(t) = 0$ e $F(x) = 0$ com ponto inicial $t_0 = 0$ e x_0, respectivamente,

$$t_{k+1} = t_k - f'(t_k)^{-1} f(t_k), \qquad x_{k+1} = x_k - F'(x_k)^{-1} F(x_k), \qquad k = 0, 1, \dots\,, \tag{5.3}$$

estão bem definidas, $\{t_k\}$ *é crescente, está contida em* $[0, t_*)$ *e converge para* t_**, a seqüência* $\{x_k\}$ *está contida em* $B(x_0, t_*)$ *e converge para um ponto* $x_* \in B[x_0, t_*]$ *o qual é o único zero de* F *em* $B[x_0, t_*]$*; valem as estimativas*

$$\|x_* - x_k\| \leq |t_* - t_k|, \qquad \|x_* - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2}\,\|x_* - x_k\|^2, \qquad k = 0, 1, \ldots, \tag{5.4}$$

e as seqüências $\{t_k\}$ *e* $\{x_k\}$ *convergem* Q*-linearmente como segue*

$$\|x_* - x_{k+1}\| \leq \frac{1}{2}\|x_* - x_k\|, \qquad t_* - t_{k+1} \leq \frac{1}{2}(t_* - t_k) \qquad k = 0, 1, \ldots. \tag{5.5}$$

Se, adicionalmente,

h4) $f'(t_*) < 0$,

então as seqüências $\{t_k\}$ *e* $\{x_k\}$ *convergem* Q*-quadráticamente como segue*

$$\|x_* - x_{k+1}\| \leqslant \frac{D_-f'(t_*)}{-2f'(t_*)}\|x_* - x_k\|^2, \qquad t_* - t_{k+1} \leq \frac{D_-f'(t_*)}{-2f'(t_*)}(t_* - t_k)^2, \quad k = 0, 1, \ldots, \tag{5.6}$$

e x_* *é o único zero de* F *em* $B(x_0, \bar{\tau})$*, onde* $\bar{\tau} > t_*$ *é definido como*

$$\bar{\tau} = \sup\{t \in [t_*, R) \; : \; f(t) \leq 0\}.$$

Observação 5.2. *Sob as hipóteses* **h1**-**h3** *do Teorema 5.1 sobre* $f : [0, R) \to \mathbb{R}$,

i. $f(t) = 0$ *tem no máximo uma raiz em* (t_*, R)*;*

ii. A condição **h4** *do Teorema 5.1 é implicada por qualquer uma das seguintes condições sobre* f*:*

h4-a) $f(t_{**}) = 0$ *para algum* $t_{**} \in (t_*, R)$,

h4-b) $f(t) < 0$ *para algum* $t \in (t_*, R)$,

onde t_* *é a menor raiz de* f *em* $[0, R)$.

Na versão usual do teorema de Kantorovich, para garantir convergência R-quadrática da seqüência $\{x_k\}$ e $\{t_k\}$, a condição **h4-a** é usada. Como discutiremos, esta condição é mais restritiva que a condição **h4** do Teorema 5.1.

Daqui em diante, assumiremos que as hipóteses do Teorema 5.1 são válidas, a exceção de **h4**, a qual será considerada válida só quando seja estabelecido explicitamente.

5.2.1 Método de Newton aplicado à função majorante

Nesta subseção estudaremos a função majorante f e provaremos todos os resultados com respeito à seqüência $\{t_k\}$.

Proposição 5.3. *A função f tem uma menor raiz $t_* \in (0, R)$, é estritamente convexa, e*

$$f(t) > 0, \quad f'(t) < 0, \qquad t < t - f(t)/f'(t) < t_*, \qquad\qquad \forall\, t \in [0, t_*). \tag{5.7}$$

Além disso, $f'(t_*) \le 0$ e

$$f'(t_*) < 0 \iff \exists\, t \in (t_*, R),\ f(t) \le 0. \tag{5.8}$$

Demonstração. Como f é contínua em $[0, R)$ e tem um zero neste intervalo **h3**, esta deve ter um menor zero t_*, o qual é maior que 0 porque $f(0) > 0$ **h1** do Teorema 5.1. Desde que f' é crescente **h2** do Teorema 5.1, f é estritamente convexa. A primeira desigualdade em (5.7) segue-se da suposição $f(0) > 0$ e da definição de t_*. Desde que f é estritamente convexa,

$$0 = f(t_*) > f(t) + f'(t)(t_* - t), \qquad t \in [0, R),\ t \ne t_*. \tag{5.9}$$

Se $t \in [0, t_*)$, então $f(t) > 0$ e $t_* - t > 0$, o qual, combinado com (5.9) resulta a segunda desigualdade em (5.7). A terceira inequação em (5.7) segue-se da primeira e da segunda desigualdade. A última desigualdade em (5.7) é obtida por divisão da inequação em (5.9) por $-f'(t)$ (a qual é estritamente positiva) e manipulações algébricas.

Como $f > 0$ em $[0, t_*)$ e $f(t_*) = 0$, devemos ter $f'(t_*) \leq 0$. Em (5.8), a implicação $\Rightarrow$ verifica-se trivialmente. Para provar a implicação $\Leftarrow$, troquemos t e t_* em (5.9) e notemos que $f(t_*) = 0$. $\square$

Em vista da primeira desigualdade em (5.7), a iteração de Newton para a função f está bem definida em $[0, t_*)$ e denotaremos por:

$$\begin{array}{rcl} n_f : [0, t_*) & \to & \mathbb{R} \\ t & \mapsto & t - f(t)/f'(t). \end{array} \tag{5.10}$$

Proposição 5.4. *A iteração de Newton n_f, aplica $[0, t_*)$ em $[0, t_*)$, e*

$$t_* - n_f(t) \leqslant \frac{1}{2}(t_* - t), \qquad \forall\, t \in [0, t_*). \tag{5.11}$$

Se f também satisfaz **h4**, *isto é, $f'(t_*) < 0$, então*

$$t_* - n_f(t) \leq \frac{D_- f'(t_*)}{-2f'(t_*)}(t_* - t)^2, \qquad \forall\, t \in [0, t_*). \tag{5.12}$$

Demonstração. A primeira afirmação segue-se trivialmente das últimas inequações em (5.7).

