Universidade Federal do Piauí
Centro de Ciências da Natureza
Pós-Graduação em Matemática
Mestrado em Matemática

Melhorias na Análise de Convergência Local do Método de Newton sob a Condição Majorante

Gilson do Nascimento Silva

Teresina - 2013

Gilson do Nascimento Silva

Dissertação de Mestrado:

Melhorias na Análise de Convergência Local do Método de Newton sob a Condição Majorante

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Matemática, da Universidade Federal do Piauí, como requisito parcial para obtenção do grau de Mestre em Matemática.

Orientador:
Prof. Dr. Paulo Sérgio Marques dos Santos

Teresina - 2013

Silva, G. N.

Melhorias na análise de convergência Local do Método de Newton sob a Condição Majorante

Gilson do Nascimento Silva– Teresina: 2013.

Orientador: Prof. Dr. Paulo Sérgio Marques dos Santos.

1. Otimização

CDD 516.36

Ao Deus de nosso Senhor Jesus Cristo, à minha família e aos meus amigos.

Agradecimentos

Primeiramente agradeço Àquele que olhou pra mim quando ainda eu não existia, que andou comigo em todo tempo, ao autor e consumador da minha fé, Jesus Cristo, pois foi Ele quem me ajudou em todo momento, indo na minha frente preparando os caminhos para eu pisar.

Agradeço aos meus pais Manuel Francisco e Josefa Belina pelo apoio em todos esses anos, o carinho, a preocupação, a educação que eles me proporcionaram, seus ensinamentos e o respeito. Aos meus irmãos Sérgio e Tatiana, pela força e coragem que me passaram, amizade, carinho e respeito. A todos os meus familiares, aos meus amigos de infância, colegas, amigos de Boa-Hora que contribuirão de qualquer forma para esse objetivo.

Ao grande amigo e Professor Paulo Sérgio quero deixar meus sinceros agradecimentos, pelo exemplo que ele passou pra mim, pelos seus ensinamentos, pela orientação nesta dissertação, por ter acreditado em mim desde o terceiro período da graduação, enfim pela nossa sincera amizade. Também não poderia deixar de agradecer à sua digníssima esposa, Professora Sissy Sousa, minha orientadora em iniciação científica, pela sua paciência, pela coragem que ela me passou e, por toda ajuda que você me proporcionou para elaborar esta dissertação e, agradeço também ao Professor Arnaldo por ter aceitado nosso convite pra participar dessa banca e por todas as suas considerações.

À minha tia Maria das Dores e sua filha Elizete, venho agradecer por terem me recebido em suas casas quando vim morar em Teresina e por terem cuidado tão bem de mim.

Aqui também quero expressar meus agradecimentos à uma grande família que Deus me deu, aos irmãos da Igreja Assembleia de Deus no Bairro de Fátima, pela recepção, grandes ensinamentos, cuidados e carinho: Raimunara, Creso, Jaqueline, Mary, Adonias, Kelly, Kenmuel, Juliana, Reinaldo, Patrícia, Elioenai, Márcia, Daniel, Carla, Igor, Thiago, Ismaias, Edilza, Otávio e família, José Pereira, enfim à todos.

É impossível esquecer que um dia morei na Residência Universitária II da UFPI, foram dias incríveis, maravilhosos, que hoje tenho muitas saudades, pois lá construir muitas amizades pra vida toda, brigamos, conversamos e oramos juntos: Douglas, Ronald, Zé Filho, Darlan, Wellington Castro, Wellington Cruz, Clóvis, Vitaliano, Gislaylson, Geovane, Luardo, Bruno, Arimatéia, Ailton, Thalisson, Jorge Malaquias, Jorge, Thiago, Teixeira, Alex, Francisco Lima, Júnior, Romário, Artur, João Magalhães e todos os demais.

Reservo agora algumas linhas para agradecer à grandes homens (alguns pilantras) e grandes mulheres que fizeram parte desta história, amigos que eu não quero esquecer jamais, que me acompanham durante toda essa trajetória, me ajudaram e me deram todo o apoio necessário: Sandoel, Vitaliano, Rui Marques, Carlos Adriano, Marilene, Marilha, Cristiane, Fernando, Lucas, Jordan, Kadu, Mário, Neylon, Simone, Samara, Jordana, Mariane, Bruno, Elesbão, Israel, Franciane, Alex, Weslay, Alberone, Ramon, Lucas Vidal, Alex doido, Valdinês, Yuri, Leonardo, Bernardo, Diêgo, Micael, Antônio Amaral, Rodolfo, Joel e Renata.

Dedico este parágrafo a quatro amigos que marcaram minha vida, pela amizade verdadeira, por todo apoio e o amor que eu tenho por eles: Carlos Adriano, Rui Marques, Sandoel Vieira e Vitaliano Amaral.

Agradeço também a todos os professores do Departamento de Matemática, pois contribuíram bastante na minha formação acadêmica.

Agradeço à CAPES pelo apoio financeiro.

"Por que desde a antiguidade não se ouviu, nem com os ouvidos se percebeu, nem com os olhos se viu um Deus além de ti, que trabalhe para aquele que nele espera.".

Isaías 64.4.

Resumo

A busca por soluções de equações não lineares em espaços de Banach, é objeto de interesse em várias áreas da ciência e engenharias. Devido a sua velocidade de convergência e eficiência computacional, o método de Newton e suas variações têm sido bastante utilizados para o propósito de obter soluções dessas equações. Nesta dissertação, apresentamos uma análise de convergência local do método de Newton baseada no princípio majorante de Kantorovich. Esta abordagem melhora os resultados até então obtidos, no seguinte sentido: com as mesmas informações iniciais são apresentadas melhores estimativas para o raio de convergência. Casos especiais e exemplos numéricos são também considerados.

Abstract

The search for solutions of nonlinear equations in Banach spaces, is object of interest in several areas of science and engineerings. Due the speed of convergence and computational efficiency, the Newton method and its variations have been sufficiently used to obtain solutions of these equations. In the dissertation we present a local convergence analysis of the Newton method based on Kantorovich's majorant principle. This approach improves the known results, in the following sense: under the same information, larger estimates of the radius of convergence are provided. Special cases and numerical examples are also provided in this study.

Sumário

Introdução

Neste trabalho estamos interessados no problema de aproximação de uma solução localmente única x^* da equação

$$F(x) = 0, \tag{1}$$

onde F é um operador Fréchet-diferenciável definido em um subconjunto Ω aberto e convexo de um espaço de Banach X, com valores em um espaço de Banach Y.

O campo das Ciências da Computação tem visto um desenvolvimento considerável em Matemática, Ciências da Engenharia e Teoria do Equilíbrio Econômico [7]. Por exemplo, sistemas dinâmicos são modelados matematicamente por equações diferenciáveis e suas soluções usualmente representam os estados dos sistemas. Por uma questão de simplicidade, suponha que o sistema tempo-invariante é dirigido pela equação $x' = T(x)$, para algum operador T adequado onde x é o estado. Então os estados de equilíbrio são determinados resolvendo uma equação do tipo (1). Equações similares são usadas no caso de sistemas discretos.[7]

Muitos problemas de Engenharia, Economia, Física e outras disciplinas podem ser apresentados na forma da equação (1) usando modelagem matemática [7]. A incógnita das equações de Engenharia podem ser funções (equações diferenciáveis e integráveis), vetoriais (sistemas de equações algébricas lineares ou não lineares), números reais ou complexos (equações algébricas simples com incógnitas simples).

Exceto em casos especiais, os métodos de solução geralmente usados são iterativos - quando se inicia a partir de uma ou várias aproximações iniciais, uma sequência é construída que converge para uma solução da equação. Métodos iterativos também são aplicados para resolver problemas de Otimização. Em tais casos, a sequência de iteração converge para uma solução ótima do problema em questão. Uma vez que todos estes métodos tem a mesma estrutura recursiva, eles podem ser introduzidos e discutidos em um quadro geral.

Notemos que em Ciências da Computação, a prática da análise numérica para encontrar soluções é essencialmente ligada a variações do Método de Newton [2].

A ideia básica do Método de Newton, para encontrar uma raiz de uma função diferenciável não linear, é bastante simples e consiste em substituir a função não linear por uma aproximação linear. Seja $f : \mathbb{R} \to \mathbb{R}$ não linear, onde queremos resolver a equação $f(x) = 0$, isto é, estamos interessados em encontrar os zeros de f. Começando com um ponto inicial x^0 podemos construir uma aproximação linear de $f(x)$ em uma vizinhança de x^0:

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

Observe que como $f(x) = 0$ a equação anterior fica

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

Portanto, ao invés de resolver $f(x) = 0$, resolvemos a equação linear

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

Se $f'(x^0) \neq 0$, a equação acima tem solução x^1 tal que

$$x^1 = x^0 - f'(x^0)^{-1} f(x^0).$$

Agora, se $f'(x^1) \neq 0$ repetimos o mesmo processo acima e encontramos

$$x^2 = x^1 - f'(x^1)^{-1} f(x^1).$$

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

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

O método de iteração que vimos acima nos motiva a deduzir que a sequência do método de Newton para resolver $f(x) = 0$, onde f é uma função diferenciável não linear, é a sequência gerada pela seguinte regra de recorrência:

$$x^{k+1} = x^k - f'(x^k)^{-1} f(x^k) \quad (k \geqslant 0) \tag{2}$$

Observe que para esta sequência fazer sentido é necessário que todos os $\{x^k\}$ estejam contidos no domínio da f e que $f'(x^k) \neq 0$, para todo $k \geqslant 0$. Também não temos garantia da convergência da sequência $\{x^k\}$, portanto necessitamos de condições adequadas sobre f e o ponto inicial x^0 para garantir que $\{x^k\}$ convirja para um zero de f.

O método descrito acima foi proposto inicialmente por Isaac Newton em 1669 para encontrar raízes de funções polinomiais. Pouco tempo depois, em 1690 J. Raphson estendeu o método para funções reais quaisquer. Por isso é muito comum, na literatura, o método ser chamado de Método de Newton-Raphson. A consolidação do método está ligada a famosos matemáticos como J. Fourier, L. A. Cauchy, entre outros. Em 1818, Fourier provou que o método convergia quadraticamente desde que o ponto inicial fosse tomado em uma vizinhança da solução procurada, enquanto Cauchy mostrou que o método se estende naturalmente para funções de várias variáveis e usou-o para provar a existência de raízes de algumas equações. Em 1916, os matemáticos Fine e Bennet deram mais algumas contribuições para o método.

Fine em [14] provou a convergência para o caso n-dimensional sem a hipótese de existência de solução. Bennet em [8] estendeu para o caso de dimensão infinita. Mais recentemente, em 1948, L. V. Kantorovich em [23] provou a existência de solução e a convergência do método para operadores $T : X \to Y$, onde X e Y são espaços de Banach e T é um operador Fréchet-diferenciável qualquer.

A grande importância desse método reside no fato de que sob algumas hipóteses é garantida a convergência, a uma taxa relativamente alta, para uma solução. O método de Newton também é usado para mostrar outros teoremas importantes na Matemática. Por exemplo, teoremas de existência e unicidade de soluções para certas equações diferenciais, veja por exemplo [21, 27], o Teorema da Função Inversa e Implícita [24] e o Teorema do Mergulho Isométrico, veja [28].

A questão da convergência do método de Newton tem sido estudada extensivamente. Os resultados obtidos podem ser distinguidos entre duas classes: convergência local e convergência semi local. Uma desvantagem da análise de convergência local é que a proximidade de uma solução, e consequentemente existência, deve ser conhecida ou determinada a priori. Por outro lado, esta análise tem a vantagem de fornecer o raio de convergência ótimo.

Recentemente, alguns resultados usando funções convexas majorantes foram estabelecidos, veja [10, 11]. Em [10, 35], foram estabelecidos, sob uma condição majorante e condição de lipschitz generalizada, respectivamente, convergência local, taxa quadrática e estimativa do melhor raio de convergência possível do método de Newton, bem como a unicidade da solução para equações não lineares. Na análise apresentada em [10], foi assumido a convexidade da derivada da função escalar majorante e, em [35], o não decrescimento da função integrável positiva que define a condição de lipschitz generalizada foi assumido.

Nesta dissertação fomos motivados pelos recentes trabalhos de Ferreira [12] e Argyros [1], onde o primeiro enfraqueceu as condições de convergência de [13, 11, 34] para a análise de convergência local do método de Newton sob a condição de Lipschitz geral e forneceu uma estimativa do erro das distâncias $\|x^k - x^*\|$ $(k \geqslant 1)$, no que ele também afirmou ser o melhor raio de convergência possível. Usando as mesmas informações, em [1] foi mostrado que em geral o raio de convergência dado em [12] não é o melhor possível, mas pode ser melhorado.

Estas observações são muito importantes em matemática computacional, uma vez que permitem uma melhor escolha da estimativa do ponto inicial x^0 e menos iterações para obtermos o erro de tolerância desejado.

Esta dissertação está organizada da seguinte forma:

No Capítulo 1, apresentamos alguns tópicos de espaços de Banach, diferenciabilidade, resultados básicos de análise convexa em $\mathbb{R}$ e provaremos uma versão do Teorema de Newton-Kantorovich [29].

No Capítulo 2, apresentamos o teorema de convergência local do método de Newton sob a condição majorante provado em Ferreira [12]. Além disso, mostramos dois casos especiais do teorema principal.

No Capítulo 3, usando as mesmas informações do resultado descrito no Capítulo 2, mostramos as melhorias obtidas em [1]. Através de dois casos especiais e exemplos ilustrativos justificamos a teoria. Finalmente, no Capítulo 4, reunimos nossas conclusões e algumas direções para trabalhos futuros.

