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

O método de Newton para resolver equações generalizadas e um teorema tipo-Kantorovich

Juliana Gomes da Silva

Teresina - 2018

Juliana Gomes da Silva

Dissertação de Mestrado:

O método de Newton para resolver equações generalizadas e um teorema tipo-Kantorovich

Dissertação submetida à Coordenação do Programa 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
Coorientadora:
Profª. Dra. Sissy da Silva Souza

Teresina - 2018

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

O Método de Newton para resolver Equações Generalizadas e um Teorema tipo-Kantorovich

JULIANA GOMES DA SILVA

Esta Dissertação foi submetida como parte dos requisitos necessários à obtenção do grau de **Mestre em Matemática**, outorgado pela Universidade Federal do Piauí.

A citação de qualquer trecho deste trabalho é permitida, desde que seja feita em conformidade com as normas da ética científica.

Dissertação aprovada em 23 de Fevereiro de 2018.

Banca Examinadora:

Paulo Sérgio Marques dos Santos
Prof. Dr. Paulo Sérgio Marques dos Santos - Presidente

Sissy da Silva Souza
Profª. Drª. Sissy da Silva Souza - Membro interno

Gilson do Nascimento Silva
Prof. Dr. Gilson do Nascimento Silva - Membro externo

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca Setorial do CCN

S586m Silva, Juliana Gomes da.
O método de Newton para resolver equações generalizadas e um teorema tipo-Kantorovich / Juliana Gomes da Silva. – Teresina, 2017.
58f.

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências da Natureza, Pós-Graduação em Matemática, 2017.
Orientador: Prof. Dr.Paulo Sérgio Marques dos Santos.
Coorientadora: Profª. Drª. Sissy da Silva Souza

1. Matemática. 2. Otimização Matemática. 3. Método de Newton. I. Título.

CDD 519.3

Aos meus pais, com carinho.

Agradecimentos

A Deus, pela saúde concedida. À minha família, em especial aos meus pais João Alves da Silva e Francisca Gomes da Silva, pelo carinho, confiança, apoio e incentivo conferidos na busca deste sonho.

A Jonathan Cardoso, pelo companheirismo de sempre, fazendo-se presente durante toda essa jornada, pela paciência e amor depositados, muito obrigada.

A todos os professores do Programa de Pós-Graduação em Matemática da Universidade Federal do Piauí. Em especial a Gleison, Humberto, Ítalo, João Carlos, José Francisco, Marcos Travaglia, Paulo Alexandre e Wilson pela matemática que tive a oportunidade de aprender com estes grandes professores. Aos professores Xavier e Rondinele que ministraram o curso de verão em análise que participei, experiência significante.

A todos os professores da graduação, em especial aos professores Cleyton Natanael, que me acompanhou em diversas disciplinas e no TCC, pela amizade, por todo o incentivo para seguir além da graduação; Pedro Jorge, pela amizade construída, pelas ótimas conversas, encorajando o crescimento acadêmico; Sissy Souza, pelas palavras de confiança; Paulo Sérgio, pelos conselhos que sempre foram de grande proveito; Carlos Augusto, por acreditar em meu potencial e Roberto Ramos, grande amigo que me acompanhou desde a primeira disciplina na graduação, instigando a busca por conhecimento.

Agradeço novamente aos professores Paulo Sérgio e Sissy Souza pela orientação inigualável, pela excelente sugestão na escolha do tema, pela confiança, pela paciência, pelo incentivo, pelas críticas e sugestões fundamentais para a construção do presente trabalho e pela amizade singular, construída desde a graduação, que buscarei sempre manter.

Agradeço ao professor Gilson do Nascimento Silva por ter aceito o convite para participar da banca que avaliou este trabalho e por suas valiosas contribuições para melhoria deste.

Às pessoas maravilhosas que tive a oportunidade de conhecer no curso, em especial Arilson, Cícero, Edimilson, José Edilson, Josimauro, Leonardo, Luan, Lucas Cassiano, Luciano (*In memoriam*), Quaresma, Rafael, Rafaelber, Raul Kanzan, Ronaldo, Ronnyê, Valéria e Yldenilson com os quais foi construída uma forte amizade. Agradeço também àqueles que já conhecia da graduação Hércules, Jeferson, Kelvin, Leandro, Lucas, Pádua, Ray e, em especial, a Tiago companheiro de estudo, que muito me ajudou. Amizades que buscarei preservar infinitamente.

Aos amigos de Parnaíba, em especial a Alexandre Merison, parceiro/irmão de longa data; Euda Passos, amiga inigualável; Amanda Siqueira, excelente pessoa; Rafael Souza, parceiro de estudo; Paulo Victor, amigo excepcional. À Laysa e Lia pela amizade ímpar. Aos amigos de Teresina, em particular a Arlete e Nayane. À Marina, amiga que tive a oportunidade de conhecer em Teresina, que se tornou essa pessoa tão querida. Enfim, a todos que direta ou indiretamente contribuiram nessa jornada, minha sincera gratidão.

Agradeço a CAPES pelo apoio financeiro.

"O insucesso é apenas uma oportunidade para recomeçar com mais inteligência".

Henry Ford

Resumo

Neste trabalho, consideramos o método de Newton para resolver a equação generalizada $F(x) + T(x) \ni 0$, onde $F : \mathbb{R}^n \to \mathbb{R}^n$ é uma função continuamente diferenciável e $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ é uma aplicação ponto-conjunto, monótona maximal entre espaços de dimensão finita. Mostramos que o método gera uma sequência que converge quadraticamente para uma solução do problema. Para isso, demonstramos um Teorema tipo-Kantorovich que usa a ideia da função majorante para suavizar a continuidade da derivada F'. Isso nos permite obter o raio ótimo de convergência, unicidade da solução e também resolver equações generalizadas sob a condição de Smale.

Palavras Chave: Método de Newton, Equação Generalizada, Operador Monótono Maximal, Teorema tipo-Kantorovich, Função Majorante.

Abstract

In this work, we consider the Newton's method for solving the generalized equation $\mathsf{F}(x) + \mathsf{T}(x) \ni 0$, where $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function and $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued, maximal monotone mapping between finite-dimensional spaces. We show that the method generates a sequence which converges quadratically to a solution of the problem. For this, we show a Kantorovich's theorem that uses the idea of majorant function to smooth the continuity of the application F'. It allows us to obtain the optimal convergence radius, uniqueness of solution and also to solve generalized equations under Smale's condition.
Key Words: Newton Method, Generalized Equation, Maximal Monotone Operator, Kantorovich's Theorem, Majorant Function.