Para provar (5.11) tomemos algum $t \in [0, t_*)$. Notemos que $f(t_*) = 0$ (Prop. 5.3). Usando também (5.10) e a continuidade de f' temos as igualdades

$$\begin{aligned} t_* - n_f(t) &= \frac{1}{f'(t)}\big[f'(t)(t_* - t) + f(t)\big] \\ &= \frac{1}{f'(t)}\big[f'(t)(t_* - t) + f(t) - f(t_*)\big] = \frac{1}{-f'(t)} \int_t^{t_*} f'(u) - f'(t)\, du. \end{aligned}$$

Como f' é convexa e $t < t_*$, segue-se da Proposição 2.10 que

$$f'(u) - f'(t) \leqslant [f'(t_*) - f'(t)]\, \frac{u - t}{t_* - t}, \qquad \forall\, u \in [t, t_*].$$

Levando em conta a positividade de $-1/f'(t)$ (segunda desigualdade em (5.7)) e combinando as duas equações acima temos a estimativa

$$t_* - n_f(t) \leqslant (-1/f'(t)) \int_t^{t_*} [f'(t_*) - f'(t)]\, \frac{u - t}{t_* - t}\, du.$$

por integração direta do último termo da inequação acima resulta a estimativa

$$t_* - n_f(t) \leqslant \frac{1}{2}\left(\frac{f'(t_*) - f'(t)}{-f'(t)}\right)(t_* - t). \tag{5.13}$$

Portanto, considerando a desigualdade acima junto com $f'(t_*) \leq 0$ e $f'(t) < 0$ implica em (5.11).

Finalmente, assumiremos que f satisfaz a hipótese **h4**. Consideremos $t \in [0, t_*)$, como f' é crescente, $f'(t_*) < 0$ e $f'(t) < 0$, obtemos as relações

$$\frac{f'(t_*) - f'(t)}{-f'(t)} \leq \frac{f'(t_*) - f'(t)}{-f'(t_*)} = \frac{1}{-f'(t_*)}\frac{f'(t_*) - f'(t)}{t_* - t}(t_* - t) \leq \frac{D_- f'(t_*)}{-f'(t_*)}(t_* - t),$$

onde a última inequação segue da Proposição 2.9 item 1. Combinando a inequação acima com (5.13) concluímos que (5.12) vale. □

A definição de $\{t_k\}$ no Teorema 5.1 é equivalente a:

$$t_0 = 0, \quad t_{k+1} = n_f(t_k), \qquad k = 0, 1, \ldots . \tag{5.14}$$

Corolário 5.5. *A seqüência $\{t_k\}$ está bem definida, é crescente e está contida em $[0, t_*)$. Além disso, satisfaz a desigualdade* (5.5) *(segunda inequação) e converge Q-linearmente para t_*. Se f também satisfaz a hipótese* **h4**, *então $\{t_k\}$ satisfaz a segunda desigualdade em* (5.6) *e converge Q-quadraticamente.*

Demonstração. Primeiro mostraremos que $t_k \in [0, t_*)$ para todo k. De fato, já que $t_0 = 0$ implica que $t_0 \in [0, t_*)$. Suponhamos agora que $t_k \in [0, t_*)$, para algum $k > 0$. Da Proposição 5.3 e Proposição 5.4 temos que, $0 \leq t_k < n_f(t_k) = t_{k+1} < t_*$, o que implica que $t_{k+1} \in [0, t_*)$ e além disso que $\{t_k\}$ é crescente. Também devido a que $f'(t) < 0$ em $[0, t_*)$ (Proposição 5.3), podemos afirmar que $\{t_k\}$ está bem definida. As demais afirmações deste corolário segue-se diretamente da Proposição 5.4 □

Assim, todas as afirmações envolvendo $\{t_k\}$ no Teorema 5.1 são válidas.

5.2.2 Convergência

Nesta subseção, provaremos que a seqüência $\{x_k\}$ gerada pelo método de Newton para resolver o sistema não-linear $F(x) = 0$ com ponto inicial x_0, está bem definida e converge para uma solução do sistema em consideração.

Proposição 5.6. *Se* $\|x - x_0\| \leq t < t_*$, *então* $F'(x)$ *é não-singular e*

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

Em particular, F' *é não-singular em* $B(x_0, t_*)$.

Demonstração. Faça $\overline{x} = x_0$ no Lema 2.18, pois $\|x - x_*\| < t < t_*$ e $f'(t) < 0$ para $0 \leq t < t_*$. □

A iteração de Newton produz pontos que são os zero da linearização de F, tais pontos, também são a primeira-ordem da expansão de Taylor para F. Assim, estudaremos o erro na linearização de F num ponto de $B(x_0,\, t)$ dada por

$$E(x, y) := F(y) - [F(x) + F'(x)(y - x)], \qquad y \in C, x \in B(x_0,\, R). \tag{5.15}$$

Limitaremos este erro pelo erro na linerização da função majorante f, dada por

$$e(t, v) := f(v) - [f(t) + f'(t)(v - t)], \qquad t, v \in [0, R). \tag{5.16}$$

Lema 5.7. *Seja*

$$x, y \in B(x_0, R) \quad e \quad 0 \leq t < v < R.$$

Se $\|x - x_0\| \leqslant t$ *e* $\|y - x\| \leqslant v - t$, *então*

$$\|F'(x_0)^{-1}E(x, y)\| \leqslant e(t, v)\frac{\|y - x\|^2}{(v - t)^2}.$$

Demonstração. Como $x, y \in B(x_0, R)$ e a bola é convexa, então

$$x + u(y - x) \in B(x_0, R) \quad \text{para} \quad 0 \leq u \leq 1.$$

Portanto, como F é continuamente diferenciável em $B(x_0, R)$, a equação (5.15) é equivalente a

$$E(x,y) = \int_0^1 [F'(x+u(y-x)) - F'(x)](y-x)\, du.$$

Combinando a igualdade acima com a hipótese em (5.1) resulta nas desigualdades

$$\begin{aligned} \|F'(x_0)^{-1}E(x,y)\| &\leq \int_0^1 \left\|F'(x_0)^{-1}[F'(x+u(y-x)) - F'(x)]\right\| \, \|y-x\| \, du \\ &\leq \int_0^1 \left[f'\left(\|x-x_0\| + u\,\|y-x\|\right) - f'\left(\|x-x_0\|\right)\right] \|y-x\|\, du\,. \end{aligned}$$