Capítulo 1

Conceitos Básicos

1.1 Espaços de Banach

Nesta seção definiremos espaço de Banach, norma de um operador linear e provaremos alguns resultados necessários, dentre os quais o famoso Lema de Banach.

Definição 1.1.1 ([26])**.** *Uma métrica num conjunto* X *é uma função* $d : X \times X \to \mathbb{R}$, *que associa a cada par ordenado de elementos* $x, y \in X$ *um número real* $d(x, y)$, *chamado a distância de* x *a* y, *de modo que sejam satisfeitas as seguintes condições para quaisquer* $x, y, z \in X$:

d_1) $d(x, x) = 0$;

d_2) *Se* $x \neq y$ *então* $d(x, y) > 0$;

d_3) $d(x, y) = d(y, x)$;

d_4) $d(x, z) \leqslant d(x, y) + d(y, z)$.

Os postulados d_1) e d_2) dizem que $d(x, y) \geqslant 0$ e que $d(x, y) = 0$ se, e somente se, $x = y$. O postulado d_3) afirma que a distância $d(x, y)$ é uma função simétrica nas variáveis x, y. A condição d_4) chama-se desigualdade triangular.

Definição 1.1.2 ([26])**.** *Um espaço métrico é um par* (X, d), *onde* X *é um conjunto e* d *é uma métrica em* X.

Veremos alguns exemplos de espaços métricos.

Exemplo 1.1.1 ([26])**.** *A reta, ou seja, o conjunto* $\mathbb{R}$ *dos números reais, é o exemplo mais importante de espaço métrico. A distância entre dois pontos* $x, y \in \mathbb{R}$ *é dada por*

$d(x, y) = |x - y|$. *As condições* $d_1)$ *a* $d_4)$ *resultam imediatamente das propriedades elementares do valor absoluto de números reais.*

Exemplo 1.1.2 ([26])**.** *O espaço euclidiano* $\mathbb{R}^n$*. Os pontos de* $\mathbb{R}^n$ *são as listas* $x = (x_1, \ldots, x_n)$ *onde cada uma das n coordenadas* x_i *é um número real. Há três maneiras naturais de se definir a distância entre dois pontos em* $\mathbb{R}^n$*. Dados* $x = (x_1, \ldots, x_n)$ *e* $y = (y_1, \ldots, y_n)$ *escreveremos:*

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2} = \left[\sum_{i=1}^{n}(x_i - y_i)^2\right]^{1/2}$$
$$d'(x, y) = |x_1 - y_1| + \ldots + |x_n - y_n| = \sum_{i=1}^{n} |x_i - y_i| \ \text{ e}$$
$$d''(x, y) = \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\} = \max |x_i - y_i|.$$

As funções $d, d', d'' : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *são métricas.*

A seguir definiremos o conceito de convergência de sequência em um espaço métrico qualquer, a qual usaremos em toda a teoria do método de Newton.

Definição 1.1.3 ([25])**.** *Uma sequência* $\{x^k\}$ *em um espaço métrico* $X = (X, \|.\|)$ *é dita convergente se existe um* $x^* \in X$ *tal que*

$$\lim_{k \to \infty} \|x^k - x^*\| = 0,$$

x^* *é chamado o limite de* $\{x^k\}$ *e escrevemos*

$$\lim_{k \to \infty} x^k = x^*.$$

No sentido usual, definimos também sequência de Cauchy, com o objetivo de analisar a convergência do método de Newton.

Definição 1.1.4 ([25])**.** *Uma sequência* $\{x^k\}$ *em um espaço métrico* $X = (X, \|.\|)$ *é chamada sequência de Cauchy se dado* $\epsilon > 0$ *existe* k_0 *tal que*

$$\|x^k - x^l\| < \epsilon, \ \ \forall \ k, l \geqslant k_0.$$

Agora definimos espaço de Banach, que é o espaço onde o método de Newton está definido, e apresentamos alguns exemplos e propriedades desses espaços.

Definição 1.1.5 ([25])**.** *Um espaço vetorial normado* X *é dito de Banach se toda sequência de Cauchy em* X *é convergente.*

Exemplo 1.1.3. *O espaço euclidiano* $\mathbb{R}^n$ *e o espaço unitário* $\mathbb{C}^n$ *são espaços de Banach com a norma usual.*

Demonstração. Veja página 33 de [25]. □

Exemplo 1.1.4 ([25])**.** *O espaço* $C[a,b]$*, das funções contínuas definidas em* $[a,b]$ *é um espaço de Banach com a norma dada por*

$$\|x\| = \max\{|x(t)|;\ t \in [a,b]\}.$$

Demonstração. Veja página 36 de [25]. □

O conceito de sequências majorantes que definiremos agora tem um papel fundamental para encontrar o principal resultado desta dissertação. Veremos ainda neste capítulo que Kantorovick usou essas sequências para resolver o método de Newton em espaços de Banach.

Definição 1.1.6 ([30])**.** *Seja* $\{x^k\}$ *uma sequência qualquer em um espaço de Banach* X*. Então a sequência* $\{t_k\} \subset [0,\infty)$ *para o qual*

$$\|x^{k+1} - x^k\| \leqslant t_{k+1} - t_k, \quad \forall\, k = 0, 1, \ldots$$

chama-se sequência majorante para $\{x^k\}$*.*

Note que qualquer sequência majorante é necessariamente monótona crescente.

Lema 1.1.1 ([29])**.** *Sejam* X *um espaço de Banach,* $\{y_k\}$ *uma sequência em* X *e* $\{t_k\}$ *uma sequência de números reais não negativos tal que*

$$\|y_{k+1} - y_k\| \leqslant t_{k+1} - t_k$$

e $t_k \to t_*$*. Então existe um* $y_* \in X$ *tal que* $y_k \to y^*$ *e*

$$\|y^* - y_k\| \leqslant t_* - t_k.$$

Demonstração. A prova segue a partir de

$$\|y_{k+p} - y_k\| \leqslant \sum_{i=1}^{p} \|y_{k+i} - y_{k+i-1}\| \leqslant t_{k+p} - t_k \leqslant t_* - t_k,$$

que mostra que $\{y_k\}$ é uma sequência de Cauchy. Como X é Banach segue que existe $y_* \in X$ tal que $y_k \to y^*$. □

Proposição 1.1.1 ([30]). *Seja* $\{x^k\}$ *uma sequência em um espaço de Banach* X. *Se existe um número real* $0 < \alpha < 1$ *tal que*

$$\|x^{k+1} - x^*\| \leqslant \alpha\|x^k - x^*\|, \ \ \forall \ k = 0, 1, ...$$

então $\{x^k\}$ *converge para* x^*.

Demonstração. Note que para $k = 0$ e $k = 1$ temos respectivamente

$\|x^1 - x^*\| \leqslant \alpha\|x^0 - x^*\|,$
$\|x^2 - x^*\| \leqslant \alpha\|x^1 - x^*\| \leqslant \alpha^2\|x^0 - x^*\|,$

e repetindo este processo k vezes obtemos

$$\|x^k - x^*\| \leqslant \alpha^k\|x^0 - x^*\|, \ \ \forall \ k = 0, 1, ...$$

Como $0 < \alpha < 1$ temos que $\lim\limits_{k\to\infty} \alpha^k = 0$ e portanto

$$\lim_{k\to\infty} \|x^k - x^*\| = 0,$$

ou seja, $\{x^k\}$ converge para x^*. □

Definição 1.1.7. *Sejam* X *um espaço de Banach e a sequência* $\{x^k\} \subset X$ *convergindo para* x^*. *Dizemos que:*
i) $\{x^k\}$ *converge linearmente se existe* $c \in [0, 1)$, *tal que*

$$\|x^{k+1} - x^*\| \leqslant c\|x^k - x^*\|, \ \ \forall \ k = 0, 1, ... \tag{1.1}$$

ii) $\{x^k\}$ *converge quadraticamente se existe* $c > 0$, *tal que*

$$\|x^{k+1} - x^*\| \leqslant c\|x^k - x^*\|^2, \ \ \forall \ k = 0, 1, ...$$

Definição 1.1.8 ([25]). *Sejam* X *e* Y *espaços de Banach,* $\pounds(X, Y)$ *o espaço dos operadores lineares de* X *em* Y *e* $T \in \pounds(X, Y)$. *Defina a norma do operador* T *como sendo*

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

Veja que a norma de T satisfaz as seguintes desigualdades:
i) $\|Tu\| \leqslant \|T\|\|u\|, \ \forall \ T \in \pounds(X, Y)$ e $u \in X$;
ii) $\|ST\| \leqslant \|S\|\|T\|$ e $\|S + T\| \leqslant \|S\| + \|T\|, \ \forall \ S, T \in \pounds(X, Y)$;
iii) $\|T^k\| \leqslant \|T\|^k$.

Proposição 1.1.2 ([30]). *Sejam* X *e* Y *espaços de Banach. Se* $T \in \mathcal{L}(X, Y)$ *é tal que* $\|T\| < 1$ *então* $I - T$ *é invertível e vale*

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

Demonstração. Considere as seguintes sequências $\{s_k\}$ e $\{t_k\}$ definidas respectivamente por $s_k = I + T + T^2 + ... + T^k$ e $t_k = 1 + \|T\| + ... + \|T\|^k$. Note que

$$\|s_{k+1} - s_k\| = \|T^{k+1}\| \leqslant \|T\|^{k+1} = t_{k+1} - t_k.$$

Como $\|T\| < 1$, temos que $\{t_k\}$ é uma sequência monótona crescente, pois $t_{k+1} - t_k = \|T\|^{k+1} \geqslant 0 \Rightarrow t_{k+1} \geqslant t_k$, e é convergente, pois $\{t_k\}$ é a soma dos termos de uma progressão geométrica de razão $\|T\| < 1$, logo

$$\lim_{k\to\infty} t_k = \frac{1}{1 - \|T\|}.$$

Então pelo Lema 1.1.1 existe $s^* \in X$ tal que $\lim_{k\to\infty} s_k = s^*$. Veja também que

$$s_k(I - T) = (I + T + ... + T^k)(I - T) = I - T^{k+1} \tag{1.2}$$

e

$$\|I - (I - T^k)\| = \|T^k\| \leqslant \|T\|^k.$$

Como $\lim_{k\to\infty} \|T\|^k = 0$ ($\|T\| < 1$) temos da última desigualdade que $\lim_{k\to\infty} (I - T^k) = I$. Então da expressão (1.2) e do último limite temos que

$$\lim_{k\to\infty} s_k(I - T) = \lim_{k\to\infty} (I - T^{k+1}) = I,$$

e como existe $\lim_{k\to\infty} s_k$, temos que

$$(I - T) \lim_{k\to\infty} s_k = I,$$

então $\lim_{k\to\infty} s_k = (I - T)^{-1}$, ou seja, $I - T$ é invertível. Agora note que

$$\begin{aligned}
\|(I - T)^{-1}\| &= \|\lim_{k\to\infty} s_k\| \\
&\leqslant \lim_{k\to\infty} \|(I + T + ... + T^k)\| \\
&\leqslant \lim_{k\to\infty} (\|I\| + \|T\| + ... + \|T\|^k) \\
&= \lim_{k\to\infty} t^k = \frac{1}{1 - \|T\|}.
\end{aligned}$$

$\square$

Provaremos agora o conhecido Lema de Banach.

Lema 1.1.2 ([30]). *(Lema de Banach) Sejam* X *e* Y *espaços de Banach,* $B \in \mathcal{L}(X, Y)$ *um operador linear e* I *o operador identidade de* X. *Se* $\|I - B\| < 1$, *então* B *é invertível e vale*

$$\|B^{-1}\| \leqslant \frac{1}{1 - \|B - I\|}.$$

Demonstração. Tomando $T = I - B$ e observando que $\|T\| = \|I - B\| < 1$, então pela proposição anterior temos que $I - T = B$ é invertível e vale

$$\|(I - T)^{-1}\| = \|B^{-1}\| \leqslant \frac{1}{1 - \|I - B\|}.$$

□

Definição 1.1.9 ([30]). *Uma aplicação* $F : D \subset X \to Y$ *é Hölder-contínua em* $D_0 \subset D$ *se existe uma constante* $c \geqslant 0$ *e* $p \in (0, 1]$ *tal que, para todo* $x, y \in D$,

$$\|F(y) - F(x)\| \leqslant c\|y - x\|^p. \tag{1.3}$$

Se $p = 1$, *então* F *é contínua-Lipschitz em* D_0.

1.2 Diferenciabilidade

Nesta seção definiremos operador Fréchet-diferenciável e Gâteaux-diferenciável em espaços de Banach e, apresentaremos algumas de suas propriedades básicas. Lembremos inicialmente a definição de diferenciabilidade em $\mathbb{R}$. Para mais detalhes veja também [37].

Definição 1.2.1 ([37]). *Uma função* $f : \mathbb{R} \to \mathbb{R}$ *é diferenciável em* x *se existe um número real* $a := f'(x)$ *tal que*

$$\lim_{t \to 0} \frac{f(x + t) - f(x) - at}{t} = 0.$$

Esta definição se estende de forma natural para espaços de Banach.

Definição 1.2.2 ([37]). *Seja* $F : U(x) \subseteq X \to Y$ *uma aplicação qualquer com* X *e* Y *espaços de Banach e* $U(x)$ *uma vizinhança de* x. *Então:*

(1) A aplicação F *é Fréchet-diferenciável (*F*-diferenciável) em* x *se, e somente se, existe uma aplicação* $T \in \mathcal{L}(X, Y)$ *tal que*

$$F(x + h) - F(x) = T(h) + o(\|h\|), \quad h \to 0 \tag{1.4}$$

para todo h *em uma vizinhança de zero. Se existir, este* T *é chamado a* F*-derivada de* F *em* x*. Denotamos* $F'(x) = T$.