Sumário

Introdução

Neste trabalho, estamos interessados na solução da equação generalizada

$$F(x) + T(x) \ni 0, \tag{1}$$

onde $F : \Omega \to \mathbb{R}^n$ é uma função diferenciável, Ω é um subconjunto aberto de $\mathbb{R}^n$ e $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ um operador ponto conjunto e monótono maximal. A equação generalizada (1) abrange uma ampla variedade de problemas em análise clássica e suas aplicações. Como, por exemplo, sistemas de equações não lineares e sistemas de inequações. Se $\sigma : \mathbb{R}^n \to (-\infty, +\infty)$ é uma função própria semicontínua inferiormente, convexa e

$$T(x) = \partial\sigma(x) = \{u \in \mathbb{R}^n : \sigma(y) \geqslant \sigma(x) + \langle u, y - x\rangle\}, \ \forall\ y \in \mathbb{R}^n,$$

então (1) torna-se o problema de inequação variacional

$$F(x) + \partial\sigma(x) \ni 0,$$

mais detalhes a respeito de tais problemas podem ser encontrados em [4], [9], [13], [27], [32], [33].

O Método de Newton é uma das mais importantes estratégias para resolver (1), o qual gera uma sequência, partindo de um ponto inicial x^0, como segue

$$F(x^k) + F'(x^k)(x^{k+1} - x^k) + T(x^{k+1}) \ni 0, \tag{2}$$

sendo $k = 0,\ 1, \ldots$.

Tal método pode ser pensado como um método tipo Newton baseado na linearização parcial, sendo objeto de pesquisa em diversos artigos que incluem [3], [4], [13], [24] e [30]. Em [4] há uma discussão interessante a respeito de métodos iterativos para resolver equações generalizadas. Percebamos que quando $T \equiv 0$, a iteração (2) torna-se o método de Newton para resolver a equação não linear $F(x) = 0$.

Kantorovich em [14], estabeleceu um resultado de convergência para uma solução do método de Newton para resolver a equação $F(x) = 0$. Impôs suposições sobre a derivada $F'(x^0)$ e o termo $\|F'(x^0)^{-1}F'(x^0)\|$ para obter tal resultado de convergência. A principal ideia de Kantorovich em sua demonstração foi o princípio majorante da sequência $\{x^k\}$ por uma sequência de escalares. Recentemente o interesse pelo Teorema de Kantorovich tem crescido, refletindo em muitos trabalhos publicados, veja por exemplo [6], [9], [11], [26], [30] e [34].

Em nossas pesquisas, encontramos o trabalho de Robinson, ver [27], que até onde sabemos, consta como o primeiro a considerar uma generalização do Teorema de Kantorovich do tipo $F(x) \in K$, onde K é um cone convexo não vazio e fechado, fornecendo resultados e "error bounds" para este método. Inspirado por [28], Robinson utilizou propriedades do processo convexo em seu trabalho.

Josephy em [13], foi o primeiro a considerar uma forma semilocal do método de Newton do tipo (2), com o intuito de resolver (1), para $T = N_C$ a aplicação cone normal de um conjunto convexo $C \subset \mathbb{R}^n$. Para garantir a boa definição do método, na teoria das equações generalizadas, a propriedade de regularidade forte de $F(x^0)+F'(x^0)(x-x^0)+T(x)$ em x^1 para 0, onde x^1, é obtido de x^0, foi usado.

Em [2] e [33], sob uma condição que generaliza a condição de Lipschitz, foi possível estabelecer a convergência local e semilocal, além da taxa quadrática do método de Newton. Uma hipótese frequentemente utilizada para obter a convergência quadrática do método de Newton (2), para resolver a equação (1), é supor a continuidade Lipschitz de F' em uma vizinhança da solução, para maiores detalhes veja [4].

Manter o controle sobre a derivada é um importante ponto na análise de convergência do método de Newton. No entanto, alguns artigos abordam a questão da análise de convergência para resolver a equação $F(x) = 0$, suavizando a hipótese da continuidade Lipschitz de F', que pode ser vista nas referências [11], [34], [35]. Atualmente, todas essas condições são equivalentes a condição de Wang, ver [34]. A vantagem de se trabalhar com uma condição majorante consiste no fato de permitir unificar vários resultados de convergência referentes ao método de Newton.

Embasados em [30], o qual reformulou a condição majorante introduzida em [11], com a finalidade de estudar as propriedades de convergência local do método de Newton (2), apresentamos a relação existente entre a função majorante e a função que define a equação generalizada. Além disso, analisamos raio de convergência ótimo e a unicidade da solução, para o método associado a condição majorante.

O presente trabalho está organizado da seguinte maneira:

No capítulo 1, apresentamos noções e conceitos de Análise, relacionando resultados básicos; Álgebra Linear, onde definimos operadores lineares, monótonos e monótonos maximais; Análise Convexa, com enfoque na caracterização destas em $\mathbb{R}$ e finalizamos com a definição de funções Analíticas, demonstrando um resultado útil nos casos especiais.

No capítulo 2, enunciamos e demonstramos, com base na referência [22], o Teorema de Kantorovich para equações, o qual assume condições semilocais para resolver a equação não linear $\mathsf{F}(x) = 0$.

No capítulo 3, estendemos o resultado anterior para equações generalizadas (1), embasados em [30]. Relacionamos o operador F e a função majorante, mostramos que em uma determinada região o método de Newton para resolver a equação generalizada está bem definido. Apresentamos o raio ótimo e taxas de convergência para a solução, mostrando ao final que tal solução é única, na região determinada.

No capítulo 4, estudamos dois casos especiais, no primeiro estudamos o Teorema 3.2.1 sob a condição de Lipschitz e no segundo apresentamos uma versão do Teorema de Smale para o método de Newton.

Por fim, fazemos algumas observações no capítulo 5.

Capítulo 1

Resultados Auxiliares e Notações

Neste capítulo, apresentamos alguns resultados de Análise e Álgebra Linear. Além disso, discorremos acerca de conceitos básicos de Análise Convexa, abordando alguns resultados de caracterização das funções convexas diferenciáveis de uma variável real. Tais assuntos serviram de fundamentação teórica para o entendimento dos capítulos subsequentes deste trabalho.

1.1 Noções Topológicas

A presente seção foi elaborada tendo por base as referências [18] e [19], na qual definimos alguns conjuntos relevantes do espaço euclidiano $\mathbb{R}^n$, sequência em $\mathbb{R}^n$ e noções de taxa de convergência de uma sequência.

A princípio, sejam dados o ponto $a \in \mathbb{R}^n$ e o número real $r > 0$. A *bola aberta* de centro a e raio r é o conjunto

$$B(a, r) := \{x \in \mathbb{R}^n : \|x - a\| < r\},$$

isto é, o conjunto de pontos $x \in \mathbb{R}^n$ cuja distância ao ponto a é menor do que r.