Agora, usando a convexidade de f', as hipóteses $\|x-x_0\| \leq t$, $\|y-x\| \leq v-t$, $v < R$, Corolário 2.8 e a Proposição 2.10 temos que, para qualquer $u \in [0,1]$.

$$\begin{aligned} f'\left(\|x-x_0\| + u\|y-x\|\right) - f'\left(\|x-x_0\|\right) &\leq f'\left(t + u\|y-x\|\right) - f'\left(t\right) \\ &\leq [f'(t+u(v-t)) - f'(t)]\frac{\|y-x\|}{v-t}\,. \end{aligned}$$

Novamente, combinando as duas últimas inequações, obtemos a estimativa

$$\|F'(x_0)^{-1}E(x,y)\| \leq \int_0^1 [f'(t+u(v-t)) - f'(t)]\frac{\|y-x\|^2}{v-t}\, du.$$

Finalmente, calculando a integral acima, obtemos o resultado desejado. □

Pela Proposição 5.6 garantimos a não-singularidade de F' em $B(x_0, t_*)$, assim podemos definir a aplicação iteração de Newton para resolver $F(x) = 0$. Chamaremos N_F à aplicação iteração de Newton (para F) nesta região, isto é,

$$\begin{array}{rcl} N_F : B(x_0, t_*) & \to & X \\ x & \mapsto & x - F'(x)^{-1}F(x). \end{array} \tag{5.17}$$

Podemos aplicar a iteração de Newton em qualquer $x \in B(x_0, t_*)$ e obter $N_F(x)$ o qual pode não pertencer a $B(x_0, t_*)$, ou até mesmo não pertencer ao domínio de F. Assim, a Proposição 5.6 garante a boa definição de uma única iteração. Para assegurarmos que a iteração de Newton pode ser repetida infinitamente, necessitaremos de algums resultados adicionais.

Primeiramente, definiremos alguns subconjuntos de $B(x_0, t_*)$. Mostraremos, que as iterações de Newton (5.17) são "bem comportadas"nestes subconjuntos. Sejam

$$K(t) := \left\{ x \in B[x_0, t] \ : \ \|F'(x)^{-1}F(x)\| \leqslant -\frac{f(t)}{f'(t)} \right\}, \qquad t \in [0, t_*) \quad e \tag{5.18}$$

$$K := \bigcup_{t \in [0, t_*)} K(t). \tag{5.19}$$

Como $0 \leqslant t < t_*$ em (5.18), temos, $f'(t) \neq 0$ e segue da Proposição 5.6 que F' é não-singular na $B[x_0, t] \subset B(x_0, t_*)$. Assim, as definições são consistentes.

Lema 5.8. *Se $t \in [0, t_*)$, então $K(t) \subset B(x_0, t_*)$ e*

$$N_F\left(K(t)\right) \subset K\left(n_f(t)\right).$$

Além disso, $K \subset B(x_0, t_*)$ e $N_F(K) \subset K$.

Demonstração. A primeira inclusão segue-se trivialmente da definição de $K(t)$. Para provar que $N_F(K(t)) \subset K(n_f(t))$ devemos mostrar que para qualquer $x \in K(t)$, valem as seguintes desigualdades:

$$\|N_F(x) - x_0\| \leq n_f(t) \quad e \quad \left\|F'(N_F(x))^{-1}F(N_F(x))\right\| \leq -\frac{f(n_f(t))}{f'(n_f(t))}. \tag{5.20}$$

Para mostrar a primeira desigualdade em (5.20), tomemos $t \in [0, t_*)$ e $x \in K(t)$. Agora, usando a definição em (5.18) e a primeira afirmação na Proposição 5.4 temos que

$$\|x - x_0\| \leq t, \qquad \|F'(x)^{-1}F(x)\| \leq -\frac{f(t)}{f'(t)}, \quad t < n_f(t) < t_*. \tag{5.21}$$

Portanto

$$\begin{aligned}\|N_F(x) - x_0\| &\leqslant \|x - x_0\| + \|N_F(x) - x\| = \|x - x_0\| + \|F'(x)^{-1}F(x)\| \\ &\leqslant t - f(t)/f'(t) = n_f(t) < t_*\end{aligned}$$

e

$$N_F(x) \in B[x_0, n_f(t)] \subset B(x_0, t_*). \tag{5.22}$$

Para mostrar a segunda desigualdade em (5.20), primeiro notemos que $N_F(x)$, $n_f(t)$ pertence aos domínios de F e f, respectivamente. Daí usando as definições da iteração de Newton em (5.10), (5.17) e erros de linearização em (5.15) e (5.16), obtemos as igualdades

$$\begin{aligned} f(n_f(t)) &= f(n_f(t)) - [f(t) + f'(t)(n_f(t) - t)] \\ &= e(t, n_f(t)) \end{aligned}$$

e

$$\begin{aligned} F(N_F(x)) &= F(N_F(x)) - [F(x) + F'(x)(N_F(x) - x)] \\ &= E(x, N_F(x))\,. \end{aligned}$$