A F*-diferencial em* x *é definida por* $dF(x, h) = F'(x)h$.

(2) A aplicação F *é Gâteaux-diferenciável (*G*-diferenciável) em* x *se, e somente se, existe uma aplicação* $T \in \mathcal{L}(X, Y)$ *tal que*

$$F(x + tk) - F(x) = tT(k) + o(t), \quad t \to 0 \tag{1.5}$$

para todo k *com* $\|k\| = 1$ *e todo número real* t *em alguma vizinhança de zero.* T *é chamado a* G*-derivada de* F *em* x*. Denotamos* $F'(x) = T$

A G*-diferencial em* x *é definida por* $d_G F(x, h) = F'(x)h$.

(3) Se a F*-derivada (respect.* G*-derivada)* F' *(respect. G') existe para todo* $x \in A$*, então a aplicação*

$$F' : A \subseteq X \to L(X, Y) \quad x \mapsto F'(x)$$

é chamada a F*-derivada (respect.* G*-derivada) de* F *em* A.

(4) As derivadas de ordem superiores são definidas sucessivamente. Assim, $F''(x)$ *é a derivada de* F' *em* x.

As equações (1.4) e (1.5) mostram que as derivadas são definidas através de uma linearização. Obviamente $F'(x)$ é determinada unicamente por (1.4) e (1.5). Segue a partir de (1.5) que a G-derivada em x pode ser definida equivalentemente através de

$$F'(x)k = \lim_{t \to 0} \frac{F(x + tk) - F(x)}{t}. \tag{1.6}$$

Se definirmos $\varphi(t) = F(x + tk)$, então para a G-derivada, (1.6) implica que

$$\varphi'(0) = F'(x)k.$$

Definição 1.2.3 ([37]). *Se* $F : D \subset X \to Y$ *é* F*-diferenciável (respect.* G*-diferenciável) em* $x \in \mathrm{int}(D)$ *então o operador linear* $T \in \mathcal{L}(X, Y)$ *para o qual vale a Definição 1.2.2 é único.*

Proposição 1.2.1 ([37]). *As seguintes afirmações são verdadeiras:*

(a) Toda F*-derivada em* x *é também* G*-derivada em* x.

(b) A G*-derivada em* x *para o qual a passagem ao limite em (1.6) é uniforme para todo* k *com* $\|k\| = 1$*, é também uma* F*-derivada em* x.

(c) Se F' *existe como* G*-derivada em alguma vizinhança de* x*, e se* F' *é contínua em* x*, então* $F'(x)$ *é também uma* F*-derivada em* x*.*

(d) Se $F'(x)$ *existe como* F*-derivada em* x*, então* F *também é contínua am* x*.*

Demonstração. Veja página 137 de [37]. □

Proposição 1.2.2 ([37])**.** *Suponha que as seguintes aplicações* $F, G : U(x) \subseteq X \to Y$ *são* F*-diferenciáveis (respect.* G*-diferenciáveis) em* x*, onde* X *e* Y *são espaços de Banach. Então para todo* $\alpha, \beta \in \mathbb{R}$ *temos que*

$$(\alpha F + \beta G)'(x) = \alpha F'(x) + \beta G'(x).$$

Demonstração. Veja página 138 de [37]. □

Proposição 1.2.3 ([37])**.** *Seja* x *fixado e* $y = F(x)$*. Suponha as seguintes aplicações:*

$$F : U(x) \subseteq X \to Y \quad e \quad G : U(y) \subseteq Y \to Z$$

com $F(U(x)) \subseteq U(y)$*, onde* X*,* Y *e* Z *são espaços de Banach. Isto define a aplicação composição* $G \circ F : U(x) \subseteq X \to Z$*. Suponha que* $F'(x)$ *e* $G'(F(x))$ *existe como* F*-derivada. Então:*

(a) A aplicação composição $H = G \circ F$ *é* F*-diferenciável em* x *e*

$$H'(x) = G'(F(x))F'(x). \tag{1.7}$$

(b) Se $F'(x)$ *existe somente como* G*-derivada em* x*, então* H *é* G*-diferenciável em* x *e vale (1.7).*

Demonstração. Veja página 138 e 139 de [37]. □

Proposição 1.2.4 ([37])**.** *Sejam* X *e* Y *espaços de Banach e* $F : D \subset X \to Y$ *F-diferenciável. Se* $\|F'(y) - F'(x)\| \leqslant L\|x - y\|$*,* $\forall\ x, y \in D$ *e* L *constante, então*

$$\|F(y) - F(x) - F'(x)(y - x)\| \leqslant \frac{L}{2}\|y - x\|^2, \quad \forall\ x, y \in D. \tag{1.8}$$

Demonstração. Ponhamos $\psi(t) = F(x + t(y - x)) - tF'(x)(y - x)$, $\forall\ t \in [0, 1]$.
Da regra da cadeia temos:

$$\begin{aligned} \psi'(t) &= F'(x + t(y - x))(y - x) - F'(x)(y - x) \\ &= \left[F'(x + t(y - x)) - F'(x)\right](y - x) \end{aligned}$$

e veja que $\psi(1) - \psi(0) = \int_0^1 \psi'(t)dt$ e $\|\psi'(t)\| \leqslant Lt\|y - x\|^2$.

Então,

$$\begin{aligned}\|F(y) - F(x) - F'(x)(y - x)\| &= \|\psi(1) - \psi(0)\| \\ &\leqslant \int_0^1 Lt\|y - x\|^2 dt \\ &= \tfrac{L}{2}\|y - x\|^2.\end{aligned}$$

□

Proposição 1.2.5 ([30]). *Seja* $F : D \subset X \to Y$ *uma aplicação continuamente diferenciável em um conjunto convexo* $D_0 \subset D$ *e suponha que, para as constantes* $\alpha \geqslant 0$ *e* $p \geqslant 0$, F' *satisfaz*

$$\|F'(y) - F'(x)\| \leqslant \alpha\|y - x\|^p, \quad \forall \ x, y \in D_0.$$

Então, para todo $x, y \in D_0$,

$$\|F(y) - F(x) - F'(x)(y - x)\| \leqslant \frac{\alpha}{p + 1}\|y - x\|^{p+1}.$$

Demonstração. Segue diretamente do Teorema Fundamental do Cálculo e de simples manipulações algébricas que

$$\begin{aligned}\|F(y) - F(x) - F'(x)(y - x)\| &= \left\| \int_0^1 [F'(x + t(y - x)) - F'(x)](y - x)dt \right\| \\ &\leqslant \int_0^1 \|F'(x + t(y - x)) - F'(x)\| \|y - x\| dt \\ &= \int_0^1 \alpha\|t(y - x)\|^p \|y - x\| dt \\ &= \alpha\|y - x\|^{p+1} \int_0^1 t^p dt \\ &= \tfrac{\alpha}{p+1}\|y - x\|^{p+1}.\end{aligned}$$

□

1.3 Resultados de Análise Convexa em $\mathbb{R}$

As seguintes afirmações e resultados elementares de Análise Convexa serão fundamentais para o desenvolvimento deste trabalho. Mais detalhes podem ser encontrados em [18, 21, 35].

Definição 1.3.1 ([32]). *Um conjunto* $D \subset \mathbb{R}$ *é dito convexo se*

$$\lambda x + (1 - \lambda)y \in D, \ \ \forall \ x, y \in D \ \text{ e } \ \lambda \in [0, 1]. \tag{1.9}$$

O ponto $\lambda x + (1 - \lambda)y$*, onde* $\lambda \in [0, 1]$*, é denominado combinação convexa de x e y.*

Definição 1.3.2 ([32])**.** *Seja* $I \subset \mathbb{R}$ *um intervalo. Uma função* $\varphi : I \to \mathbb{R}$ *é dita convexa quando, para todo* $x, y \in I$ *e todo* $\lambda \in [0, 1]$

$$\varphi(\lambda x + (1 - \lambda)y) \leqslant \lambda\varphi(x) + (1 - \lambda)\varphi(y). \tag{1.10}$$

Quando a desigualdade acima é estrita, dizemos que a função φ *é estritamente convexa.*

Proposição 1.3.1 ([17])**.** *Sejam* $I \subset \mathbb{R}$ *um intervalo e* $\varphi : I \to \mathbb{R}$ *uma função convexa. Então dados* a, b *e* $c \in I$ *com* $a < b < c$*, temos*

$$\frac{\varphi(b) - \varphi(a)}{b - a} \leqslant \frac{\varphi(c) - \varphi(a)}{c - a} \leqslant \frac{\varphi(c) - \varphi(b)}{c - b}. \tag{1.11}$$

Demonstração. Como $a, b, c \in I$ e φ é convexa com $a < b < c$, temos

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

com $\frac{c-b}{c-a} < 1$ e $\frac{b-a}{c-a} < 1$. Então pela convexidade de φ temos:

$$\varphi(b) \leqslant \frac{c - b}{c - a}\varphi(a) + \frac{b - a}{c - a}\varphi(c)$$

e

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

Então $\frac{\varphi(b)-\varphi(a)}{b-a} \leqslant \frac{\varphi(c)-\varphi(a)}{c-a}$. A segunda desigualdade é obtida de modo análogo. □

Proposição 1.3.2 ([17])**.** *Sejam* $\epsilon > 0$ *e* $\lambda \in [0, 1]$*. Se* $\varphi : [0, \epsilon) \to \mathbb{R}$ *é convexa, então* $l : (0, \epsilon) \to \mathbb{R}$ *definida por*

$$l(x) = \frac{\varphi(x) - \varphi(\lambda x)}{x} \tag{1.12}$$

é não decrescente.

Demonstração. Dados $x, y \in (0, \epsilon)$ com $0 < x < y$. Se $\lambda = 1$, então $l(x) = 0$. Se $\lambda = 0$, então $l(x) = \frac{\varphi(x)-\varphi(0)}{x}$ e o resultado segue como consequência da primeira desigualdade (1.11) com $a = 0$, $b = x$ e $c = y$. Agora suponha que $\lambda \in (0, 1)$, então $\lambda < 1 \Rightarrow \lambda x < x < y$ e aplicando novamente a primeira desigualdade (1.11) temos:

$$\frac{\varphi(x) - \varphi(\lambda x)}{x - \lambda x} \leqslant \frac{\varphi(y) - \varphi(\lambda x)}{y - \lambda x}.$$

Novamente $\lambda < 1 \Rightarrow \lambda x < \lambda y < y$ e aplicando a segunda desigualdade (1.11) obtemos:

$$\frac{\varphi(y) - \varphi(\lambda x)}{y - \lambda x} \leqslant \frac{\varphi(y) - \varphi(\lambda y)}{y - \lambda y}$$

e, juntando as duas desigualdades, obtemos:

$$\frac{\varphi(x)-\varphi(\lambda x)}{x-\lambda x} \leqslant \frac{\varphi(y)-\varphi(\lambda x)}{y-\lambda x} \leqslant \frac{\varphi(y)-\varphi(\lambda y)}{y-\lambda y}.$$

Portanto $l(x) \leqslant l(y)$ para todo $x < y \in (0, \epsilon)$. □

Proposição 1.3.3 ([32]). *Sejam* $I \subset \mathbb{R}$ *um intervalo e* $\varphi : I \to \mathbb{R}$ *uma função diferenciável. Então* φ *é convexa se, e somente se,*

$$\varphi(y) \geqslant \varphi(x) + \varphi'(x)(y-x), \forall\ y \in I, x \in \mathrm{int}(I). \tag{1.13}$$

A função φ *é estritamente convexa se, e somente se, vale a desigualdade estrita.*

Demonstração. ($\Rightarrow$) Dados $y \in I$ e $x \in \mathrm{int}(I)$ temos por hipótese que

$$\varphi(\lambda y + (1-\lambda)x) \leqslant \lambda\varphi(y) + (1-\lambda)\varphi(x), \ \ \forall\ \lambda \in (0,1).$$

Então,

$$\frac{\varphi(x+\lambda(y-x)) - \varphi(x)}{\lambda} \leqslant \varphi(y) - \varphi(x).$$

Fazendo $\lambda \to 0$ obtemos

$$\varphi'(x)(y-x) + \varphi(x) \leqslant \varphi(y),$$

que prova a primeira parte.

($\Leftarrow$) Dados $x, y \in I$. Se $\lambda = 0$ e $\lambda = 1$ a convexidade é imediata.