De modo análogo, definimos a *bola fechada* de centro a e raio r como o conjunto

$$B[a, r] := \{x \in \mathbb{R}^n : \|x - a\| \leqslant r\}.$$

Uma sequência $\{x^k\} \subset \mathbb{R}^n$ é uma aplicação $x : \mathbb{N} \to \mathbb{R}^n$, o valor que tal aplicação assume em cada número natural k é indicado com x^k, intitulado k-ésimo termo da sequência. Dizemos que uma sequência $\{x^k\}$ é limitada quando o conjunto de seus termos é limitado em $\mathbb{R}^n$, ou seja, quando existe um número real $c > 0$ tal que $\|x^k\| \leqslant c$ para todo $k \in \mathbb{N}$.

Definimos uma sequência $\{x^k\}$ monótona não decrescente quando temos $x^k \leqslant x^{k+1}$ para todo $k \in \mathbb{N}$, e $\{x^k\}$ monótona não crescente quando temos $x^{k+1} \leqslant x^k$ para todo $k \in \mathbb{N}$.

Seja $X \subset \mathbb{R}^n$, dizemos que X é aberto quando todos os seus pontos são interiores, isto é, para cada $a \in X$ existe $\varepsilon > 0$ tal que $B(a, \varepsilon) \subset X$, denotamos por $\mathrm{Int}(A)$ o interior de A. Por outro lado, um conjunto $X \subset \mathbb{R}^n$ é dito fechado se todos os seus pontos são de aderência. Lembramos que um ponto a é aderente ao conjunto $X \subset \mathbb{R}^n$ quando este é limite de alguma sequência de pontos $x^k \in X$. Segue que todo ponto $a \in X$ é aderente a X, bastando tomar $x^k = a$ para todo $k \in \mathbb{R}^n$. Chamamos de *fecho* de um conjunto X o conjunto $\overline{X}$, formado por todos os pontos aderentes a X. Portanto, um conjunto X é fechado se, e somente se, todo ponto aderente a X pertence a X, isto é, $X = \overline{X}$.

Seja $X \subset \mathbb{R}^n$. Dizemos que um ponto $a \in \mathbb{R}^n$ é de acumulação do conjunto X quando toda bola aberta de centro a contém algum ponto do conjunto X diferente do próprio a, posto de outro modo, para todo $\varepsilon > 0$, deve existir $x \in X$ tal que $0 < \|x - a\| < \varepsilon$. Denotamos por X' o conjunto dos pontos de acumulação.

Definição 1.1.1. *Dizemos que uma sequência* $\{x^k\}$ *converge para* $x^* \in \mathbb{R}^n$*, se dado* $\varepsilon > 0$ *existe* n_0 *tal que*

$$\|x^k - x^*\| < \varepsilon, \quad \forall\ k \geqslant n_0.$$

Como consequência da Definição 1.1.1 temos que

$$\lim x^k = x^* \Longleftrightarrow \lim \|x^k - x^*\| = 0,$$

que reduz a convergência em $\mathbb{R}^n$ à convergência de números reais não negativos.

Definição 1.1.2. *Uma sequência* $\{x^k\} \in \mathbb{R}^n$ *diz-se uma sequência de Cauchy, se dado* $\varepsilon > 0$ *existe* n_0 *tal que*

$$\|x^m - x^k\| < \varepsilon, \quad \forall\ m,\ k \geqslant n_0.$$

Toda sequência de Cauchy é limitada. De fato, basta tomar $\varepsilon = 1$ na Definição 1.1.2 observamos que existe n_0 tal que, exceto possivelmente os pontos $x^1,\ x^2, \ldots, x^{n_0}$, os demais termos pertencem a $B(x^{n_{0+1}}, 1)$. Assim, o conjunto de termos da sequência é limitado.

Podemos reformular a condição de que a sequência $\{x^k\}$ seja de Cauchy, escrevendo $\lim_{k,m \in \mathbb{N}} \|x^k - x^m\| = 0$. Concluímos que se $\mathbb{N}' \subset \mathbb{N}$ é um subconjunto infinito, isto é, se $\{x^m\}_{m \in \mathbb{N}'}$ é uma subsequência de $\{x^k\}$, temos que $\lim_{k \in \mathbb{N}, m \in \mathbb{N}'} \|x^k - x^m\| = 0$.

Proposição 1.1.1. *Uma sequência em $\mathbb{R}^n$ converge se, e somente se, é de Cauchy.*

Demonstração. Consideremos $\{x^k\}$ uma sequência de Cauchy em $\mathbb{R}^n$. Portanto limitada, logo possui uma subsequência convergente $\{x^m\}_{m\in\mathbb{N}'}$. Se $a = \lim_{m\in\mathbb{N}'} x^m$, temos que $\lim_{m\in\mathbb{N}'} \|x^m - a\| = 0$, como $\{x^k\}$ é de Cauchy segue que $\lim_{k\in\mathbb{N}, m\in\mathbb{N}'} \|x^k - x^m\| = 0$, usando a desigualdade triangular, temos que $\|x^k - a\| \leqslant \|x^k - x^m\| + \|x^m - a\|$, concluímos que $\lim_{k\in\mathbb{N}} \|x^k - a\| = 0$, isto é, $\lim_{k\in\mathbb{N}} x^k = a$. Por outro lado, se $\{x^k\}$ é convergente, com $\lim x^k = a$, como $\|x^k - x^m\| \leqslant \|x^k - x^m\| + \|x^m - a\|$, segue que $\lim_{k,m\to\infty} \|x^k - x^m\| = 0$, isto é $\{x^k\}$ é de Cauchy. □

Proposição 1.1.2. *Seja $\{x^k\}$ uma sequência em $\mathbb{R}^n$. Se existe um número real $\zeta < 1$ tal que*

$$\|x^k - x^*\| \leqslant \zeta \|x^{k-1} - x^*\|, \quad \forall\, k \in \mathbb{N}, \tag{1.1}$$

Então $\{x^k\}$ converge para x^*.

Demonstração. Segue de (1.1) que $\|x^k - x^*\| \leqslant \zeta^k \|x^0 - x^*\|$. Fazendo $k \to \infty$ na desigualdade, temos $\lim_{k\to\infty} \|x^k - x^*\| = 0$. Concluindo a convergência. □

A seguir definimos dois tipos de taxa de convergência, que constam como maneiras de medir a velocidade de convergência de uma sequência. Mostraremos em seguida que a velocidade da taxa de convergência quadrática supera a linear.