Logo, das duas últimas equações, de (5.21) e do Lema 5.7, temos as relações

$$\begin{aligned} \|F'(x_0)^{-1}F(N_F(x))\| &= \|F'(x_0)^{-1}E(x, N_F(x))\| \\ &\leq e(t, n_f(t))\frac{\|F'(x)^{-1}F(x)\|^2}{(f(t)/f'(t))^2} \leq e(t, n_f(t)) = f(n_f(t)). \end{aligned}$$

Como $\|N_F(x) - x_0\| \leq n_f(t)$, segue-se da Proposição 5.6 que $F'(N_F(x))$ é não-singular e

$$\|F'(N_F(x))^{-1}F'(x_0)\| \leq -1/f'(n_f(t)).$$

Combinando as duas inequações acima concluímos que

$$\begin{aligned} \|F'(N_F(x))^{-1}F(N_F(x))\| &\leqslant \|F'(N_F(x))^{-1}F'(x_0)\|\,\|F'(x_0)^{-1}F(N_F(x))\| \\ &\leqslant -f(n_f(t))/f'(n_f(t)). \end{aligned}$$

Este resultado, junto com (5.22) prova que $N_F(x) \in K(n_f(t))$, portanto vale a segunda inclusão deste lema.

A seguinte inclusão (primeira em a segunda sentença), segue-se trivialmente das definições em (5.18) e (5.19). Para verificar a última inclusão, tomemos $x \in K$. Então $x \in K(t)$ para algum $t \in [0, t_*)$. Usando a primeira parte deste lema, temos que $N_F(x) \in K(n_f(t))$. Para finalizar a prova, notemos que $n_f(t) \in [0, t_*)$ e usemos a definição de K. $\square$

Finalmente, estamos pronto para provar o resultado principal desta seção o qual é uma conseqüência imediata do último resultado. Primeiramente notemos que a seqüência $\{x_k\}$ (veja (5.3)) satisfaz a igualdade

$$x_{k+1} = N_F(x_k), \qquad k = 0, 1, \dots, \tag{5.23}$$

a qual é de fato uma definição equivalente desta seqüência.

Corolário 5.9. *A seqüência $\{x_k\}$ está bem definida, está contida em $B(x_0, t_*)$, converge para o ponto $x_* \in B[x_0, t_*]$,*

$$\|x_* - x_k\| \leq t_* - t_k, \qquad k = 0, 1, \dots,$$

e $F(x_*) = 0$.

Demonstração. De (5.2) e da hipótese **h1** do Teorema 5.1, temos que

$$x_0 \in K(0) \subset K,$$

onde a inclusão segue trivialmente de (5.19). Usando (5.23), as inclusões $N_F(K) \subset K$ e Lema 5.8, concluímos que a seqüência $\{x_k\}$ está bem definida e permanece em K. Da primeira inclusão da segunda parte do Lema 5.8 temos trivialmente que $\{x_k\}$ está contida em $B(x_0, t_*)$.

Provaremos, por indução que

$$x_k \in K(t_k), \qquad k = 0, 1, \dots. \tag{5.24}$$

A inclusão acima, mostra que (5.24) vale para $k = 0$. Assumamos agora que $x_k \in K(t_k)$, para algum $k > 0$. Assim, usando o Lema 5.8, (5.23) e (5.14) temos que $x_{k+1} \in K(t_{k+1})$, a qual completa a prova de (5.24).

Agora, usando (5.24) e (5.18), temos que

$$\|F'(x_k)^{-1}F(x_k)\| \leq -f(t_k)/f'(t_k), \qquad k = 0, 1, \dots,$$

a qual, por (5.3), é equivalente a

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \qquad k = 0, 1, \ldots . \tag{5.25}$$

Daqui, e do Lema 2.13 podemos afirmar que $\{x_k\}$ é uma seqüência de Cauchy em $B(x_0, t_*)$ e assim, converge para algum $x_* \in B[x_0, t_*]$. Além disso temos que $\|x_* - x_k\| \leq t_* - t_k$, para todo k.

Resta provar que $F(x_*) = 0$. Primeiro, notemos que

$$\begin{aligned}\|F'(x_k)\| &\leqslant \|F'(x_0)\| + \|F'(x_k) - F'(x_0)\| \\ &\leq \|F'(x_0)\| + \|F'(x_0)\|\|F'(x_0)^{-1}\left[F'(x_k) - F'(x_0)\right]\|.\end{aligned}$$

Já que $\|x_k - x_0\| \leq t_k$ e $t_k < t_* < R$ temos que

$$\|F'(x_0)^{-1}\left[F'(x_k) - F'(x_0)\right]\| \leq f'(\|x_k - x_0\|) - f'(0) \leq f'(t_*) - f'(0).$$

Combinando as duas inequações acima temos que $\{\|F'(x_k)\|\}$ é limitada. Por outro lado, segue-se de (5.3) e (5.25) que

$$\begin{aligned}\|F(x_k)\| &\leqslant \|F'(x_k)\|\,\|F'(x_k)^{-1}F(x_k)\| \\ &\leqslant \|F'(x_k)\|\,\left(t_{k+1} - t_k\right).\end{aligned}$$

Pelo fato que $\{\|F'(x_k)\|\}$ é limitada e $\{t_k\}$ converge, podemos tomar limite na última inequação para concluir que

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

Finalmente, a continuidade de F em $B[x_0, t_*]$, $\{x_k\} \subset B(x_0, t_*)$ e $\{x_k\}$ converge para x_*, implicam que

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

□

5.2.3 Unicidade e taxa de Convergência

Provamos até agora, que a seqüência $\{x_k\}$ converge para uma solução x_* de $F(x) = 0$ e $x_* \in B[x_0, t_*]$. Assim, provaremos que esta convergência para x_* é pelo menos Q-linearmente e x_* é a única solução de $F(x) = 0$ na região $B[x_0, t_*]$. Além disso, considerando que f satisfaz a hipótese **h4**, também provaremos que $\{x_k\}$ converge Q-quadraticamente para x_* e a região de unicidade cresce de $B[x_0, t_*]$ para $B(x_0, \bar{\tau})$. Os resultados serão obtidos como uma conseqüência do seguinte lema.