Suponha que $\lambda \in (0,1)$ e seja $z = \lambda x + (1-\lambda)y \in \mathrm{int}(I)$ então de (1.13) temos

$$\varphi(\lambda x + (1-\lambda)y)\lambda(y-x) + \varphi(\lambda x + (1-\lambda)y) \leqslant \varphi(y) \tag{1.14}$$

e

$$\varphi(\lambda x + (1-\lambda)y)(1-\lambda)(x-y) + \varphi(\lambda x + (1-\lambda)y) \leqslant \varphi(x). \tag{1.15}$$

Multiplicando (1.14) por $(1-\lambda)$, (1.15) por λ e adicionando os resultados obtemos que

$$\varphi(\lambda x + (1-\lambda)y) \leqslant \lambda\varphi(x) + (1-\lambda)\varphi(y) \tag{1.16}$$

e portanto φ é convexa. □

Proposição 1.3.4 ([32]). *Sejam* $I \subset \mathbb{R}$ *um intervalo e* $\varphi : I \to \mathbb{R}$ *uma função diferenciável. Então* φ *é convexa se, e somente se,*

$$(\varphi'(x) - \varphi'(y))(x-y) \geqslant 0, \ \ \forall\ x, y \in \mathrm{int}(I).$$

Demonstração. ($\Rightarrow$) Tome $x, y \in \text{int}(I)$, como φ é convexa em I temos pela proposição anterior que

$$\varphi'(x)(y-x)+\varphi(x) \leqslant \varphi(y) \;\; e \;\; \varphi'(y)(x-y)+\varphi(y) \leqslant \varphi(x) \;\; \forall \;\; x,y \in \text{int}(I). \quad (1.17)$$

Então segue de (1.17) que

$$\varphi(x)+\varphi'(x)(y-x) \leqslant \varphi(y) \leqslant \varphi(x)+\varphi'(y)(y-x)$$

e isto implica que

$$(\varphi'(x)-\varphi'(y))(x-y) \geqslant 0, \;\; \forall \;\; x,y \in \text{int}(I).$$

($\Leftarrow$) O resultado é óbvio para $x = y$.
Sejam $x, y \in I$ com $x \neq y$ e $\lambda \in [0,1]$, temos pelo Teorema do Valor Médio que, para todo $a \in (x, y)$, existem c' e c'' com $c' \in (x, a)$ e $c'' \in (a, y)$ tais que

$$\varphi(a)-\varphi(x) = \varphi'(c')(a-x) \text{ e } \varphi(y)-\varphi(a) = \varphi'(c'')(y-a).$$

Multiplicando a primeira igualdade acima por $-\lambda$, a segunda por $(1-\lambda)$, tomando $a = \lambda x + (1-\lambda)y$ e somando o resultado obtemos depois de algumas manipulações algébricas que

$$\lambda\varphi(x)+(1-\lambda)\varphi(y)-\varphi(\lambda x+(1-\lambda)y) = \varphi'(c'')(y-x)(1-\lambda)\lambda-\lambda\varphi'(c')(1-\lambda)(y-x).$$

Note que existe $k > 0$ tal que $c''-c' = k(y-x)$, ou seja, $y-x = \frac{c''-c'}{k}$. Assim a igualdade anterior se reduz a

$$\lambda\varphi(x)+(1-\lambda)\varphi(y)-\varphi(\lambda x+(1-\lambda)y) = \frac{1-\lambda}{k}\lambda(\varphi'(c'')-\varphi'(c'))(c''-c').$$

Como $k > 0$ e $\lambda \in [0,1]$ segue da igualdade anterior e de

$$(\varphi'(x)-\varphi'(y))(x-y) \geqslant 0, \;\; \forall \;\; x,y \in \text{int}(I)$$

que

$$\varphi(\lambda x+(1-\lambda)y) \leqslant \lambda\varphi(x)+(1-\lambda)\varphi(y),$$

e portanto φ é convexa. $\square$

1.4 O Teorema de Newton-Kantorovich

O Teorema de Kantorovich assume condições semi-locais para garantir existência e unicidade de uma solução da equação não-linear $F(x) = 0$, onde F é uma aplicação diferenciável entre espaços de Banach. Este teorema usa construtivamente o método de Newton e também garante a convergência para uma solução deste processo iterativo. Além de sua elegância, ele tem muitas aplicações teóricas e práticas. Kantorovich forneceu basicamente duas provas diferentes deste resultado, primeiro usando relações de recorrência [30] e segundo por funções majorantes [22]. O propósito desta seção é apresentar uma prova modificada do segundo método. Para mais detalhes veja [29].

Teorema 1.4.1. *Sejam X e Y espaços de Banach e $F : D \subset X \to Y$. Suponha que em um subconjunto aberto convexo $D_0 \subset D$, F é Fréchet Diferenciável e*

$$\|F'(x) - F'(y)\| \leqslant K\|x - y\|, \quad \forall \ x, y \in D_0. \tag{1.18}$$

Para algum $x^0 \in D_0$, suponha que $\Gamma_0 \equiv [F'(x^0)]^{-1}$ é definida em todo Y e que $h \equiv \beta K \eta \leqslant \frac{1}{2}$ onde $\|\Gamma_0\| \leqslant \beta$ e $\|\Gamma_0 F x^0\| \leqslant \eta$. Seja

$$t_* = \frac{1}{\beta K}(1 - \sqrt{1 - 2h}), \quad t_{**} = \frac{1}{\beta K}(1 + \sqrt{1 - 2h}) \tag{1.19}$$

e suponha que $S \equiv \{x; \|x - x^0\| \leqslant t_*\} \subset D_0$. Então as iterações de Newton $x^{k+1} = x^k - F'(x^k)^{-1} F x^k$ estão bem definidas, estão contidas em S e convergem para uma solução x^* de $F(x) = 0$ que é única em $D_0 \cap \{x; \|x - x^0\| \leqslant t_{**}\}$. Se $h < \frac{1}{2}$ então a ordem de convergência é no mínimo quadrática.

A prova segue como consequência do Lema 1.1.1 e dos seguintes lemas.

Lema 1.4.1. *Para todo $x \in Q \equiv \{x; \|x - x^0\| < 1/\beta K\} \cap D_0$, $F'(x)^{-1}$ é definido em todo Y e*

$$\|[F'(x)]^{-1}\| \leqslant \frac{\beta}{(1 - \beta K \|x - x^0\|)}. \tag{1.20}$$

Se x e $N(x) \equiv x - [F'(x)]^{-1} F(x)$ estão em Q, então

$$\|N(Nx) - Nx\| \leqslant \frac{1}{2} \frac{\beta K \|x - Nx\|^2}{1 - \beta K \|x^0 - Nx\|}. \tag{1.21}$$

Demonstração. A primeira afirmação segue a partir do Lema de Banach. Para provar a segunda parte note primeiro que $F(x) + F'(x)(Nx - x) = 0$ e

$$\begin{aligned} \|N(Nx) - Nx\| &= \|Nx - F'(Nx)^{-1}F(Nx) - Nx\| \\ &= \|F'(Nx)^{-1}F(Nx)\| \\ &\leqslant \tfrac{\beta}{1-\beta K\|x^0 - Nx\|}\|F(Nx)\| \\ &= \tfrac{\beta}{1-\beta K\|x^0 - Nx\|}\|F(Nx) - Fx - F'(x)(N(x) - x)\|, \end{aligned}$$

logo pela Proposição 1.2.4

$$\|F(Nx) - Fx - F'(x)(N(x) - x)\| \leqslant \frac{K}{2}\|Nx - x\|^2$$

e portanto

$$\|N(Nx) - Nx\| \leqslant \frac{1}{2}\frac{\beta K\|x - Nx\|^2}{1 - \beta K\|x^0 - Nx\|}.$$

□

Lema 1.4.2. *A sequência de Newton* $\{x^k\}$ *está bem definida e é majorada pela sequência definida por*

$$t_{k+1} = t_k - \frac{(\beta K/2)t_k^2 - t_k + \eta}{\beta K t_k - 1}, \quad t_0 = 0. \tag{1.22}$$

Além disso, $t_k \to t_*$*, onde* t_* *é definido por (1.19).*

Demonstração. Primeiro veja que t_k são as iteradas de Newton para o polinômio $\frac{\beta K}{2}t^2 - t + \eta$ com raízes t_* e t_{**}, então pelo método de Newton $t_k \to t_*$.

Agora suponha que existem x^1, ..., x^k e $\|x^i - x^{i-1}\| \leqslant t_i - t_{i-1}$, $i = 1, ..., k$. Isto vale para $k = 1$. E, mais ainda,

$\|x^k - x^0\| \leqslant \sum_{i=1}^k \|x^i - x^{i-1}\| \leqslant \sum_{i=1}^k (t_i - t_{i-1}) = t_k - t_0 \leqslant t_*$. Assim $\{x^k\} \in S$ e pelo Lema (1.4.1) temos:

$$\begin{aligned} \|x^{k+1} - x^k\| &= \|N(Nx^{k-1}) - Nx^{k-1}\| \\ &\leqslant \tfrac{1}{2}\tfrac{\beta K\|x^k - x^{k-1}\|^2}{1-\beta K\|x^0 - x^k\|} \\ &\leqslant \tfrac{1}{2}\tfrac{\beta K(t_k - t_{k-1})^2}{1-\beta K t_k} \\ &= t_{k+1} - t_k. \end{aligned}$$

□

Agora vamos provar o Teorema de Newton-Kantorovick.

Demonstração. Segue dos Lemas 1.1.1 e 1.4.2 que existe um $x^* \in S$ tal que $x^k \to x^*$. Que x^* é a solução segue a partir de

$$\begin{aligned}
\|F(x^k)\| &= \|F'(x^k)(x^{k+1} - x^k)\| \\
&= \|F'(x^k)(x^{k+1} - x^k) - F'(x^0)(x^{k+1} - x^k) + F'(x^0)(x^{k+1} - x^k)\| \\
&\leqslant [\|F'(x^0)\| + \|F'(x^0) - F'(x^k)\|]\|x^{k+1} - x^k\| \\
&\leqslant [\|F'(x^0)\| + Lt_*]\|x^{k+1} - x^k\|.
\end{aligned}$$

Portanto da continuidade de F em S e $\|x^{k+1} - x^k\| \to 0$, concluímos que $\|F(x^k)\| \to 0$ quando $k \to \infty$. Desse modo, $F(x^*) = 0$ e a prova do Teorema está concluída. $\square$

Capítulo 2

Convergência Local do Método de Newton sob a Condição Majorante

Neste Capítulo serão apresentados resultados obtidos em [12] que generalizam o Teorema 2.1 de [10] e, consequentemente, fornecem o raio de convergência da sequência gerada pelo método. Conforme [12], sob certas condições as sequências geradas pelo método de Newton estão bem definidas e convergem para uma solução do problema em foco. Encerramos apresentando duas aplicações. Na primeira, um resultado de convergência para operador do tipo-Hölder e a segunda, um resultado de convergência para o operador satisfazendo uma condição de Lipschitz generalizada.

2.1 Análise Local para o Método de Newton

Teorema 2.1.1. *Sejam* X, Y *espaços de Banach,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função continuamente diferenciável. Seja* $x^* \in \Omega$, $R > 0$ *e* $\kappa := \sup\{t \in [0, R) : B(x^*, t) \subset \Omega\}$. *Suponha que* $F(x^*) = 0$, $F'(x^*)$ *é invertível e que existe* $f : [0, R) \to \mathbb{R}$ *continuamente diferenciável tal que*

$$\|F'(x^*)^{-1}[F'(x) - F'(x^* + \tau(x - x^*))]\| \leqslant f'(\|x - x^*\|) - f'(\tau\|x - x^*\|), \tag{2.1}$$

para todo $\tau \in [0, 1]$, $x \in B(x^*, \kappa)$ *e*

(h_1) $f(0) = 0$ *e* $f'(0) = -1$*;*

(h_2) f' *é estritamente crescente.*

Sejam

$$\nu := \sup\{t \in [0, R) : f'(t) < 0\},$$

$$\rho := \sup\left\{\delta \in (0,\nu) : \left(\frac{f(t)}{f'(t)} - t\right)\frac{1}{t} < 1, \;\; t \in (0,\delta)\right\}$$

e

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

Então as sequências com pontos inicial $x^0 \in B(x^*, r)\backslash\{x^*\}$ *e* $t_0 = \|x^0 - x^*\|$, *definidas respectivamente por*

$$x^{k+1} = x^k - F'(x^k)^{-1}F(x^k), \qquad\qquad t_{k+1} = \left|t_k - \frac{f(t_k)}{f'(t_k)}\right|, \quad k = 0, 1, ... \tag{2.2}$$

estão bem definidas; $\{t_k\}$ *é estritamente decrescente, está contida em* $(0, r)$ *e converge para* 0 *e,* $\{x^k\}$ *está contida em* $B(x^*, r)$ *e converge para* x^* *que é o único zero de* F *em* $B(x^*, \sigma)$, *onde* $\sigma := \sup\{t \in (0,\kappa) : f(t) < 0\}$ *e vale:*

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0 \qquad\qquad \lim_{k\to\infty} \frac{t_{k+1}}{t_k} = 0. \tag{2.3}$$

Por outro lado, se $\frac{f(\rho)}{\rho f'(\rho)} - 1 = 1$ *e* $\rho < \kappa$ *então* $r = \rho$ *é o melhor raio de convergência possível.*

Além disso, tomando $0 \leqslant p \leqslant 1$

(h_3) *a função* $(0,\nu) \ni t \mapsto \left[\frac{f(t)}{f'(t)} - t\right]\frac{1}{t^{p+1}}$ *é estritamente crescente,*

então a sequência $\{\frac{t_{k+1}}{t_k^{p+1}}\}$ *é estritamente decrescente e vale*

$$\|x^{k+1} - x^*\| \leqslant \left[\frac{t_{k+1}}{t_k^{p+1}}\right]\|x^k - x^*\|^{p+1}. \tag{2.4}$$

Observação 2.1.1. *A primeira equação em (2.3) diz que* $\{x^k\}$ *converge superlinearmente para* x^*. *Além disso, por que a sequência* $\{t_{k+1}/t_k^{p+1}\}$ *é estritamente decrescente temos* $t_{k+1}/t_k^{p+1} \leqslant t_1/t_0^{p+1}$, *para* $k = 0, 1, ...$ *Assim a desigualdade em (2.4) implica que* $\|x^{k+1} - x^*\| \leqslant [t_1/t_0^{p+1}]\|x^k - x^*\|^{p+1}$, *para* $k = 0, 1, ...$ *Como consequência, se* $p = 0$ *então* $\|x^k - x^*\| \leqslant t_0[t_1/t_0]^k$ *para* $k = 0, 1, ...$ *e se* $0 < p \leqslant 1$ *então*

$$\|x^k - x^*\| \leqslant t_0(t_1/t_0)^{[(p+1)^k - 1]/p}, k = 0, 1, ...$$

2.2 Casos Especiais

Nesta seção vamos apresentar, como aplicações, dois casos especiais do Teorema 2.1.1. Primeiramente vamos mostrar um teorema de convergência para o método de Newton sob uma condição invariante afim do tipo-Hölder que apareceu em [19] e [36].