Definição 1.1.3. *Seja $\{x^k\}$ uma sequência que converge para x^* em $\mathbb{R}^n$. Dizemos que a convergência é Q-linear se existem uma constante $\mu \in (0,1)$ e $k_0 > 0$ tais que*

$$\|x^{k+1} - x^*\| \leqslant \mu \|x^k - x^*\|, \quad \forall\, k \geqslant k_0.$$

A convergência é Q-quadrática se existem uma constante $M > 0$ e $k_0 > 0$ tais que

$$\|x^{k+1} - x^*\| \leqslant M \|x^k - x^*\|^2, \quad \forall\, k \geqslant k_0.$$

Denotamos por $e_k = \|x^k - x^*\|$ o erro cometido na k-ésima iteração. Façamos uma pequena análise relacionando a velocidade de convergência dessas taxas.

- Convergência Q-linear: Para cada k natural, temos

$$e_k \leqslant \mu e_{k-1} \Rightarrow e_k \leqslant \mu^k e_0.$$

Suponha que na k-ésima iterada obtemos que o erro é da ordem de 10^{-n}, isto é, que temos pelo menos n casas exatas.

$$e_k \leqslant \mu^k e_0 \leqslant 10^{-n}.$$

Segue que

$\log(\mu^k e_0) \leqslant \log(10^{-n}) \Leftrightarrow k\log(\mu) + \log(e_0) \leqslant -n \Leftrightarrow k\log(\mu) \leqslant -n - \log(e_0)$.
Como $\mu \in (0,1)$, temos $\log(\mu) < 0$, dito isso, segue algumas equivalências
$k\log(\mu) \leqslant -n - \log(e_0) \Leftrightarrow k(-\log(\mu)) \geqslant n + \log(e_0) \Leftrightarrow k \geqslant \dfrac{n}{-\log(\mu)}\dfrac{\log(e_0)}{-\log(\mu)}$.
Concluímos que para dobrar o número de casas exatas é necessário dobrar o número de iterações.

- Convergência Q-quadrática: Para cada k natural, temos

$$e_k \leqslant Me_{k-1}^2 \Rightarrow e_k \leqslant M^{2^k-1}e_0^{2^k}.$$

Para mostrar que vale a implicação, basta fazer por indução, para $k = 0$ é imediato. Suponha que vale para k, e provemos que vale para $k+1$.

$$e_{k+1} \leqslant M(e_k^2) \leqslant M(M^{2^k-1}e_0^{2^k})^2 = M^{2^{k+1}-1}e_0^{2^{k+1}},$$

portanto, vale para $k+1$. Vamos supor que na k-ésima iteração tenhamos

$$e_k \leqslant M^{2^k-1}e_0^{2^k} \leqslant 10^{-n}.$$

Segue que

$\log(M^{2^k-1}e_0^{2^k}) \leqslant \log(10^{-n}) \Leftrightarrow (2^k - 1)\log(M) + 2^k\log(e_0) \leqslant -n \Leftrightarrow$
$2^k(\log(M) + \log(e_0)) \leqslant -n + \log(M) \Leftrightarrow 2^k(\log(Me_0)) \leqslant -n + \log(M) \Leftrightarrow$
$2^k \geqslant \dfrac{n}{-\log(Me_0)} + \dfrac{\log(M)}{-\log(Me_0)}$.
Assumindo, sem perda de generalidade, $Me_0 \in (0,1)$. Observamos que para dobrar o número de casas exatas é necessário apenas mais uma iteração.

Finalizamos, assim, a discussão acerca de resultados básicos de Topologia no espaço euclidiano $\mathbb{R}^n$.

1.2 Transformações Lineares e Matrizes

Seção estruturada segundo as referências [7], [15], [19], [20] e [23]. Onde apresentamos conceitos fundamentais de transformações lineares e matrizes, que são necessárias para demonstrações de resultados presentes neste trabalho.

Iniciamos por denotar $\mathbb{R}^{m\times n}$ o conjunto das matrizes $m \times n$, e dado $A \in \mathbb{R}^{m\times n}$, indicamos a transposta da matriz A por A^T. Ao espaço dos operadores lineares de $\mathbb{R}^n$ em $\mathbb{R}^m$ designamos por $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Sendo que se $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, então $A \in \mathbb{R}^{m\times n}$. Quando $m = n$, dizemos apenas que $A \in \mathcal{L}(\mathbb{R}^n)$.

Definimos o *posto segundo linhas* da matriz $A \in \mathbb{R}^{m\times n}$ como o números máximo de linhas linearmente independentes em A, em que este número é igual à dimensão do espaço $\mathbb{R}^n$. De modo similar definimos o *posto segundo colunas* de uma matriz $A \in \mathbb{R}^{m\times n}$ sendo o número máximo de colunas linearmente independentes em A, tal número é igual à dimensão do espaço $\mathbb{R}^m$.

Mesmo que os vetores coluna e vetores linha de A sejam subespaços de espaços vetoriais distintos, temos que vale o seguinte resultado.

Proposição 1.2.1. *Para toda matriz $A \in \mathbb{R}^{m\times n}$, o posto segundo linhas e o posto segundo colunas são iguais.*

Demonstração. Ver o Teorema 8.2, da referência [15]. □

Com os dados anteriores podemos definir formalmente o posto de uma matriz.

Definição 1.2.1. *O posto de $A \in \mathbb{R}^{m\times n}$, é o número de linhas ou de colunas linearmente independentes da matriz A. Denotamos $\rho(A)$.*

A seguir definimos matrizes inversíveis.

Definição 1.2.2. *Uma matriz $A \in \mathbb{R}^{n\times n}$ é dita inversível ou não singular, se existe uma matriz $B \in \mathbb{R}^{n\times n}$ tal que $AB = BA = I_n$, em que I_n é a matriz identidade de ordem n. Chamamos de inversa de A a matriz B. Notacionamos A^{-1} a inversa de A.*

Quando a matriz A não possue inversa, dizemos que A é singular ou não inversível.

Proposição 1.2.2. *Uma matriz $A \in \mathbb{R}^{n\times n}$ admite inversa se, e somente se, $\rho(A) = n$.*

Demonstração. Ver o Corolário 3.8.2, da referência [20] □

Agora, com base nas referências [7], [19] e [23], definimos norma de matrizes, abordamos algumas de suas principais propriedades e, por fim, demonstramos o famoso Lema de Banach.

Consideremos $\mathsf{T},\ \mathsf{S} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Uma norma $||\cdot||$ é uma aplicação que associa a cada matriz um número real não negativo satisfazendo as seguintes propriedades.

N1. $\mathsf{T} \neq 0 \Rightarrow ||\mathsf{T}|| > 0$;

N2. $||\alpha \mathsf{T}|| = |\alpha|\ ||\mathsf{T}||, \quad \forall\ \alpha \in \mathbb{R}^n$;

N3. $||\mathsf{T} + \mathsf{S}|| \leqslant ||\mathsf{T}|| + ||\mathsf{S}||$.

Segue um exemplo de norma matricial induzida pela norma vetorial.