Lema 5.10. *Sejam* $x, y \in B(x_0, R)$ *e* $0 \le t < v < R$. *Se*

$$t < t_*, \qquad \|x - x_0\| \le t, \qquad \|y - x\| \le v - t, \qquad f(v) \le 0, \quad e \quad F(y) = 0,$$

então,

$$\|y - N_F(x)\| \leqslant [v - n_f(t)] \frac{\|y - x\|^2}{(v - t)^2}.$$

Demonstração. De simples manipulações algébricas temos que

$$\begin{aligned} y - N_F(x) &= y - x + F'(x)^{-1}F(x) - F'(x)^{-1}F(y) \\ &= -F'(x)^{-1}\left[F(y) - F(x) - F'(x)(y - x)\right] = -F'(x)^{-1}E(x, y), \end{aligned}$$

onde usamos que $F(y) = 0$ na primeira igualdade e a definição em (5.15) na última igualdade. Da equação acima temos trivialmente que

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

Tomando norma em ambos lados da equação acima e usando a Proposição 5.6, o Lema 5.7 junto com as hipóteses do lema obtemos a estimativa

$$\|y - N_F(x)\| \le -\frac{1}{f'(t)} \; e(t, v) \, \frac{\|y - x\|^2}{(v - t)^2}.$$

Como $0 \le t < t_*$, tem-se $f'(t) < 0$. Agora, usando também (5.16) e a hipótese $f(v) \le 0$ temos que

$$\begin{aligned} (-1/f'(t)) \;\; e(t, v) &= v - t + f(t)/f'(t) - f(v)/f'(t) \\ &\le v - t + f(t)/f'(t) = v - n_f(t). \end{aligned}$$

Para finalizar a prova, combinemos as duas equações acima. □

Corolário 5.11. *As seqüências* $\{x_k\}$ *e* $\{t_k\}$ *satisfazem a condição*

$$\|x_* - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2}\,\|x_* - x_k\|^2, \qquad \textit{para} \quad k = 0, 1, \ldots . \tag{5.26}$$

Em particular,

$$\|x_* - x_{k+1}\| \leqslant \frac{1}{2}\|x_* - x_k\|, \qquad \textit{para} \quad k = 0, 1, \ldots \ . \tag{5.27}$$

Além disso, se f *satisfaz a hipótese* **h4**, *então*

$$\|x_* - x_{k+1}\| \leqslant \frac{D_- f'(t_*)}{-2f'(t_*)}\|x_* - x_k\|^2, \qquad \textit{para} \quad k = 0, 1, \ldots \ . \tag{5.28}$$

Demonstração. Tomemos um k arbitrario e apliquemos o Lema 5.10 com $x = x_k$, $y = x_*$, $t = t_k$ e $v = t_*$, para obter a desigualdade

$$\|x_* - N_F(x_k)\| \leq [t_* - n_f(t_k)]\, \frac{\|x_* - x_k\|^2}{(t_* - t_k)^2}.$$

A inequação (5.26) segue da inequação acima, e das definições em (5.23) e (5.14).

Notemos que, por (5.11), (5.14) e Corolário 5.9, para qualquer k

$$\frac{t_* - t_{k+1}}{t_* - t_k} \leq 1/2 \quad \text{e} \quad \frac{\|x_* - x_k\|}{t_* - t_k} \leq 1.$$

Combinando estas inequações com (5.26) temos (5.27).

Agora, assumamos que, a hipótese **h4** vale. Então, pelo Corolário 5.5, vale a segunda inequação em (5.6), a qual, combinado com (5.26) implica (5.28). □

Corolário 5.12. *O limite* x_* *da seqüência* $\{x_k\}$ *é o único zero de* F *em* $B[x_0, t_*]$.

Além disso, se f *satisfaz a hipótese* **h4**, *então* x_* *é o único zero de* F *em* $B(x_0, \bar{\tau})$, *onde* $\bar{\tau}$ *é definido como no Teorema 5.1, i.e.,*

$$\bar{\tau} = \sup\{t \in [t_*, R) \ : \ f(t) \leq 0\,\}.$$

Demonstração. Seja y_* um zero de F em $B[x_0, t_*]$. Isto implica que:

$$\|y_* - x_0\| \leq t_*, \qquad F(y_*) = 0.$$

Provaremos por indução que

$$\|y_* - x_k\| \leq t_* - t_k, \qquad k = 0, 1, \dots . \tag{5.29}$$

Para $k = 0$ a inequação acima vale trivialmente, porque $t_0 = 0$. Agora, assumamos que a inequação vale para algum $k > 0$. De (5.24) temos $\|x_k - x_0\| \leq t_k$. Logo, podemos aplicar o Lema 5.10 com $x = x_k$, $y = y_*$, $t = t_k$ e $v = t_*$ e obter a desigualdade

$$\|y_* - N_F(x_k)\| \leq [t_* - n_f(t_k)] \frac{\|y_* - x_k\|^2}{(t_* - t_k)^2}.$$

Usando a hipótese indutiva na última inequação (para estimar o quociente no último termo), (5.14) e (5.23) obtemos que (5.29) também vale para $k + 1$. Isto completa a prova de (5.29). Já que $\{x_k\}$ converge a x_* e $\{t_k\}$ converge a t_*, de (5.29) concluímos que $y_* = x_*$. Portanto, x_* é o único zero de F em $B[x_0, t_*]$.