2.2.1 Resultado de Convergência sob uma Condição tipo-Hölder

Teorema 2.2.1. *Sejam* X, Y *espaços de Banach,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função continuamente diferenciável. Sejam* $x^* \in \Omega$, $R > 0$ *e* $\kappa := \sup\{t \in [0, R) : B(x^*, t) \subset \Omega\}$. *Suponha que* $F(x^*) = 0$, $F'(x^*)$ *é invertível e que existe uma constante* $K > 0$ *e* $0 < p \leqslant 1$ *tal que*

$$\|F'(x^*)^{-1}[F'(x) - F'(x^* + \tau(x - x^*))]\| \leqslant K(1 - \tau^p)\|x - x^*\|^p, \tag{2.5}$$

para todo $\tau \in [0, 1]$, $x \in B(x^*, \kappa)$.
Seja $r = \min\left\{\kappa, (\frac{p+1}{(2p+1)K})^{1/p}\right\}$. *Então, as sequências com ponto inicial* $x^0 \in B(x^*, r)\backslash\{x^*\}$ *e* $t_0 = \|x^0 - x^*\|$ *definidas respectivamente por,*

$$x^{k+1} = x^k - F'(x^k)^{-1}F(x^k), \quad t_{k+1} = \frac{Kpt_k^{p+1}}{(p+1)(1 - Kt_k^p)}, \ k = 0, 1, ... \tag{2.6}$$

estão bem definidas, $\{t_k\}$ *é estritamente decrescente, está contida em* $(0, r)$ *e converge para 0,* $\{x^k\}$ *está contida em* $B(x^*, r)$, *converge para* x^*, *que é o único zero de F em* $B(x^*, [(p + 1)/K]^{1/p})$ *e vale*

$$\|x^{k+1} - x^*\| \leqslant \frac{Kp}{(p + 1)(1 - Kt_k^p)}\|x^k - x^*\|^{p+1}.$$

Além disso, se

$$\left[\frac{p + 1}{(2p + 1)K}\right]^{1/p} < \kappa$$

então $r = \left[\frac{p+1}{(2p+1)K}\right]^{1/p}$ *é o melhor raio de convergência possível.*

Observação 2.2.1. *Desde que o Teorema 2.2.1 é um caso especial do Teorema 2.1.1 segue a partir da Observação 2.1.1 que*

$$\|x^k - x^*\| \leqslant \left[\frac{Kp\|x^0 - x^*\|^p}{(p + 1)[1 - K\|x^0 - x^*\|^p]}\right]^{[(p+1)^k - 1]/p} \|x^0 - x^*\|.$$

2.2.2 Resultado de Convergência sob uma Condição de Lipschitz Generalizada

Nesta subseção, iremos apresentar um teorema de convergência local do método de Newton sob a condição de Lipschitz generalizada devido a X. Wang, que apareceu em [36] e também em [35]. Vale a pena ressaltar que o resultado desta subseção não assume que a função que define a condição de Lipschitz generalizada é não decrescente.

Teorema 2.2.2. *Sejam* X, Y *espaços de Banach,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função continuamente diferenciável. Sejam* $x^* \in \Omega$, $R > 0$ *e* $\kappa := \sup\{t \in [0, R) : B(x^*, t) \subset \Omega\}$. *Suponha que* $F(x^*) = 0$, $F'(x^*)$ *é invertível e que existe uma função integrável positiva* $L : [0, R) \to \mathbb{R}$ *tal que*

$$\|F'(x^*)^{-1}[F'(x) - F'(x^* - \tau(x - x^*))]\| \leqslant \int_{\tau\|x-x^*\|}^{\|x-x^*\|} L(u)du, \tag{2.7}$$

para todo $\tau \in [0, 1]$, $x \in B(x^*, \kappa)$. *Seja* $\bar{\nu} > 0$ *a constante definida por*

$$\bar{\nu} := \sup\left\{t \in [0, R) : \int_0^t L(u)du - 1 < 0\right\}, \tag{2.8}$$

e seja $\bar{\rho} > 0$ *e* $\bar{r} > 0$ *as constantes definidas por*

$$\bar{\rho} := \sup\left\{t \in (0, \delta) : \int_0^t L(u)udu/t(1 - \int_0^t L(u)du) < 1, t \in (0, \delta)\right\}, \tag{2.9}$$

$$\bar{r} = \min\{\kappa, \bar{\rho}\}. \tag{2.10}$$

Então, as sequências com ponto inicial $x^0 \in B(x^*, \bar{r})\backslash\{x^*\}$ *e* $t_0 = \|x^0 - x^*\|$, *definidas respectivamente por*

$$x^{k+1} = x^k - F'(x^k)^{-1}F(x^k), \ t_{k+1} = \int_0^{t_k} L(u)udu/\left(1 - \int_0^{t_k} L(u)du\right) \tag{2.11}$$

estão bem definidas, $\{t_k\}$ *é estritamente decrescente, está contida em* $(0, \bar{r})$ *e converge para 0,* $\{x^k\}$ *está contida em* $B(x^*, \bar{r})$, *converge para* x^* *que é o único zero de* F *em* $B(x^*, \bar{\sigma})$, *onde*

$$\bar{\sigma} := \sup\left\{t \in (0, \kappa) : \int_0^t L(u)(t - u)du - t < 0\right\} \tag{2.12}$$

e valem:

$$\lim_{k\to\infty} \frac{t_{k+1}}{t_k} = 0$$

e

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0.$$

Além disso, se

$$\frac{\int_0^{\bar{\rho}} L(u)udu}{\bar{\rho}(1 - \int_0^{\bar{\rho}} L(u)du)} = 1, \tag{2.13}$$

e $\bar{\rho} < \kappa$ *então* $\bar{r} = \bar{\rho}$ *é o melhor raio de convergência possível.*

Se, adicionando a hipótese $0 \leqslant p \leqslant 1$

(h) *a função* $(0, \nu) \ni t \mapsto t^{1-p}L(t)$ *é não decrescente,*

então a sequência $\{\frac{t_{k+1}}{t_k^{p+1}}\}$ *é estritamente decrescente e vale*

$$\|x^{k+1} - x^*\| \leqslant \frac{t_{k+1}}{t_k^{p+1}}\|x^k - x^*\|^{p+1}, \ \ k = 0, 1, ... \tag{2.14}$$

Capítulo 3

Melhorias na Análise de Convergência

O próximo resultado usa as mesmas informações (x^*, F, f) que o Teorema 2.1.1, uma vez que f_0 introduzido em (3.1) é apenas um caso especial de f e que sua existência é garantida na continuidade Lipschitz geral, quando $\tau = 0$. Vamos mostrar a partir de agora que a nova sequência escalar $\{s_k\}$ converge a zero mais rápido que a sequência escalar $\{t_k\}$ e, através de casos especiais, que o raio de convergência aqui é maior que aquele apresentado no capítulo anterior.

3.1 Análise Local para o Método de Newton

Teorema 3.1.1. *Sejam* X, Y *espaços de Banach,* $\Omega \subseteq X$ *um conjunto aberto e convexo e* $F : \Omega \to Y$ *Fréchet-diferenciável. Seja* $x^* \in \Omega$, $R > 0$ *e* $\kappa := \sup\{t \in [0, R) : B(x^*, t) \subset \Omega\}$. *Suponha que* $F(x^*) = 0$, $F'(x^*)$ *é invertível e que existem* $f, f_0 : [0, R) \to \mathbb{R}$ *continuamente diferenciáveis tais que:*

$$\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leqslant f_0'(\|x - x^*\|) - f_0'(0), \tag{3.1}$$

$$\|F'(x^*)^{-1}[F'(x) - F'(x^* + \tau(x - x^*))]\| \leqslant f'(\|x - x^*\|) - f'(\tau\|x - x^*\|), \tag{3.2}$$

para todo $\tau \in [0, 1]$, $x \in B(x^*, \kappa)$ *e*

(H_1) $f_0(0) = f(0) = 0$ *e* $f_0'(0) = f'(0) = -1$*;*

(H_2) f_0', f' *são estritamente crescentes satisfazendo*

$$f_0(t) \leqslant f(t) \quad \text{e} \quad f_0'(t) \leqslant f'(t) \quad t \in [0, R), \tag{3.3}$$

Considere também

$$\nu_0 := \sup\{t \in [0, R) : f_0'(t) < 0\}, \tag{3.4}$$

$$\nu := \sup\{t \in [0, R) : f'(t) < 0\}, \tag{3.5}$$

$$f_1(t) := \frac{f'(t)}{f_0'(t)}, \tag{3.6}$$

$$\rho_0 := \sup\left\{\delta \in [0, \nu) : \left(\frac{f(t)}{f'(t)} - t\right)\frac{f_1(t)}{t} < 1, t \in [0, \delta)\right\}, \tag{3.7}$$

$$r_0 := \min\{\kappa, \rho_0\}, \tag{3.8}$$

$$s_0 = \|x^0 - x^*\|, \quad s_{k+1} = \left|\left(s_k - \frac{f(s_k)}{f'(s_k)}\right) f_1(s_k)\right| \quad (k \geqslant 0). \tag{3.9}$$

Então as seguintes afirmações são verdadeiras:

(a) $\{s_k\}$ *está bem definida, é estritamente decrescente, está contida em* $(0, r_0)$*, converge para zero e*

$$\lim_{k\to\infty} \frac{s_{k+1}}{s_k} = 0. \tag{3.10}$$

(b) A sequência $\{x^k\}$ *gerada pelo método de Newton, com ponto inicial* $x^0 \in B(x^*, r_0)$*, está bem definida, está contida em* $B(x^*, r_0)$ *e converge para* x^**, que é a única solução de* $F(x) = 0$ *em* $B(x^*, \sigma_0)$*, onde*

$$\sigma_0 := \sup\{t \in [0, \kappa) : f_0(t) < 0\} \tag{3.11}$$

e

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^k\|}{\|x^k - x^*\|} = 0. \tag{3.12}$$

(c) Se

$$\left(\frac{f(\rho_0)}{\rho_0 f'(\rho_0)} - 1\right) f_1(\rho_0) = 1 \ \text{ e } \ \rho_0 < \kappa \tag{3.13}$$

então $r_0 = \rho_0$ *é o melhor raio de convergência possível.*

(d) Se a sequência escalar $\{t_k\}$ *é dada por*

$$t_0 = \|x^0 - x^*\|, \quad t_{k+1} = \left|t_k - \frac{f(t_k)}{f'(t_k)}\right| \tag{3.14}$$

então

$$s_k \leqslant t_k \ \text{ para todo } \ k \geqslant 0 \tag{3.15}$$

e a desigualdade estrita vale para $k > 1$ *se* $f_0'(t) < f'(t),\ t \in [0, R)$*.*

Se adicionarmos, $0 \leqslant p \leqslant 1$

(H_3) *A função* $t \to \left(\frac{f(t)}{f'(t)} - t\right)\frac{f_1(t)}{t^{p+1}}$ *é estritamente crescente em* $(0, \nu_0)$*, então*
(e) A sequência $\{\frac{s_{k+1}}{s_k^{p+1}}\}$ *é estritamente decrescente, assim*

$$\|x^{k+1} - x^*\| \leqslant \left\{\frac{s_{k+1}}{s_k^{p+1}}\right\}\|x^k - x^*\|^{p+1}. \tag{3.16}$$

Além disso, para $k \geqslant 0$,

$$\|x^k - x^*\| \leqslant \begin{cases} s_0[\frac{s_1}{s_0}]^k, & \textit{se} \quad p = 0 \\ s_0(\frac{s_1}{s_0})^{\frac{(p+1)^k - 1}{p}}, & \textit{se} \quad p \neq 0 \end{cases}$$

A prova do teorema 3.1.1 é baseada nos seguintes lemas:

Lema 3.1.1. *As constantes* κ, ν, σ_0 *são positivas e* $(t - \frac{f(t)}{f'(t)})f_1(t) < 0$ *para todo* $t \in (0, \nu)$.

Demonstração. Como Ω é aberto e $x^* \in \Omega$, temos que κ é positivo. Agora veja que como f' é contínua em $[0, R)$ e $f'(0) = -1$, então existe $\delta > 0$ tal que $f'(t) < 0$ para todo $t \in (0, \delta)$ e como $\nu := \sup\{t \in [0, R) : f'(t) < 0\}$ isto implica que $\nu \geqslant \delta > 0$. Desde que $f(0) = 0$ e $f'(0) = -1$, existe $\delta > 0$ tal que $f(t) < 0$ para todo $t \in (0, \delta)$. Portanto, temos que $\sigma = \sup\{t \in [0, k) : f(t) < 0\} > 0$ e então, por (H_2), $\sigma_0 \geqslant \sigma > 0$, $t \in (0, \sigma_0)$.
De (H_1), (H_2) e Proposição 1.3.3 temos que $0 = f(0) > f(t) - tf'(t)$ para todo $t \in [0, R)$. Se $t \in (0, \nu)$ então $f'(t) < 0$. Logo dividindo a última expressão por $f'(t)$ obtemos $t - \frac{f(t)}{f'(t)} < 0$ e, como $f_1(t) > 0$ para todo $t \in (0, \nu)$, segue o lema. $\square$

Segue de (H_2), das definições de ν_0 e ν que $f_0'(t) < 0$, $f'(t) < 0$ para todo $t \in [0, \nu)$, desde que $\nu \leqslant \nu_0$. Assim, a função f_1 está bem definida em $(0, \nu_0)$. Portanto, a função iteração de Newton

$$n_{f,f_0} : [0, \nu) \to (-\infty, 0] \tag{3.17}$$

$$t \to \left(t - \frac{f(t)}{f'(t)}\right) f_1(t) \tag{3.18}$$

está bem definida.