Exemplo 1.2.1. *Seja* $\mathsf{T} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. *Consideremos a norma das transformações lineares* $||\cdot||$ *como sendo*

$$||\mathsf{T}|| = \sup_{||x|| \neq 0} \frac{||\mathsf{T}x||}{||x||}. \tag{1.2}$$

A verificação que (1.2) satisfaz as propriedades de norma de matriz são facilmente constatadas.

Temos ainda que a aplicação norma matricial induzida pela norma vetorial goza das seguintes propriedades.

Lema 1.2.1. *Dados* $\mathsf{T},\ \mathsf{S} \in \mathcal{L}(\mathbb{R}^n)$ *e* $x \in \mathbb{R}^n$, *então são válidas as seguintes propriedades:*

1. $||\mathsf{T}x|| \leqslant ||\mathsf{T}||\ ||x||$;

2. $||\mathsf{TS}|| \leqslant ||\mathsf{T}||\ ||\mathsf{S}||$;

3. $||\mathsf{T}^k|| \leqslant ||\mathsf{T}||^k, \quad \forall\ k = 0,\ 1, \ldots$.

Demonstração. Temos que se x for o vetor nulo o resultado segue imediato. Consideremos $x \neq 0$ e $y = x/||x||$, usando (1.2) temos,

$$||\mathsf{T}|| \leqslant ||\mathsf{T}y|| = \frac{1}{||x||}||\mathsf{T}x||.$$

Logo $||\mathsf{T}x|| \leqslant ||\mathsf{T}||\ ||x||$, provando o primeiro item.

Combinando (1.2), item 1 e propriedades de supremo, obtemos

$$||\mathsf{TS}|| = \sup_{||x|| \neq 0} \frac{||\mathsf{TS}x||}{||x||} \leqslant \sup_{||x|| \neq 0} \frac{||\mathsf{T}||\ ||\mathsf{S}x||}{||x||} = ||\mathsf{T}||\ ||\mathsf{S}||,$$

o que demonstra o item 2. Por fim, temos que o item 3 segue como consequência direta do item anterior. $\square$

O próximo resultado é de suma importância para demonstrar a boa definição do método de Newton.

Lema 1.2.2. *(Lema de Banach) Seja* $\mathsf{B} \in \mathcal{L}(\mathbb{R}^n)$ *um operador linear e* I *o operador identidade de* $\mathbb{R}^n$*. Se* $||\mathsf{B} - \mathrm{I}|| < 1$ *então* B *é inversível e*

$$||\mathsf{B}^{-1}|| \leqslant \frac{1}{1 - ||\mathsf{B} - \mathrm{I}||}. \tag{1.3}$$

Demonstração. Vamos mostrar que se $\mathsf{T} \in \mathcal{L}(\mathbb{R}^n)$ é tal que $||\mathsf{T}|| < 1$ então $(\mathrm{I} - \mathsf{T})$ é inversível e vale a estimativa

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

Consideremos as sequências $\{\mathsf{S}_k\}$ e $\{t_k\}$ definidas por

$$\mathsf{S}_k = \mathrm{I} + \mathsf{T} + \cdots + \mathsf{T}^k, \quad t_k = 1 + ||\mathsf{T}|| + \cdots + ||\mathsf{T}||^k.$$

Donde temos que $||\mathsf{S}_{k+1} - \mathsf{S}_k|| = ||\mathsf{T}^{k+1}|| \leqslant ||\mathsf{T}||^{k+1} = t_{k+1} - t_k$. Como $||\mathsf{T}|| < 1$, segue que $\{t_k\}$ é uma sequência monótona e convergente, com limite $t_* = \dfrac{1}{1 - ||\mathsf{T}||}$. Então, $\{\mathsf{S}_k\}$ é uma sequência de Cauchy em $\mathcal{L}(\mathbb{R}^n)$, portanto, existe o limite $\lim_{k\to\infty} \mathsf{S}_k$.

Por outro lado, temos

$$\mathsf{S}_k(\mathrm{I} - \mathsf{T}) = (\mathrm{I} + \mathsf{T} + \cdots + \mathsf{T}^k)(\mathrm{I} - \mathsf{T}) = \mathrm{I} - \mathsf{T}^{k+1}.$$

Ainda, $\lim_{k\to\infty}(\mathrm{I} - \mathsf{T}^k) = \mathrm{I}$, pois $||\mathrm{I} - (\mathrm{I} + \mathsf{T}^k)|| \leqslant ||\mathsf{T}||^k \to 0$.

Logo, obtemos

$$\mathsf{S}_k(\mathrm{I} - \mathsf{T}) = \mathrm{I} - \mathsf{T}^{k+1} \to \mathrm{I}.$$

Concluímos que o limite de $\{\mathsf{S}_k\}$ existe, $\lim_{k\to\infty} \mathsf{S}_k = (\mathrm{I} - \mathsf{T})^{-1}$, portanto, $(\mathrm{I} - \mathsf{T})$ é inversível.

Vejamos agora a estimativa

$$\begin{aligned}
||(\mathrm{I} - \mathsf{T})^{-1}|| &= \left\|\lim_{k\to\infty} \mathsf{S}_k\right\| \\
&\leqslant \lim_{k\to\infty} ||\mathrm{I} + \mathsf{T} + \cdots + \mathsf{T}^k|| \\
&\leqslant \lim_{k\to\infty} (||\mathrm{I}|| + ||\mathsf{T}|| + \cdots + ||\mathsf{T}||^k) \\
&= \lim_{k\to\infty} t_k \\
&= \frac{1}{1 - ||\mathsf{T}||}.
\end{aligned}$$

Assim, fazendo $\mathsf{T} = \mathrm{I} - \mathsf{B}$ e tendo em vista que $||\mathsf{B} - \mathrm{I}|| < 1$, temos que $(\mathrm{I} - \mathsf{T}) = \mathsf{B}$ é inversível e vale a estimativa (1.3). $\square$

Como consequência do Lema de Banach temos

Corolário 1.2.1. *Sejam* $A,\ C \in \mathcal{L}(\mathbb{R}^n)$ *e assuma que* A *é inversível com* $\|A^{-1}\| \leqslant \beta$*. Se* $\|A - C\| \leqslant \gamma$ *e* $\beta\gamma < 1$*, então* C *é inversível e*

$$\|C^{-1}\| \leqslant \frac{\beta}{1-\beta\gamma}$$

Demonstração. Notemos que

$$\|I - A^{-1}C\| = \|A^{-1}(A - C)\| \leqslant \beta\gamma < 1.$$

Logo pelo Lema de Banach $A^{-1}C$ é inversível. Como $A^{-1}C = I - (I - A^{-1}C)$ temos que C é inversível. Além disso,

$$\begin{aligned}
\|C^{-1}\| &= \|(I - (I - A^{-1}C))^{-1}A^{-1}\| \\
&\leqslant \beta \sum_{i=0}^{\infty} \|I - A^{-1}C\|^i \\
&\leqslant \beta \sum_{i=0}^{\infty} (\beta\gamma)^i \\
&= \frac{\beta}{1-\beta\gamma}.
\end{aligned}$$

□

O que encerra a argumentação acerca de transformações lineares e matrizes.