Agora, suponhamos que f satisfaz a hipótese **h4** e assim, equivalentemente, $t_* < \bar{\tau}$. Já mostramos que se $y_* \in B[x_0, t_*]$ então $y_* = x_*$. Resta provar que F não tem zeros em $B(x_0, \bar{\tau}) \setminus B[x_0, t_*]$. Provaremos este fato por contradição, assumamos que F tem um zero, i.e., existe $y_* \in X$, tal que

$$t_* < \|y_* - x_0\| < \bar{\tau}, \qquad F(y_*) = 0.$$

Mostraremos que a suposição acima não pode verificar-se. Primeiramente, usando o Lema 5.7 com $x = x_0$, $y = y_*$, $t = 0$ e $v = \|y_* - x_0\|$ obtemos que

$$\|F'(x_0)^{-1} E(x_0, y_*)\| \leqslant e(0, \|y_* - x_0\|) \frac{\|y_* - x_0\|^2}{\|y_* - x_0\|^2} = e(0, \|y_* - x_0\|).$$

Como estamos aceitando que $F(y_*) = 0$, usando também (5.15) e (5.2) concluímos que

$$\begin{aligned}\|F'(x_0)^{-1}E(x_0, y_*)\| &= \|F'(x_0)^{-1}[-F(x_0) - F'(x_0)(y_* - x_0)]\| \\ &= \|y_* - x_0 + F'(x_0)^{-1}F(x_0)\| \\ &\geq \|y_* - x_0\| - \|F'(x_0)^{-1}F(x_0)\| \\ &\geq \|y_* - x_0\| - f(0).\end{aligned}$$

Logo, de (5.16) e da hipótese **h1** do Teorema 5.1 temos que

$$e(0, \|y_* - x_0\|) = f(\|y_* - x_0\|) - f(0) + \|y_* - x_0\|.$$

Agora, combinando esta equação com as duas inequações acima é fácil ver que

$$f(\|y_* - x_0\|) - f(0) + \|y_* - x_0\| \geq \|y_* - x_0\| - f(0),$$

ou equivalentemente, $f(\|y_* - x_0\|) \geq 0$. Deste modo f será estritamente convexa, é estritamente positiva no intervalo $(\|y_* - x_0\|,\, R)$. Assim, $\bar{\tau} \leq \|y_* - x_0\|$, em contradição com as suposicões acima. Portanto, F não tem zeros em $B(x_0, \bar{\tau}) \setminus B[x_0, t_*]$ e x_* é o único zero de F em $B(x_0, \bar{\tau})$. □

Por conseguinte, segue-se de Corolário 5.5, Corolário 5.9, Corolário 5.11 e Corolário 5.12 que todas as afirmacões no Teorema 5.1 são válidas.

5.2.4 Caso limite para o teorema de Kantorovich

Para provar convergência e estimar a sua taxa, só as regiões $[0, t_*]$ e $B[x_0, t_*]$ foram consideradas. De fato, para obter convergência Q-quadrática, a ação de f mais do que t_* não foi usada. Assim, daremos agora uma formulação a qual envolve só as regiões mencionadas.

Teorema 5.13. *Sejam X, Y espaços de Banach, $C \subseteq X$ e $F : C \to Y$ uma função contínua, continuamente diferenciável no $int(C)$.*

Tomemos $x_0 \in \text{int}(C)$ *com* $F'(x_0)$ *não-singular. Suponhamos que exista* $t_* > 0$ *e uma função continuamente diferenciável* $f : [0, t_*] \to \mathbb{R}$ *tal que,* $B[x_0, t_*] \subseteq C$,

$$\|F'(x_0)^{-1}\,[F'(y) - F'(x)]\,\| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|),$$

para $x, y \in B(x_0, t_*)$, $\|x - x_0\| + \|y - x\| < t_*$,

$$\|F'(x_0)^{-1}F(x_0)\| \leq f(0)$$

e

h1') $f(0) > 0$, $f'(0) = -1$;

h2') f' *é convexa e crescente;*

h3') $f(t) > 0$ *em* $[0, t_*)$ *e* $f(t_*) = 0$.

Então as seqüências geradas pelo método de Newton para resolver $f(t) = 0$ *e* $F(x) = 0$ *com pontos iniciais* $t_0 = 0$ *e* x_0, *respectivamente,*

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

estão bem definidas, $\{t_k\}$ *é crescente, está contida em* $[0, t_*)$, *e converge para* t_*, $\{x_k\}$ *está contida em* $B(x_0, t_*)$ *e converge para um ponto* $x_* \in B[x_0, t_*]$ *o qual é o único zero de* F *em* $B[x_0, t_*]$,

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

e as seqüências $\{t_k\}$ *e* $\{x_k\}$ *convergem* Q*-linearmente como segue*

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

Se, além disso,

h4') $f'(t_*) < 0$,

então as seqüências $\{t_k\}$ *e* $\{x_k\}$ *convergem* Q*-quadraticamente como segue*

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

5.3 Casos especiais

O teorema de Kantorovich foi usado em Wang [26] para provar o teorema de Smale [23], e foi usado em Alvarez, Botle e Munier [1] para provar o teorema de Nesterov-Nemirovskii [17]. Apresentaremos estas provas agora.

Iniciaremos com a definição.

Definição 5.14. *Sejam X, Y espaços de Banach, $C \subseteq X$ e $F : C \to Y$ uma função contínua, contínuamente diferenciável em int(C). Seja $x_0 \in \operatorname{int}(C)$ com $F'(x_0)$ não-singular. Uma função contínuamente diferenciável $f : [0, R) \to \mathbb{R}$ é dita uma função majorante para F em x_0 se $B(x_0, R) \subseteq C$ e* (5.1)*,* (5.2)*,* **h1***,* **h2***,* **h3***,* **h4** *do Teorema 5.1. são satisfeitas.*

Os seguintes resultados mostrarão uma condição mais fácil para verificar do que a condição (5.1), quando temos funções duas vezes continuamente diferenciáveis.

Lema 5.15. *Sejam X, Y espaços de Banach, $C \subseteq X$ e $F : C \to Y$ uma função contínua, duas vezes continuamente diferenciável no int(C). Seja $f : [0, R) \to \mathbb{R}$ uma função duas vezes continuamente diferenciável com derivada f' convexa. Então F satisfaz* (5.1) *se, só se,*

$$\|F'(x_0)^{-1}F''(x)\| \leqslant f''(\|x - x_0\|), \tag{5.30}$$

para todo $x \in int(C)$ tal que $\|x - x_0\| < R$.