Lema 3.1.2. *As seguintes afirmações são verdadeiras:*

$$\lim_{t \to 0} \frac{|n_{f,f_0}(t)|}{t} = 0 \tag{3.19}$$

$$\rho_0 > 0 \tag{3.20}$$

e

$$|n_{f,f_0}(t)| < t \ \ \forall \ \ t \in (0, \rho_0). \tag{3.21}$$

Demonstração.

$$\begin{aligned}
\frac{|n_{f,f_0}(t)|}{t} &= \left[\frac{f(t)}{f'(t)} - t\right]\frac{f_1(t)}{t}\\
&= \left(\frac{f(t)-tf'(t)}{tf'(t)}\right) f_1(t)\\
&= \left(\frac{1}{f'(t)}\frac{f(t)}{t} - 1\right) f_1(t)\\
&= \left(\frac{1}{f'(t)}\frac{f(t)-f(0)}{t-0} - 1\right) f_1(t) \ \ \forall \ \ t \in (0, \nu).
\end{aligned}$$

Como f é continuamente diferenciável, $f'(0) = -1$ e $f_1(0) = 1$, obtemos fazendo $t \to 0$ que $\lim\limits_{t\to 0} \dfrac{|n_{f,f_0}(t)|}{t} = 0$. Logo existe $\delta > 0$ tal que

$$\left(\frac{f(t)}{f'(t)} - t\right)\frac{f_1(t)}{t} < 1 \ \ \forall \ \ t \in (0, \delta).$$

Portanto, da definição de ρ_0, concluímos que $\rho_0 > 0$ e $|n_{f,f_0}(t)| < t \ \ \forall \ \ t \in (0, \rho_0)$. □

É fácil ver que a sequência $\{s_k\}$ pode ser definida como:

$$s_0 = \|x^0 - x^*\|, \quad s_{k+1} = |n_{f,f_0}(s_k)|. \tag{3.22}$$

Lema 3.1.3. *A sequência* $\{s_k\}$ *está bem definida, é estritamente decrescente e está contida em* $(0, \rho_0)$*. Alem disso,* $\{s_k\}$ *converge para 0 com taxa de convergência superlinear, isto é,* $\lim\limits_{k\to\infty} \dfrac{s_{k+1}}{s_k} = 0$*. Se adicionarmos* (H_3)*, então a sequência* $\{\frac{s_{k+1}}{s_k^{p+1}}\}$ *é estritamente decrescente.*

Demonstração. Como $0 < s_0 = \|x^0 - x^*\| < r \leqslant \rho_0$, concluímos a partir do Lema 3.1.2 e (3.17) que a sequência $\{s_k\}$ está bem definida. Como $s_{k+1} = |n_{f,f_0}(s_k)|$ temos do Lema 3.1.2 que $s_{k+1} = |n_{f,f_0}(s_k)| < s_k$ e portanto $\{s_k\}$ é estritamente decrescente e está contida em $(0, \rho_0)$.

Como $\{s_k\} \subset (0, \rho_0)$ e é estritamente decrescente segue que $\{s_k\}$ é convergente. Assim, $\lim\limits_{k\to+\infty} s_k = s^*$ com $0 \leqslant s^* < \rho_0$.

De $s_{k+1} = |n_{f,f_0}(s_k)|$ e da continuidade de n_{f,f_0}, obtemos fazendo $k \to \infty$ que $0 \leqslant s^* = |n_{f,f_0}(s^*)|$. Se $s^* \neq 0$, i.e, $s^* \in (0, \rho_0)$, então do lema anterior segue que $|n_{f,f_0}(s^*)| < s^*$, que gera uma contradição. Portanto $\lim\limits_{k\to+\infty} s_k = 0$. Agora veja que

$$\lim_{k\to+\infty} \frac{s_{k+1}}{s_k} = \lim_{k\to+\infty} \frac{|n_{f,f_0}(s_k)|}{s_k} = 0$$

pelo lema anterior. Desde que $\{s_k\}$ é estritamente decrescente, a última afirmação segue imediatamente como consequência de (H_3). □

Observação 3.1.1. *Para ver que as sequências* $\{s_k\}$ *e* $\{t_k\}$ *satisfazem* $s_k \leqslant t_k$*, note que* $s_0 = \|x^0 - x^*\| = t_0$*, que* $|f_1(\theta)| \leqslant 1$*, e, por conseguinte, que as funções de iteração* $|n_f|$ *e* $|n_{f,f_0}|$ *satisfazem* $|n_{f,f_0}(\theta)| \leqslant |n_f(\theta)|$ *para* $\theta \in [0, \nu_0)$*.*

3.2 Relação entre a Função Majorante e o Operador Não Linear

Lema 3.2.1. *Se* $x \in B(x^*, t)$, $t \in [0, \min\{k, \nu_0\})$, $\|x - x^*\| \leqslant \min\{k, \nu_0\}$, *então* $F'(x)^{-1}$ *existe e*

$$\|F'(x)^{-1}F'(x^*)\| \leqslant -\frac{1}{f_0'(\|x - x^*\|)} \leqslant -\frac{1}{f_0'(t)}. \tag{3.23}$$

Demonstração. Seja $x \in B(x^*, t)$ e $t \in [0, \min\{k, \nu_0\})$. Usando que $f_0'(0) = -1$, (3.1) e que f_0' é estritamente crescente, obtemos

$$\begin{aligned} \|F'(x^*)^{-1}(F'(x) - F'(x^*))\| &\leqslant f_0'(\|x - x^*\|) - f_0'(0) \\ &= f_0'(\|x - x^*\|) + 1 \\ &\leqslant f_0'(t) + 1 < 1, \end{aligned}$$

onde a última desigualdade segue a partir das definições de k, ν_0 e da escolha de t. Assim, o Lema de Banach garante que $F'(x^*)^{-1}F'(x)$ é invertível e portanto $F'(x)$ também é. Também do Lema de Banach temos que

$$\begin{aligned} \|F'(x)^{-1}F'(x^*)\| &\leqslant \frac{1}{1 - \|F'(x^*)^{-1}(F'(x) - F'(x^*))\|} \\ &\leqslant \frac{1}{1 - (f_0'(\|x - x^*\|) - f_0'(0))} \\ &= -\frac{1}{f_0'(\|x - x^*\|)}, \end{aligned}$$

pois $f'(0) = -1$. Portanto

$$\|F'(x)^{-1}F'(x^*)\| \leqslant -\frac{1}{f_0'(\|x - x^*\|)} \leqslant -\frac{1}{f_0'(t)}. \tag{3.24}$$

□

Vamos estudar agora os erros de linearização do operador F e da função majorante f em um ponto de Ω, dados por:

$$E_F(x, y) := F(y) - [F(x) + F'(x)(y - x)], \quad x, y \in \Omega \tag{3.25}$$

$$e_f(t, u) := f(u) - [f(t) + f'(t)(u - t)], \quad t, u \in [0, R). \tag{3.26}$$

Lema 3.2.2. *Se* $\|x^* - x\| < \kappa$, *então* $\|F'(x^*)^{-1}E_F(x, x^*)\| \leqslant e_f(\|x - x^*\|, 0)$.

Demonstração. Desde que $B(x^*, \kappa)$ é convexa, temos que $x^* + \tau(x - x^*) \in B(x^*, \kappa)$ para $0 \leqslant \tau \leqslant 1$ e $x \in B(x^*, \kappa)$. Como F é continuamente diferenciável em Ω segue da definição de E_F e do Teorema Fundamental do Cálculo que:

$$\begin{aligned}
\|F'(x^*)^{-1}E_F(x, x^*)\| &= \|F'(x^*)^{-1}[F(x^*) - F(x) - F'(x)(x^* - x)]\| \\
&= \left\| \int_0^1 F'(x^*)^{-1}[F'(x) - F'(x^* + \tau(x - x^*))](x - x^*)d\tau \right\| \\
&\leqslant \int_0^1 \|F'(x^*)^{-1}[F'(x) - F'(x^* + \tau(x - x^*))]\| \|(x - x^*)\| d\tau.
\end{aligned}$$

Agora, do Teorema 3.1.1, do Teorema Fundamental do Cálculo e da definição de e_f segue que

$$\begin{aligned}
\|F'(x^*)^{-1}E_F(x, x^*)\| &\leqslant \int_0^1 [f'(\|x - x^*\|) - f'(\tau\|x - x^*\|)]\|x - x^*\|d\tau \\
&= f'(\|x - x^*\|)\|x - x^*\| - f(\|x - x^*\|) + f(0) \\
&= e_f(\|x - x^*\|, 0).
\end{aligned}$$

□

O Lema 3.2.1 garante, em particular, que F' é invertível em $B(x^*, r_0)$ e consequentemente:

$$\begin{aligned}
N_F : B(x^*, r_0) &\to Y \\
x &\mapsto x - F'(x)^{-1}F(x)
\end{aligned} \tag{3.27}$$

é um operador bem definido.

Lema 3.2.3. *Se* $\|x - x^*\| < r_0$ *então* $\|N_F(x) - x^*\| \leqslant |n_{f,f_0}(\|x - x^*\|)|$. *Como consequência,*

$$N_F(B(x^*, r_0)) \subset B(x^*, r_0).$$

Demonstração. Como $F(x^*) = 0$, então a desigualdade é direta para $x = x^*$. Suponha que $0 < \|x - x^*\| \leqslant r_0$. Assim o Lema 3.2.1 garante que $F'(x)$ é invertível. Novamente, de $F(x^*) = 0$ obtemos:

$$\begin{aligned}
x^* - N_F(x) &= x^* - x + F'(x)^{-1}F(x) \\
&= F'(x)^{-1}F'(x)(x^* - x) + F'(x)^{-1}F(x) \\
&= -F'(x)^{-1}[F(x^*) - F(x) - F'(x)(x^* - x)] \\
&= -F'(x)^{-1}E_F(x, x^*).
\end{aligned}$$

Então, dos Lemas 3.2.1 e 3.2.2, segue que

$$\begin{aligned}\|x^* - N_F(x)\| &\leqslant \|-F'(x)^{-1}F'(x^*)\|\|F'(x^*)^{-1}E_F(x,x^*)\| \\ &\leqslant \frac{e_f(\|x-x^*\|,0)}{|f_0'(\|x-x^*\|)|}.\end{aligned}$$

De $f(0) = 0$ e das definições de e_f e n_{f,f_0}, obtemos

$$\begin{aligned}\frac{e_f(\|x-x^*\|,0)}{|f_0'(\|x-x^*\|)|} &= \frac{f(0) - f(\|x-x^*\|) + f'(\|x-x^*\|)\|x-x^*\|}{-f_0'(\|x-x^*\|)} \\ &= \frac{f(\|x-x^*\|)}{f_0'(\|x-x^*\|)} - \frac{f'(\|x-x^*\|)}{f_0'(\|x-x^*\|)}\|x-x^*\| \\ &= \frac{f(\|x-x^*\|)}{f_0'(\|x-x^*\|)}\frac{f'(\|x-x^*\|)}{f'(\|x-x^*\|)} - \frac{f'(\|x-x^*\|)}{f_0'(\|x-x^*\|)}\|x-x^*\| \\ &= \left(\frac{f(\|x-x^*\|)}{f'(\|x-x^*\|)} - \|x-x^*\|\right)\frac{f'(\|x-x^*\|)}{f_0'(\|x-x^*\|)} \\ &= |n_{f,f_0}(\|x-x^*\|)|.\end{aligned}$$

Portanto $\|x^* - N_F(x)\| \leqslant |n_{f,f_0}(\|x-x^*\|)|$.

Agora tome $y \in N_F(B(x^*, r_0))$. Então existe $x \in B(x^*, r_0)$ tal que $N_F(x) = y$. Como $x \in B(x^*, r_0)$ implica que $\|x - x^*\| < r_0$ e $r \leqslant \rho_0$, a primeira parte desta proposição junto com a segunda parte do Lema 3.1.2 implica que

$$\|N_F(x) - x^*\| \leqslant |n_{f,f_0}(\|x-x^*\|)| < \|x-x^*\| < r_0$$

e, portanto, $N_F(B(x^*, r_0)) \subset B(x^*, r_0)$. □

Lema 3.2.4. *Supondo* (H_3) *e* $\|x - x^*\| \leqslant t < r_0$ *então*

$$\|N_F(x) - x^*\| \leqslant \frac{|n_{f,f_0}(t)|}{t^{p+1}}\|x-x^*\|^{p+1}. \tag{3.28}$$

Demonstração. A desigualdade é direta se $x = x^*$. Suponha que $0 < \|x - x^*\| \leqslant t$ e (H_3) vale. Então da definição de n_{f,f_0} temos que a aplicação $t \mapsto \frac{|n_{f,f_0}(t)|}{t^{p+1}}$ é estritamente crescente e consequentemente

$$\frac{|n_{f,f_0}(\|x-x^*\|)|}{\|x-x^*\|^{p+1}} \leqslant \frac{|n_{f,f_0}(t)|}{t^{p+1}}.$$

Portanto, do Lema 3.2.3, temos:

$$\|N_F(x) - x^*\| \leqslant |n_{f,f_0}(\|x-x^*\|)| \leqslant \frac{|n_{f,f_0}(t)|}{t^{p+1}}\|x-x^*\|^{p+1}.$$

□

3.3 Unicidade e Raio de Convergência Ótimo

Agora vamos obter a unicidade da solução e o raio de convergência ótimo.

Lema 3.3.1. *O ponto* x^* *é o único zero de* F *em* $B(x^*, \sigma_0)$.