1.3 Funções Contínuas

Apresentamos nesta seção resultados interessantes a respeito de funções contínuas, as demonstrações foram omitidas, por tratar de teoremas e proposições clássicas que podem ser encontradas, por exemplo, em [16], [18] e [19]. Começamos por definir formalmente a continuidade de uma função.

Dizemos que uma aplicação $\varphi : X \subset \mathbb{R}^m \to \mathbb{R}^n$ é contínua no ponto $a \in X$, se para qualquer $\varepsilon > 0$ dado, podemos obter $\delta > 0$ tal que $\|x - a\| < \delta$ implica $\|\varphi(x) - \varphi(a)\| < \varepsilon$. Quando φ é contínua em todos os pontos do conjunto X, dizemos apenas que φ é uma função contínua.

Definição 1.3.1. *Uma função* $\varphi : X \subset \mathbb{R}^m \to \mathbb{R}^n$ *diz-se Lipschitziana quando existem* $\delta,\ L > 0$ *tais que,* $\|y - x\| \leqslant \delta$ *para quaisquer* $x,\ y \in X$*, temos*

$$\|\varphi(y) - \varphi(x)\| \leqslant L\|y - x\|.$$

Observação 1.3.1. *Toda função Lipschitziana é contínua. Para verificar tal fato basta tomar* $\delta = \varepsilon/L$.

Definição 1.3.2. *Seja* $\varphi : X \to \mathbb{R}^n$, *definida no conjunto aberto* $X \subset \mathbb{R}^m$. *Dizemos que* φ *é diferenciável no ponto* $x \in X$ *quando existe uma transformação linear* $\varphi' : \mathbb{R}^m \to \mathbb{R}^n$ *tal que*

$$\varphi(x+h) = \varphi(x) + \varphi'(x)h + s(h), \quad \textit{onde} \quad \lim_{h\to 0} \frac{s(h)}{||h||} = 0.$$

Se $\varphi : X \to \mathbb{R}^n$ *é diferenciável em todos os pontos de* X, *dizemos simplesmente que* φ *é uma função diferenciável.*

A aplicação $\varphi'(x)$ diz-se a derivada de φ no ponto x.

Definição 1.3.3. *Seja* $\varphi : X \to \mathbb{R}^n$, *uma aplicação definida no aberto* $X \subset \mathbb{R}^m$. *Dizemos que* φ *é continuamente diferenciável em* X, *ou que* φ *é de classe* C^1, *e escrevemos* $\varphi \in C^1$, *quando* φ *for diferenciável e, além disso,* $\varphi' : X \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ *for contínua.*

Proposição 1.3.1. *(Teorema Fundamental do Cálculo) Seja* $X \subset \mathbb{R}^m$ *um conjunto aberto. Dada* $\varphi : X \to \mathbb{R}^n$ *de classe* C^1, *vamos supor que o segmento de reta* $[x, x+h]$ *esteja em* X. *Então*

$$\varphi(x+h) - \varphi(x) = \int_0^1 \varphi'(x+th)h \; dt$$

Proposição 1.3.2. *(Teorema de Permanência do Sinal) Sejam* $X \subset \mathbb{R}$, $\varphi : X \to \mathbb{R}$ *e* $a \in X'$. *Se* $\lim_{x\to a} \varphi(x) = L$ *e* $\alpha < L < \beta$ *com* α, $\beta \in \mathbb{R}$, *então existe* $\delta > 0$ *tal que para todo* $x \in X$ *e* $0 < |x - a| < \delta$, *temos* $\alpha < \varphi(x) < \beta$.

Proposição 1.3.3. *(Teorema do Valor Médio) Seja* $\varphi : [a, b] \to \mathbb{R}$ *uma aplicação contínua e diferenciável no intervalo aberto* (a, b). *Então existe* $c \in (a, b)$ *para o qual*

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

1.4 Elementos de Análise Convexa

Nesta seção, construída a partir das referências [12], [16], [23], [28] e [31], definimos formalmente conjunto convexo e funções convexas. Abordamos fatos básicos a respeito de tais conjuntos e funções, importantes em todo o desenvolvimento do presente trabalho.

Definição 1.4.1. *Um conjunto* $D \subset \mathbb{R}^n$ *é dito convexo se para quaisquer* $x,\ y \in D$ *e* $\lambda \in [0,1]$, *tem-se*

$$\lambda x + (1-\lambda)y \in D.$$

Isto é, um conjunto D é dito convexo se contém todos os segmentos com extremos pertencentes ao conjunto D. O ponto $\lambda x + (1-\lambda)y$, chama-se *combinação convexa* de x e y (com parâmetro λ).

Exemplo 1.4.1. *O conjunto vazio e o espaço euclidiano* $\mathbb{R}^n$, *são exemplos triviais de conjuntos convexos.*

Exemplo 1.4.2. *Uma bola em* $\mathbb{R}^n$ *é um conjunto convexo.*

Demonstração. Consideremos o caso de uma bola aberta em $\mathbb{R}^n$, sejam $x^1,\ x^2 \in B(a,r)$ e $\lambda \in [0,1]$, temos que

$$\begin{aligned} \|\lambda x^1 + (1-\lambda)x^2 - a\| &= \|\lambda(x^1 - a) + (1-\lambda)(x^2 - a)\| \\ &\leqslant \lambda\|x^1 - a\| + (1-\lambda)\|x^2 - a\| \\ &< \lambda r + (1-\lambda)r = r. \end{aligned} \tag{1.4}$$

Por meio de um raciocínio análogo ao anterior provamos que a bola fechada em $\mathbb{R}^n$ também é um conjunto convexo. $\square$

Exemplo 1.4.3. *Um semi-espaço qualquer em* $\mathbb{R}^n$ *é um conjunto convexo, isto é,*

$$C = \{x \in \mathbb{R}^n : \langle a, x\rangle \leqslant c\},$$

sendo $a \in \mathbb{R}^n$ *e* $c \in \mathbb{R}$, *é convexo.*

Demonstração. Consideremos $x,\ y \in C$, então $\langle a, x\rangle \leqslant c$ e $\langle a, y\rangle \leqslant c$. Logo, para $tx + (1-t)y$, com $t \in [0,1]$, temos:

$$\begin{aligned} \langle a, tx + (1-t)y\rangle &= t\langle a, x\rangle + (1-t)\langle a, y\rangle \\ &\leqslant tc + (1-t)c \\ &= c \end{aligned} .$$

Portanto, $tx + (1-t)y \in C$ e, assim, C é um conjunto convexo. $\square$

Com a ideia de conjunto convexo em mente, estamos aptos a estabelecer a definição de função convexa.