Demonstração. Se F satisfaz (5.1), então (5.30) verifica-se trivialmente.

Reciprocamente, tomando $x, y \in C$ tal que $\|x - x_0\| + \|y - x\| < R$, obtemos que

$$\|F'(x_0)^{-1}[F'(y) - F'(x)]\| \leqslant \int_0^1 \|F'(x_0)^{-1}F''(x + \tau(y - x))\| \|y - x\| d\tau.$$

Agora, como f satisfaz (5.30) e f' é convexa, obtemos da última inequação que

$$\begin{aligned}\|F'(x_0)^{-1}[F'(y)-F'(x)]\| &\leqslant \int_0^1 f''(\|(x-x_0)+\tau(y-x)\|)\|y-x\|d\tau \\ &\leqslant \int_0^1 f''(\|x-x_0\|+\tau\|y-x\|)\|y-x\|d\tau \\ &= f'(\|x-x_0\|+\|y-x\|)-f'(\|x-x_0\|),\end{aligned}$$

a qual implica que F satisfaz (5.1), e o lema esta provado. □

Teorema 5.16. *(Teorema de Smale). Sejam X, Y espaços de Banach, $C \subseteq X$ e $F : C \to Y$ uma função contínua, analítica no $int(C)$. Tomemos $x_0 \in int(C)$ com $F'(x_0)$ não-singular e definimos*

$$\gamma := \sup_{k>1} \left\| \frac{F'(x_0)^{-1}F^{(k)}(x_0)}{k!} \right\|^{\frac{1}{k-1}}.$$

Suponhamos que $B(x_0, 1/\gamma) \subseteq C$ e que exista $\beta > 0$ tal que

$$\left\|F'(x_0)^{-1}F(x_0)\right\| \leqslant \beta,$$

e $\alpha := \beta\gamma \leqslant 3-2\sqrt{2}$. Então a seqüência gerada pelo método de Newton para resolver $F(x) = 0$ com ponto inicial x_0

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

está bem definida, $\{x_k\}$ está contida em $B(x_0, t_*)$ e converge para um ponto x_* o qual é o único zero de F em $B[x_0, t_*]$, onde $t_* := (\alpha + 1 - \sqrt{(\alpha+1)^2 - 8\alpha})/(4\gamma)$. Além disso, $\{x_k\}$ converge Q-linearmente como segue

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

Adicionalmente, se $\alpha < 3 - 2\sqrt{2}$, então $\{x_k\}$ converge Q-quadraticamente como segue

$$\|x_* - x_{k+1}\| \leqslant \frac{\gamma}{(1-\gamma t_*)[2(1-\gamma t_*)^2 - 1]}\|x_* - x_k\|^2, \qquad k = 0, 1, \dots ,$$

e x_* é o único zero de F em $B(x_0, t_{**})$, onde $t_{**} := (\alpha + 1 + \sqrt{(\alpha+1)^2 - 8\alpha})/(4\gamma)$.

Demonstração. É fácil mostrar que $f : [0, 1/\gamma) \to \mathbb{R}$ definida por

$$f(t) = t/(1 - \gamma t) - 2t + \beta,$$

satisfaz as seguintes condições: $f(0) = \beta$, $f'(t) = 1/(1 - \gamma t)^2 - 2$, $f'(0) = -1$,

$$f''(t) = \frac{2\gamma}{(1 - \gamma t)^3}, \qquad f'''(t) = \frac{6\gamma}{(1 - \gamma t)^4} > 0.$$

Da hipótese do teorema temos que, $\|F'(x_0)^{-1}F(x_0)\| \leqslant \beta = f(0)$ e $f(t_*) = 0$. Disto e das relações acima podemos afirmar que f, F e x_0 satisfazem (5.2), **h1**, **h2**, **h3**, **h4** do Teorema 5.1.

Agora, seja $x \in B(x_0, 1/\gamma)$. Já que F é analítica em x, o Lema 4.19 e a equação para a segunda derivada para f dada acima implicam que

$$\left\|F'(x_0)^{-1}F''(x)\right\| \leq \frac{2\gamma}{(1 - \gamma \|x - x_0\|)^3} = f''(\|x - x_0\|).$$

Portanto, do Lema 5.15, concluimos que F, f e x_0 satisfazem (5.1). Consequentemente, da definicão 5.14, $f(t) = t/(1 - \gamma t) - 2t + \beta$, é uma função majorante para F em x_0, com raizes iguais a t_* e t_{**} . Assim, os resultados deste teorema seguem do Teorema 5.1.

□

Teorema 5.17. *(Teorema Nesterov-Nemirovskii). Seja $C \subset \mathbb{R}^n$ um conjunto convexo aberto e seja $g : C \to \mathbb{R}$ uma função estritamente convexa, três vezes continuamente diferenciável no int(C). Tomemos $x_0 \in$ int(C) com $g''(x_0)$ não-singular. Suponhamos que g é a-auto-concordante, então a seqüência gerada pelo método de Newton para resolver $g'(x) = 0$ (ou equivalentemente, para minimizar g) com ponto inicial x_0*

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

está bem definida, $\{x_k\}$ está contida em $W_{t_*}(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} < t_*\}$ e converge para um ponto x_* em $W_{t_*}[x_0] = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} \leq t_*\}$ onde $t_* := (\beta + 1 -$

$\sqrt{(\beta+1)^2-8\beta})/4$. *Além disso,* $\{x_k\}$ *converge* Q*-linearmente como segue*

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

Adicionalmente, se $\beta < 3 - 2\sqrt{2}$*, então* $\{x_k\}$ *converge* Q*-quadraticamente como segue*

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

e x_* *está em* $W_{t_{**}}(x_0)$*, onde* $t_{**} := (\beta + 1 + \sqrt{(\beta+1)^2-8\beta})/4$.