Demonstração. Tome $y \in B(x_*, t)$ e $F(y) = 0$. Como $F(x_*) = 0$ e $F(y) = 0$, temos:

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

para $u \in (0, 1)$. Usando a relação entre a função majorante e F no Teorema 3.1.1 com $x = x_* + u(y - x_*)$ e $\tau = 0$, obtemos

$$\|y - x_*\| \leqslant \int_0^1 [f_0'(u\|y - x_*\|) - f_0'(0)]\|y - x_*\|du = f_0(\|y - x_*\|) - f_0(0) - f_0'(0)\|y - x_*\|.$$

Como $f_0(0) = 0$ e $f_0'(0) = -1$, temos

$$\|y - x_*\| \leqslant f_0(\|y - x_*\|) + \|y - x_*\|,$$

logo $f_0(\|y - x_*\|) \geqslant 0$.

Agora, como f_0 é estritamente convexa e $f_0(t) < 0$ devemos ter $f_0 < 0$ em $(0, t]$, i.e, 0 é o único zero de f_0 em [0,t], então da última desigualdade concluímos que $\|y - x_*\| = 0$, i.e, $y = x_*$. Assim, x_* é o único zero de F em $B(x_*, t)$, onde $t \in (0, k]$.

Portanto, da definição de σ_0, temos x_* é o único de zero de F em $B(x_*, \sigma_0)$. □

Lema 3.3.2. *Se*

$$\left(\frac{f(\rho_0)}{\rho_0 f'(\rho_0)} - 1\right) f_1(\rho_0) = 1$$

e $\rho_0 < \kappa$, *então* $r_0 = \rho_0$ *é o raio de convergência ótimo.*

Demonstração. Suponha que $\left(\frac{f(\rho_0)}{\rho_0 f'(\rho_0)} - 1\right) f_1(\rho_0) = 1$ e $\rho_0 < \kappa$. Defina a função $F : (-\kappa, \kappa) \to \mathbb{R}$ dada por

$$F(t) = \begin{cases} -f(-t), & \text{se} \quad t \in (-\kappa, 0) \\ f(t), & \text{se} \quad t \in [0, \kappa) \end{cases}$$

onde

$$f(t) = \frac{et^2}{2} - t \quad \text{e} \quad f_0(t) = \frac{(e-1)t^2}{2} - t.$$

Usando o software Scientific Workplace conseguimos que

$$\rho_0 = \frac{2}{3e - 2} \simeq 0,3249.$$

Veja que $F(0) = 0$, $F'(0) = -1$ e $|F'(0)^{-1}[F'(t) - F'(\theta t)]| \leqslant f'(|t|) - f'(\theta|t|)$, $\theta \in [0,1]$ e $t \in (-\kappa, \kappa)$. Assim F, $X = Y = \mathbb{R}$, $\Omega = (-\kappa, \kappa)$ e $x^* = 0$ satisfazem as hipóteses do Teorema 3.1.1.

Agora note que $N_F(t)$ é crescente para $t > \frac{2}{e}$. De fato,

$$\begin{aligned}(N_F(t))' &= 1 - \frac{(f'(t)^2) - f(t)f''(t)}{(f'(t)^2)} \\ &= \frac{f(t)f''(t)}{(f'(t)^2)}.\end{aligned}$$

Veja que $f''(t) = e$ e $f(t) > 0$ se, e somente se, $t > \frac{2}{e}$. Logo $(N_F(t))' > 0 \ \ \forall t > \frac{2}{e}$. Como $\rho_0 < \kappa$, é suficiente provar que a aplicação do método de Newton para resolver $F(t) = 0$ com ponto inicial $t_0 = -\rho_0$ não converge. Assim,

$$t_1 = -\rho_0 + \frac{f(\rho_0)}{f'(\rho_0)} = \frac{\rho_0}{f_1(\rho_0)} \simeq 1,2297.$$

Então $t_1 > \frac{2}{e} \simeq 0,7357$. Como $N_F(t)$ é crescente para $t > \frac{2}{e}$ obtemos que $t_k \geqslant \frac{2}{e} \ \ \forall \ k$, e assim o método de Newton para resolver $F(t) = 0$ iniciando com $t_0 = -\rho_0$ não converge. Portanto ρ_0 é o melhor raio de convergência possível. □

3.4 A Sequência de Newton

Finalmente vamos provar as afirmações no Teorema 3.1.1 a respeito da sequência de Newton $\{x^k\}$. Observe primeiro que sequência $\{x^k\}$ satisfaz

$$x^{k+1} = N_F(x^k) \ \ (k \geqslant 0), \tag{3.29}$$

que de fato é uma definição equivalente para esta sequência.

Lema 3.4.1. *A sequência* $\{x^k\}$ *está bem definida, está contida em* $B(x^*, r_0)$, *converge para o ponto* x^* *que é o único zero de* F *em* $B(x^*, \sigma_0)$ *e satisfaz*

$$\lim_{k\to+\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0. \tag{3.30}$$

Se adicionarmos (H_3), *então as sequências* $\{x^k\}$ *e* $\{s_k\}$ *satisfazem*

$$\|x^{k+1} - x^*\| \leqslant \frac{s_{k+1}}{s_k^{p+1}}\|x^k - x^*\|^{p+1}, \quad (k \geqslant 0). \tag{3.31}$$

Demonstração. Como $x^0 \in B(x^*, r_0)$ e $r_0 \leqslant \nu_0$, então dos Lemas 3.2.3 e 3.2.1 respectivamente temos que $N_F(B(x^*, r_0)) \subset B(x^*, r_0)$ e $F'(x^*)$ é invertível, logo $\{x^k\}$ está bem definida e está contida em $B(x^*, r_0)$.

Vamos provar agora que $x^k \to x^*$. De fato, como $0 < \|x^k - x^*\| < r_0 \leqslant \rho_0$, temos do Lema 3.2.3 e Lema 3.1.2 que

$$\|x^{k+1} - x^*\| = \|N_F(x^k) - x^*\| \leqslant |n_{f,f_0}(\|x^k - x^*\|)| < \|x^k - x^*\|. \tag{3.32}$$

Assim, $\{\|x^k - x^*\|\}$ é estritamente decrescente e portanto convergente. Seja $l_* = \lim_{k\to+\infty} \|x^k - x^*\|$. Como $\{\|x^k - x^*\|\} \subset (0, \rho_0)$ e é estritamente decrescente, temos que $0 \leqslant l_* < \rho_0$. Então, da continuidade de n_{f,f_0} em $[0, \rho_0)$ e (3.32) temos $0 \leqslant l_* = |n_{f,f_0}(l_*)|$. Se $l_* \neq 0$ então $0 < l_* < \rho_0$ e pelo Lema 3.1.2 teríamos $|n_{f,f_0}(l_*)| < l_*$, uma contradição, pois $|n_{f,f_0}(l_*)| = l_*$. Portanto $l_* = 0$.

Novamente de (3.32) temos

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \leqslant \frac{|n_{f,f_0}(\|x^k - x^*\|)|}{\|x^k - x^*\|}$$

e, como $\lim_{k\to+\infty} \|x^k - x^*\| = 0$, segue do Lema 3.1.2 que $\lim_{k\to+\infty} \dfrac{|n_{f,f_0}(\|x^k - x^*\|)|}{\|x^k - x^*\|} = 0$ e portanto

$$\lim_{k\to+\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0.$$

Provemos agora a última desigualdade. Primeiro vamos mostrar via indução que as sequências $\{s_k\}$ e $\{x^k\}$ satisfazem $\|x^k - x^*\| \leqslant s_k$.

Para $k = 0$, temos $s_0 = \|x^0 - x^*\|$. Agora assuma que para $k = n$, $\|x^n - x^*\| \leqslant s_n$. Usando (3.32), Lema 3.2.4, hipótese de indução e $|n_{f,f_0}(s_n)| = s_{n+1}$ temos

$$\begin{aligned}
\|x^{n+1} - x^*\| &= \|N_F(x^n) - x^*\| \\
&\leqslant \tfrac{|n_{f,f_0}(s_n)|}{s_n^{p+1}} \|x^n - x^*\|^{p+1} \\
&\leqslant \tfrac{|n_{f,f_0}(s_n)|}{s_n^{p+1}} s_n^{p+1} \\
&= |n_{f,f_0}(s_n)| \\
&= s_{n+1}
\end{aligned}$$

e segue o resultado via indução.

Portanto

$$\|x^{k+1} - x^*\| = \|N_F(x^k) - x^*\| \leqslant \tfrac{|n_f(t_k)|}{t_k^{p+1}} \|x^k - x^*\|^{p+1} = \tfrac{t_{k+1}}{t_k^{p+1}} \|x^k - x^*\|^{p+1}.$$

$\square$

Assim (a), (b), (c) e (d) seguem respectivamente dos Lemas 3.1.3, 3.4.1, 3.3.2 e Observação 3.1.1 e, (e) segue também do Lema 3.4.1. Portanto fica provado o Teorema 3.1.1.

Observação 3.4.1. *Veja que se* $f_0(t) = f(t) \ \ \forall \ \ t \in [0, R)$, *então o Teorema 3.1.1 reduz-se ao Teorema 2.1.1. E neste caso temos que,*

$$s_k = t_k \ \ (k \geqslant 0), \quad \rho_0 = \rho, \quad \sigma_0 = \sigma \quad e \quad r_0 = r.$$

3.5 Casos Especiais sob uma Condição Enfraquecida

3.5.1 Resultado de Convergência sob uma Condição tipo-Hölder Fraca

Teorema 3.5.1. *Sejam* X, Y *espaços de Banach,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função continuamente diferenciável. Seja* $x^* \in \Omega$, $R > 0$ *e* $\kappa := \sup\{t \in [0, R) : B(x^*, t) \subset \Omega\}$. *Suponha que* $F(x^*) = 0$, $F'(x^*)$ *é invertível e que existem* $L_0 > 0$, $L > 0$ *e* $0 < p \leqslant 1$ *tal que*

$$\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leqslant L_0\|x - x^*\|^p, \tag{3.33}$$

$$\|F'(x^*)^{-1}(F'(x) - F'(x^* + \lambda(x - x^*)))\| \leqslant L(1 - \lambda^p)\|x - x^*\|^p, \tag{3.34}$$

para todo $\lambda \in [0, 1]$, $x \in B(x^*, \kappa)$.
Seja

$$r_0 = \min\left\{\kappa, \left(\frac{p + 1}{Lp + L_0(p + 1)}\right)^{1/p}\right\},$$

$$x^0 \in B(x^*, r_0)\backslash\{x^*\} \quad s_0 = \|x^0 - x^*\| \quad s_{k+1} = \frac{Lps_k^{p+1}}{(p + 1)(1 - L_0 s_k^p)}.$$

Então, valem as seguintes afirmações:

(a) $\{s_k\}$ *está bem definida, é estritamente decrescente, está contida em* $(0, r_0)$, *converge para 0 e*

$$\lim_{k\to\infty} \frac{s_{k+1}}{s_k} = 0.$$

(b) A sequência $\{x^k\}$ *dada pelo método de Newton, iniciando com* $x^0 \in B(x^*, r_0)\backslash\{x^*\}$ *está bem definida, está contida em* $B(x^*, r_0)$ *para todo* $k \geqslant 0$ *e converge para* x^*, *que é a única*

solução de $F(x) = 0$ *em* $B\left(x^*, \left(\frac{p+1}{L_0}\right)^{1/p}\right)$, *assim para* $k \geqslant 0$*:*

$$\|x^{k+1} - x^*\| \leqslant \frac{Lp}{(p+1)(1 - L_0 s_k^p)} \|x^k - x^*\|^{p+1}$$

e

$$\|x^k - x^*\| \leqslant \left(\frac{Lp\|x^0 - x^*\|^p}{(p+1)(1 - L_0\|x^0 - x^*\|^p)}\right)^{((p+1)^k - 1)/p} \|x^0 - x^*\|.$$

Além disso, se

$$\Upsilon_0 = \left(\frac{p+1}{Lp + L_0(p+1)}\right)^{1/p} < \kappa,$$

então $r = \Upsilon_0$ *é o melhor raio de convergência possível.*

Demonstração. Considere as funções $f_0, f : [0, \kappa] \to \mathbb{R}$ definidas por

$$f_0(t) = \frac{L_0 t^{p+1}}{p+1} - t \quad e \quad f(t) = \frac{L t^{p+1}}{p+1} - t.$$

Agora observe que

$f_0'(t) = L_0 t^p - 1 \;\; e \;\; f'(t) = L t^p - 1,$

$f_0' \;\; e \;\; f'$ são estritamente crescentes,

$\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leqslant L_0\|x - x^*\|^p,$

$\|F'(x^*)^{-1}(F'(x) - F'(x^* + \lambda(x - x^*)))\| \leqslant L(1 - \lambda^p)\|x - x^*\|^p,$

$f_0(t) \leqslant f(t) \Leftrightarrow L_0 \leqslant L,$

$f_0'(t) \leqslant f'(t) \Leftrightarrow L_0 \leqslant L,$

$f_0'(0) = f'(0) = -1 \;\; e \;\; f_0(0) = f(0) = 0.$

Então F, x^*, f_0, f, H_1, H_2 e H_3 satisfazem as condições do Teorema 3.1.1, e consequentemente temos que

$$r_0 = \min\left\{\kappa, \left(\frac{p+1}{Lp + L_0(p+1)}\right)^{1/p}\right\}$$

e

$$s_{k+1} = \frac{Lp s_k^{p+1}}{(p+1)(1 - L_0 s_k^p)}.$$

Portanto, o resultado segue a partir do Teorema 3.1.1. □

Observação 3.5.1. *Se* $L = L_0$, *então o Teorema 3.5.1 torna-se o Teorema 2.2.1. Além disso, se* $L_0 < L$, *temos que*

$$\Upsilon_0 = \left(\frac{p+1}{Lp + L_0(p+1)}\right)^{1/p} > r = \left(\frac{p+1}{(2p+1)L}\right)^{1/p}$$

e

$$\|x^{k+1} - x^*\| \leqslant \frac{Lp}{(p+1)(1 - L_0 t_k^p)} \|x^k - x^*\|^{p+1} \quad (k \geqslant 0).$$

Veremos agora alguns exemplos ilustrativos.