Definição 1.4.2. *Se* $D \subset \mathbb{R}^n$ *é um conjunto convexo, dizemos que a função* $\psi : D \to \mathbb{R}$ *é convexa em* D *quando*

$$\psi(\lambda x + (1-\lambda)y) \leqslant \lambda\psi(x) + (1-\lambda)\psi(y).$$

para quaisquer $x,\ y \in D$ *e* $\lambda \in [0,1]$.

Definição 1.4.3. *Se* $D \subset \mathbb{R}^n$ *é um conjunto convexo, dizemos que a função* ψ *é estritamente convexa quando a desigualdade anterior é estrita para todos* $x,\ y$ *distintos em* D *e* $\lambda \in (0,1)$.

Definição 1.4.4. *Seja* $\psi : \mathbb{R}^n \to \mathbb{R}$ *função convexa. Dizemos que* $y \in \mathbb{R}^n$ *é um subgradiente de* ψ *no ponto* $x \in \mathbb{R}^n$ *se*

$$\psi(z) \geqslant \psi(x) + \langle y, z - x\rangle, \forall\ z \in \mathbb{R}^n.$$

Ao conjunto de todos os subgradientes de ψ em x, chamamos de subdiferencial de ψ em x, o denotamos por $\partial\psi(x)$.

A seguir vamos discorrer sobre algumas das propriedades de funções convexas de uma variável.

Proposição 1.4.1. *(Desigualdades Fundamentais) Consideremos* $I \subset \mathbb{R}$ *um intervalo e* $\psi : I \to \mathbb{R}$ *uma função convexa. Então dados* $a < b < c \in I$*, temos que*

$$\frac{\psi(b)-\psi(a)}{b-a} \leqslant \frac{\psi(c)-\psi(a)}{c-a} \leqslant \frac{\psi(c)-\psi(b)}{c-b}. \tag{1.5}$$

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

$$s(x) = \frac{\psi(x)-\psi(d)}{x-d},\ \ d \in \mathit{Int}(I)$$

é não decrescente. Se ψ *for estritamente convexa as desigualdades anteriores são estritas e a função* s *é crescente.*

Demonstração. Temos que

$$b = b\frac{c-a}{c-a} = \frac{ac-ab+bc-ac}{c-a} = \frac{c-b}{c-a}a + \frac{b-a}{c-a}c;$$

Como $a < b < c$ então $(c-b)/(c-a) < 1$ e $(b-a)/(c-a) < 1$. Combinando este fato com a igualdade anterior e a convexidade de ψ, obtemos

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

Por meio de algumas manipulações constatamos que

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

equivalentemente

$$\frac{\psi(b) - \psi(a)}{b-a} \leqslant \frac{\psi(c) - \psi(a)}{c-a}.$$

Concluímos, assim, a primeira desigualdade de (1.5). Obtemos a segunda desigualdade de modo análogo. Se considerarmos $\alpha,\ \beta \in I$ com $\alpha < \beta$, segue das desigualdades em (1.5) que $s(\alpha) \leqslant s(\beta)$, portanto, s é não decrescente. As outras afirmações seguem da convexidade estrita de ψ. □

As desigualdades em (1.5) significam que, para $a < b < c$ a secante ab tem inclinação menor que a secante ac, esta por sua vez, tem inclinação menor que do que a secante bc.

Corolário 1.4.1. *Sejam* $I \subset \mathbb{R}$ *um intervalo e* $\psi : I \to \mathbb{R}$ *uma função convexa. Se* $u,\ v,\ w \in I$ *sendo* $u < w$ *e* $u \leqslant v \leqslant w$*, então*

$$\psi(v) - \psi(u) \leqslant [\psi(w) - \psi(u)]\,\frac{v-u}{w-u}.$$

Demonstração. Resultado segue da Proposição 1.4.1. □

Corolário 1.4.2. *Suponha* $\psi : I \to \mathbb{R}$ *convexa no intervalo* $I \subset \mathbb{R}$*, então existem as derivadas laterais* $\psi_+(d)$ *e* $\psi_-(d)$*, para todo* $d \in$ *Int*(I)*.*

Demonstração. Como $d \in \mathrm{Int}(I)$ existe $a \in I$ tal que $a < d$, segue da Proposição 1.4.1 que

$$s(a) = \frac{\psi(a) - \psi(d)}{a-d} \leqslant \frac{\psi(x) - \psi(d)}{x-d} = s(x),\ \forall\ x > d,$$

ou seja, s é limitada inferiormente. Pela monotonicidade da função s existe o limite

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

De modo análogo provamos que existe $\psi_-(d)$. □

Proposição 1.4.2. *As seguintes afirmações sobre a função* $\psi : I \to \mathbb{R}$*, diferenciável em* $I \subset \mathbb{R}$*, são equivalentes:*

(i) ψ *é convexa.*

(ii) A derivada $\psi' : I \to \mathbb{R}$ *é monótona não decrescente.*

(iii) Para quaisquer $a,\ x \in I$, *temos que* $\psi(x) \geqslant \psi(a) + \psi'(a)(x-a)$.

Demonstração. Mostraremos que *(i)* $\Rightarrow$ *(ii)* $\Rightarrow$ *(iii)* $\Rightarrow$ *(i)*.

(i) $\Rightarrow$ *(ii)* Sejam $a < x < b$ em I. Nas desigualdades fundamentais (1.5), fazendo primeiro $x \to a^+$, e depois $x \to b^-$ obtemos

$$\psi_+(a) \leqslant \frac{\psi(b) - \psi(a)}{b-a} \leqslant \psi_-(b).$$

Como ψ é diferenciável temos $a < b \Rightarrow \psi'(a) \leqslant \psi'(b)$. Portanto, ψ é monótona não decrescente.

(ii) $\Rightarrow$ *(iii)* Supondo $a < x$ em I, pelo Teorema do Valor Médio, existe $z \in (a, x)$ tal que

$$\psi(x) = \psi(a) + \psi'(z)(x-a).$$

Como ψ' é monótona não decrescente, temos que $\psi'(z) \geqslant \psi'(a)$. Então,

$$\psi(x) \geqslant \psi(a) + \psi'(a)(x-a).$$

Para $x < a$, o raciocínio é análogo.