Demonstração. Seja $X := (\mathbb{R}^n, \|.\|_{x_0})$ um espaço de Banach. Definamos $f : [0,1) \to \mathbb{R}$ tal que $f(t) = t/(1-t) - 2t + \beta$. É fácil verificar que f satisfaz as seguintes condições: $f(0) = \beta$, $f'(t) = 1/(1-t)^2 - 2$, $f'(0) = -1$,

$$f''(t) = \frac{2}{(1-t)^3}, \qquad f'''(t) = \frac{6}{(1-t)^4} > 0.$$

Da hipótese do teorema temos que, $\|g''(x_0)^{-1}g'(x_0)\| \leqslant \beta = f(0)$ e $f(t_*) = 0$. Disto e das relações acima, podemos afirmar que f, g' e x_0 satisfazem (5.2), **h1**, **h2**, **h3**, **h4** do Teorema 5.1.

Usando o Lema 2.28 com $\bar{x} = x_0$ e a expresão de $f''(t)$ dada acima obtemos as relações

$$\|g''(x_0)^{-1}g'''(x)\|_{x_0} \leq \frac{2}{(1-\|x-x_0\|_{x_0})^3} = f''(\|x-x_0\|_{x_0}), \qquad \forall\, x \in W_1(x_0).$$

Portanto, do Lema 5.15, f, g' e x_0 satisfazem (5.1). Consequentemente, da definicão 5.14, $f(t) = t/(1-t) - 2t + \beta$, é uma função majorante para g' em x_0, com raizes iguais a t_* e t_{**}, veja Alvarez, Botle e Munier [1]. Assim, os resultados deste teorema seguem do Teorema 5.1.

$\square$

5.4 Conclusões finais

Neste trabalho foi apresentado uma nova análise de convergência local e semi-local do método de Newton baseada no princípio majorante de Kantorovich, deixando muito claro, a relação da função majorante com o operador não linear em consideração. Mediante esta nova análise mostramos também que teoremas anteriormente não relacionados com os teoremas local e semi-local do método de Newton são relacionados. Assim, temos que o teorema de Smale (achar zeros de uma função analítica) e o teorema de Nesterov-Nemirovskii (achar o minimizador de uma função a-auto-concordante), podem ser mostrados também utilizando nosso análise de convergência local e semilocal do método de Newton.

Conseqüentemente, podemos fazer as seguintes perguntas: qual será o maior conjunto de funções para o qual o método de Newton funciona bem?; é possível tirar alguns condições impostas para a função majorante em nosso análise?. Acho que esta dissertação será de grande utilidade para estudar as perguntas acima.

Referências Bibliográficas

[1] Alvarez, F., Botle, J. e Munier, J., *A Unifying Local Convergence Result for Newton's Method in Riemannian Manifolds* INRIA, Rapport de recherche, N. 5381, (2004).

[2] Blum, L., Cucker, F. Shub, M. e Smale, S. *Complexity and real computation*, Springer-Verlag, New York, (1997).

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

[4] Ferreira, O.P., *Local Convergence of Newton's Method in Banach Space from the Viewpoint of the Majorant Principle*, submetida (2007).

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

[6] Ferreira, O.P. e Svaiter, B. F., *Kantorovich's Majorants Principle for Newton's Method*, to appear in Computational Optimization and Applications, (2006).

[7] Hiriart-Urruty, J.-B e Lemaréchal, C. *Convex analysis and minimization algorithms* I, Berlin Heidelberg, Springer-Verlag (1993).

[8] Izmailov, A., e Solodov, M., *Otimizção-volume 1 Condicoes de Otimalidade, Elementos de Análise Convexa e de Dualidade* IMPA (2005).

[9] Kantorovich, L.V., *Principe of majorants and Newton's method* Doklady AN SSSR, **76**, 1 (1951), pp. 104-144.

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

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

[12] Kolmogorov, A.N., e Fomin, S.V.,*Elementos da teoria das funções e de análise funcional*, Moscou, (1982).

[13] Li, Chong; Wang, Jinhua *Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition.* IMA J. Numer. Anal. 26 (2006), no. 2, 228–251.

[14] Lima, E.*Curso de Análise Volumen 1,* IMPA, (2002).

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

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

[17] Nesterov, Y. e Nemirovskii, A. *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, 13, Philadelphia, (1994).

[18] Ortega, J.M., *Numerical Analysis: A second couse*, SIAM Classics in Applied Mathematics 3, Philadelphia, 1990.

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

[20] Polyak, B. T., *Newton's method and its use in optimization.* To appear in European Journal of Operational Research (2006), doi:10.1016/j.ejor.2005.06.076.

[21] Rall, L. B., *A note on the convergence of Newton's Method*, SIAM J. Numer. Anal. **11**, 1 (19974), pp. 34–36.

[22] Shub,M. e Smale, S., *Complexity of Bezout's theorem.I:Geometric aspects*, Journal of the American Mathematical Society, **6**, 2 (1993), pp. 459–499.

[23] Smale, S. , *Newton method estimates from data at one point*, The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics (Ewing, R., Gross, K., and Martin, C. eds.), Springer-Verlag, New York, (1986), pp. 185–196.

[24] Traub, J. F. e H. Wozniakowski, ,*Convergence and complexity of Newton iteration for operator equation*, Journal of the Association Computing Machinery, **26**, 2 (1979), pp. 250 - 258.

[25] Wang, Jin-hua; Li, Chong *Uniqueness of the singular points of vector fields on Riemannian manifolds under the γ-condition*, J. Complexity 22 (2006), no. 4, 533–548.

[26] Wang, X., *Convergence of Newton's method and inverse function theorem in Banach space*, Math. Comp., **68**, 225 (1999), pp.169-186.

[27] Wang, X., *Convergence of Newton's method and uniqueness of the solution of equation in Banach space*, IMA Journal of Numerical Analysis. **20**, (2000), pp.123-134.

[28] Wayne, C. E., *An introduction to KAM theory. Dynamical systems and probabilistic methods in partial differential equation*, Lectures in Appl. Math., 31, Amer. Math. Soc., Providence, 1996.