Exemplo 3.5.1. *Seja* $X = Y = \mathbb{R}$. *Defina* F *em* $\Omega = (-1, -1)$, *dada por*

$$F(x) = e^x - 1.$$

Então, para $x^* = 0$, *temos* $F(x^*) = 0$ *e* $F'(x^*) = 1$. *Portanto, as hipóteses do Teorema 3.5.1 valem para* $p = 1$, $L = e > L_0 = e - 1$. *E veja que*

$$\frac{L}{L_0} = \frac{e}{e-1} = 1,581976707.$$

e tomando $\rho = \left(\frac{p+1}{(2p+1)L}\right)^{1/p}$ *obtemos*

$$\rho = \frac{2}{3L} = 0,245252 < \rho_0 = \frac{2}{2L_0 + L} = 0,324947.$$

Também podemos apresentar uma tabela de comparação, usando o software Maple 13, entre as sequências s_k e t_k e iniciando com $x_0 = 0,7158$.

Tabela de comparação			
k	$\|x_{k+1} - x_k\|$	s_k	t_k
0	0,2473838936	0,2842	0,2842
1	0,03614663422	0,2145495033	0,4826134043
2	0,0006692478074	0,09909547154	1,015025071
3	2,2399998e-7	0,01608560415	0,7960154923
4	0	0,0003616695761	0,7399991923
5	~	1,778927982e-7	0,7357830417
6	~	4,299999235e-14	0,7357588833
7	~	0	0,7357588824
8	~	~	0,7357588825

Exemplo 3.5.2. *Seja* $X = Y = \mathbb{R}$. *Defina a função* F *em* $\Omega = (1, 3)$, *dada por*

$$F(x) = \frac{2}{3}x^{3/2} - x.$$

Então, o zero de F *é* $x^* = \frac{9}{4} = 2,25$. *Usando as hipóteses do Teorema 3.5.1,* $F'(x^*) = 0,5$, $L = 2 > L_0 = 1$ *e* $p = 0,5$, *obtemos:*

$$\rho = \frac{9}{64} = 0,140625 < \rho_0 = \frac{9}{25} = 0,360000.$$

3.5.2 Resultado de Convergência Fraca sob uma Condição de Lipschitz Generalizada

Teorema 3.5.2. *Sejam* X, Y *espaços de Banach,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função continuamente diferenciável. Seja* $x^* \in \Omega, R > 0$ *e* $\kappa := \sup\{t \in [0, R) : B(x^*, t) \subset \Omega\}$. *Suponha que* $F(x^*) = 0, F'(x^*)$ *é invertível e que existem as funções integráveis positivas* $L_0, L : [0, R) \to \mathbb{R}$ *tais que*

$$\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leqslant \int_0^{\|x-x^*\|} L_0(u)du, \tag{3.35}$$

$$\|F'(x^*)^{-1}(F'(x) - F'(x^* + \tau(x - x^*)))\| \leqslant \int_{\tau\|x-x^*\|}^{\|x-x^*\|} L(u)du, \tag{3.36}$$

para todo $\tau \in [0, 1]$, $x \in B(x^*, \kappa)$.

Sejam $\bar{\nu}_0 > 0$, $\bar{\rho}_0 > 0$ *e* $\bar{r}_0 > 0$ *as constantes definidas por*

$$\bar{\nu}_0 := \sup\left\{t \in [0, R) : \int_0^t L_0(u)du - 1 < 0\right\},$$

$$\bar{\rho}_0 := \sup\left\{t \in (0, \bar{\nu}_0) : \frac{\int_0^t L(u)udu}{t(1 - \int_0^t L_0(u)du)} < 1\right\},$$

$$\bar{r}_0 = \min\{\kappa, \bar{\rho}_0\}.$$

Sejam

$$x^0 \in B(x^*, r_0)\backslash\{x^*\}, \quad s_0 = \|x^0 - x^*\|, \quad s_{k+1} = \frac{\int_0^{s_k} L(u)udu}{1 - \int_0^{s_k} L_0(u)du}.$$

Então, as seguintes afirmações são verdadeiras:

(a) A sequência $\{s_k\}$ *está bem definida, é estritamente decrescente, está contida em* $(0, \bar{r}_0)$, *converge para 0 e*

$$\lim_{k\to\infty} \frac{s_{k+1}}{s_k} = 0.$$

(b) A sequência $\{x^k\}$ *gerada pelo método de Newton, com ponto inicial* $x^0 \in B(x^*, \bar{r}_0)\backslash\{x^*\}$, *está bem definida, está contida em* $B(x^*, \bar{r}_0)$ *e converge para* x^* *que é o único zero de* F *em* $B(x^*, \bar{\sigma}_0)$, *onde*

$$\bar{\sigma}_0 := \sup\left\{t \in (0, \kappa) : \int_0^t L_0(u)(t - u)du - t < 0\right\},$$

e vale:

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0.$$

Além disso, se

$$\Upsilon = \frac{\int_0^{\bar{\rho}_0} L(u)u du}{\bar{\rho}_0(1 - \int_0^{\bar{\rho}_0} L_0(u) du)} = 1$$

e $\bar{\rho}_0 < \kappa$, *então* $\bar{r}_0 = \bar{\rho}_0$ *é o melhor raio de convergência possível.*

Demonstração. Considere as funções

$$\bar{f}(t) = \int_0^t L(u)(t-u)du - t$$

e

$$\bar{f_0}(t) = \int_0^t L_0(u)(t-u)du - t.$$

Note que a derivada de $\bar{f}$ e $\bar{f_0}$ são dadas respectivamente por:

$$\bar{f}'(t) = \int_0^t L(u)du - 1$$

e

$$\bar{f_0}'(t) = \int_0^t L_0(u)du - 1.$$

Como L e L_0 são funções integráveis temos que $\bar{f}'$ e $\bar{f_0}'$ são contínuas e, portanto, $\bar{f}$ e $\bar{f_0}$ são continuamente diferenciáveis.

Veja também que

$$\bar{f}(0) = \bar{f_0}(0) = 0 \quad \mathsf{e} \quad \bar{f}'(0) = \bar{f_0}'(0) = -1,$$

$$\bar{f_0}'(\|x - x^*\|) - \bar{f_0}'(0) = \int_0^{\|x-x^*\|} L_0(u)du - 1 + 1 = \int_0^{\|x-x^*\|} L_0(u)du$$

e

$$\begin{aligned} \bar{f}'(\|x - x^*\|) - \bar{f}'(\tau\|x - x^*\|) &= \int_0^{\|x-x^*\|} L(u)du - 1 - \int_0^{\tau\|x-x^*\|} L(u)du + 1 \\ &= \int_{\tau\|x-x^*\|}^{\|x-x^*\|} L(u)du. \end{aligned}$$

Logo, (3.1) torna-se (3.35) com $f' = \bar{f}'$ e $f_0' = \bar{f_0}'$. Além disso, como L e L_0 são positivas, e pelos resultados acima temos que as funções $f = \bar{f}$ e $f_0 = \bar{f_0}$ satisfazem as condições (H_1) e (H_2) no Teorema 3.1.1. Portanto, o resultado segue a partir do Teorema 3.1.1 com $f = \bar{f}$, $f_0 = \bar{f_0}$, $\nu_0 = \bar{\nu}_0$, $\rho_0 = \bar{\rho}_0$, $r = \bar{r}$ e $\sigma = \bar{\sigma}$. □

Capítulo 4

Conclusão e Trabalhos Futuros

Nesta dissertação, apresentamos uma análise de convergência local do método de Newton baseada no princípio majorante de Kantorovich. Esta abordagem melhorou os resultados até então obtidos, no seguinte sentido: com as mesmas informações iniciais são apresentadas melhores estimativas para o raio de convergência. Casos especiais e exemplos numéricos também foram considerados.

No momento, estamos estudando como generalizar a técnica usada em [1] para obter melhores resultados de convergência.

Referências Bibliográficas

[1] Argyros, I. K., Hilout, S. - *Improved local Convegence of Newton's Method Under Weak Majorant Condition*. Journal of Computational and Applied Mathematics, (236) 1892–1902, 2012.

[2] Argyros, I. K. - *Computational Theory of Iterative Methods*. Studies in Computational Mathematics - vol.15, Elsevier, New York, USA, 2007.

[3] Argyros, I. K. - *Concerning the Semilocal Convergence of Newton's Method and Convex Majorants*. Rend. Circ. Mat. Palermo, 2 (57) 331–341, 2008.

[4] Argyros, I. K. - *On the Semilocal Convergence of Inexact Newton Methods in Banach Spaces*. Journal of Computational and Applied Mathematics, (228) 434–443, 2009.

[5] Argyros, I. K. - *A Semilocal Convergence Analysis for Directional Newton Methods*. Mathematics of Computation, vol.80. AMS, 337–343, 2011.

[6] Argyros, I. K. - *Local Convergence of Newton's Method Using Kantorovich's Convex Majorants*. Rev. Anal. Numér. Théor. Approx., (39) 97–106, 2010.

[7] Argyros, I. K., Hilout, S., Tabatabai, M. A. - *Mathematical Modelling with Applications in Biosciences and Engineering*. Nova Science Pub., New York, 2011.

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

[9] Bertsekas, D. - *Network Optimization Continuos and Discrete Models*. Athena Scientific, EUA, 1998.

[10] Ferreira, O. P. - *Local Convergence of Newton's Method in Banach Space from the Viewpoint of the Majorant Principle*. IMA J. Numer.Anal. 3 (29) 746–759, 2009.

[11] Ferreira, O. P., Svaiter, B. F. - *Kantorovich's Majorants Principle for Newton's Method*. Computational Optimization Applied (42) 213–229, 2009.

[12] Ferreira, O. P. - *Local Convergence of Newton's Method under Majorant Condition*. Journal of Computational and applied Mathematics (235) 1515–1522, 2011.

[13] Ferreira, O. P., Gonçalves, M. L. N. - *Local Convergence Analysis of Inexact Newton-Like Methods under Majorant Principle*. Comput. Optim. Appl., (48) 1–21, 2011.

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

[15] Gutiérrez, J. M., Hernández, M. A. - *Newton's Method under Weak Kantorovich Conditions*. IMA J. Numer. Anal, (20) 521–532, 2000.

[16] Gutiérrez, J. M., Hernández, M. A., Salanova, M. A. - *Accessibility of Solutions by Newton's Method*. Int. J. Comput. Math., (57) 239–247, 1995.

[17] Hiriart-Urruty, J. B., Lemaréchal, C. - *Convex Analysis and Minimization Algorithms I*. Springer-Verlag, 1993.

[18] Huang, Z. - *Newton Method under Weak Lipschitz Continuos Derivative in Banach Spaces*. Appl. Math. Comput., (140) 115–126, 2003.

[19] Huang, Z. - *The Convergence Ball of Newton's Method and the Uniqueness Ball of Equations's under Hölder-Type Continuos Derivatives*. Appl. Math. Comput., (47) 247–251, 2007.

[20] Izmailov, A., Solodov, M. - *Otimização- volume 1 Codições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. IMPA, 2009.

[21] Kantorovich, L. V., Akilov, G. P. - *Functional Analysis in Normed Spaces*. Pergamon Press, Oxford, 1982.

[22] Kantorovich, L. V. - *Functional Analysis and Applied Mathematics*. Uspehi Mat. Nauk, (3) 89–185, 1948.

[23] Kantorovich, L, V. - *On Newton's Method for Functional Analysis*. Dokl. Akad. Nauk SSSR, 7 (59) 1237–1240, 1948.

[24] Krantz, S. G., Parks, H. R. - *The Implicit Function Theorem: History, Theory and Applications*. Boston, Birkhäuser, 2002.

[25] Kreyszig, E. - *Introductory Functional Analysis With Application.*New York, John Wiley e Sons, 1978.

[26] Lima, E. L. - *Espaços Métricos*. IMPA, 2009.

[27] Moser, J. - *A New Techniquess for the Construction of Solutions of Nonlinear Differential Equations*. Proc. Nat. Acad. Sci. USA, (47) 1824–1831, 1961.

[28] Nash, J. - *The Embedding Problem for Riemannian Manifolds*. Ann. of Math., (63) 20–63, 1956.

[29] Ortega, J. M. - *The Newton-Kantorovich Theorem*. MAA The Amarican Mathematical Marthly, vol.75 no.6, 658–660, 1998.

[30] Ortega, J. M., Rheinboldt, W. C. - *Iterative Solution of Nonlinear Equations in Several Variables.*New York, SIAM, 1970.

[31] Rockafellar, R. T. - *convex analysis*. Princeton University Press, Princeton, 1970.

[32] Tiel, J. V. - *Convex Analysis an Introductory Text*. Royal Netherlands Meteorological Institute, 1984.

[33] Smale, S. - *Newton Method Estimates From Data at One Point*. Applied and Computational Mathematics, Springer-Verlag, New York, 185–196, 1986.

[34] Wang, X. - Convergence on Newton's Method and Inverse Function Theorem in Banach Space. Mathematical Computational (68) 169–186, 1999.

[35] Wang, X. - *Convergence of Newton's Method and Uniqueness of the Solution of Equation in Banach Space*. IMA j. Numer.Anal. (20) 123–134, 2000.

[36] Wang, X., Li, C. - *Convergence of Newton's Method and Uniqueness of the Solution of Equations in Banach Spaces II*. Acta Math. Sin. (Engl. Ser.) 2 (19) 405–412, 2003.

[37] Zeidler, E. - *Nonlinear Functional Analysis and its Applications I Fixed-point Theorems*. Springer-Verlag, New York, 1985.