(iii) $\Rightarrow$ *(i)* Consideremos $a < c < b$ em I. Sejam $\alpha(x) = \psi(c) + \psi'(c)(x-c)$ e o semiplano superior $H = \{(x,y) \in \mathbb{R}^2 : y \geqslant \alpha(x)\}$ determinado pela reta $y = \alpha(x)$, tangente ao gráfico de ψ no ponto $(c, \psi(c))$. Como H é um conjunto convexo, e os pontos $(a, \psi(a))$ e $(b, \psi(b))$ pertencem a H, temos que o segmento de reta com extremos nesses pontos está contido no semiplano H. Então o ponto de tal segmento que tem c como abcissa pertence a H, ou seja, tem ordenada $\geqslant \alpha(c)$, que por sua vez é igual a $\psi(c)$. Logo

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

Como $a,\ b,\ c \in I$ foram tomados arbitrários, concluímos que ψ é convexa. □

Notemos que o item *(iii)* da proposição anterior significa que o gráfico de ψ está situado acima de qualquer uma de suas tangentes. Como consequência da Proposição 1.4.2, temos os seguintes Corolários.

Corolário 1.4.3. *Todo ponto crítico de uma função convexa é um ponto de mínimo absoluto.*

Demonstração. Suponhamos $a \in I$ ponto crítico da função $\psi : I \to \mathbb{R}$, então $\psi'(a) = 0$. Como ψ é convexa a condição *(iii)* do Teorema 1.4.2 assegura $\psi(x) \geqslant \psi(a)$ para todo $x \in I$, portanto, a é ponto de mínimo absoluto. □

Corolário 1.4.4. *Uma função* $\psi : I \to \mathbb{R}$*, duas vezes diferenciável no intervalo* I*, é convexa se, e somente se,* $\psi''(x) \geqslant 0$ *para todo* $x \in I$*.*

Demonstração. De fato, afirmar que $\psi''(x) \geqslant 0$ para todo $x \in I$ equivale a afirmar que $\psi' : I \to \mathbb{R}$ é monótona não decrescente. □

Finalizamos, aqui, a discussão acerca de funções convexas de uma variável.

1.5 Elementos da Teoria dos Operadores

Nesta seção, baseada nas referências [1], [15], [21], [29], e [30], abordamos um pouco da teoria dos operadores lineares em espaços de dimensão finita, apresentamos definições de suma importância no desenvolvimento deste trabalho. Começamos com uma pequena discussão a respeito de operadores adjuntos.

Dado um operador linear $G : \mathbb{R}^n \to \mathbb{R}^m$, podemos associar a este um operador $G^* : \mathbb{R}^m \to \mathbb{R}^n$, o qual denotamos adjunto de G, que satisfaz a seguinte igualdade $\langle Gx, y\rangle = \langle x, G^*y\rangle$, para vetores quaisquer $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$.

Definição 1.5.1. *O operador linear* $G : \mathbb{R}^n \to \mathbb{R}^n$ *é dito auto-adjunto quando* $G = G^*$*, ou seja, se vale a igualdade*

$$\langle Gx, y\rangle = \langle x, Gy\rangle.$$

Definição 1.5.2. *Um operador linear* $G : \mathbb{R}^n \to \mathbb{R}^n$ *é dito positivo se* G *for auto-adjunto e* $\langle Gx, x\rangle \geqslant 0$*, para todo* $x \in \mathbb{R}^n$*.*

Lema 1.5.1. *Seja* G *um operador positivo. Suponhamos que* G^{-1} *existe, então para cada* $x \in \mathbb{R}^n$ *temos que*

$$\langle Gx, x\rangle \geqslant \frac{||x||^2}{||G^{-1}||}.$$

Demonstração. Seja λ autovalor de G $(\lambda > 0)$. Vamos mostrar que

$$\langle Gx, x\rangle \geqslant \lambda_{min(G)}\langle x, x\rangle \geqslant \frac{||x||^2}{||G^{-1}||}.$$

A primeira desigualdade segue imediata. Sabemos que $||Gx|| \leqslant ||G||\ ||x||$, então

$$||G|| \geqslant \left\|\frac{Gx}{||x||}\right\| = \lambda\left\|\frac{x}{||x||}\right\| = \lambda.$$

Logo, $\lambda_{max(G^{-1})} \leqslant ||G^{-1}||$, então $\dfrac{1}{\lambda_{min(G)}} \leqslant ||G^{-1}||$. Concluímos que $\dfrac{1}{||G^{-1}||} \leqslant \lambda_{min(G)}$, segue o resultado. □

Seja $\mathsf{G} : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear monótono, convencionamos que

$$\widehat{\mathsf{G}} := \frac{1}{2}(\mathsf{G} + \mathsf{G}^*),$$

onde G^* é o operador adjunto de G.

Denotamos na Seção 1.2 por $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ o espaço consistindo de todas as aplicações lineares e contínuas $\mathsf{A} : \mathbb{R}^n \to \mathbb{R}^m$ e partir desse instante fixemos o operador norma de A definido por

$$||\mathsf{A}|| := \sup\{||\mathsf{A}x|| : \ ||x|| \leqslant 1\}.$$

Para finalizar esta seção, introduzimos o conceito e discutimos algumas propriedades de operadores monótonos ponto conjunto e operadores monótonos maximais em $\mathbb{R}^n$.

Definição 1.5.3. *Dizemos que* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *é um operador ponto conjunto se a cada elemento* $x \in \mathbb{R}^n$ *associamos um subconjunto* $\mathsf{T}(x) \subseteq \mathcal{P}(\mathbb{R}^n)$.

Segue a listagem de algumas definições referentes a operadores ponto conjunto.

Definição 1.5.4. *Seja* $\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador ponto conjunto.*

i) O gráfico de T, $\mathsf{G}(\mathsf{T})$, *é definido por*

$$\mathsf{G}(\mathsf{T}) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : y \in \mathsf{T}(x)\}.$$

ii) O domínio de T, $\mathsf{Dom}(\mathsf{T})$, *é definido por*

$$\mathsf{Dom}(\mathsf{T}) := \{x \in \mathbb{R}^n : \mathsf{T}(x) \neq \emptyset\}.$$

iii) A imagem de T, $\mathsf{Im}(\mathsf{T})$, *é definida por*

$$\mathsf{Im}(\mathsf{T}) := \{u \in \mathbb{R}^n : u \in \mathsf{T}(x) \textit{ para algum } x \in \mathbb{R}^n\}.$$

Definição 1.5.5. *Seja* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador ponto conjunto. Dizemos que* T *é monótono se para quaisquer* $x,\ y \in \mathsf{Dom}(\mathsf{T})$ *e* $u \in \mathsf{T}(x)$, $v \in \mathsf{T}(y)$ *vale a desigualdade*

$$\langle v - u, y - x\rangle \geqslant 0.$$

Se o operador T for ponto a ponto, dizemos que será monótono quando para quaisquer $x,\ y \in \mathsf{Dom}(\mathsf{T})$ tivermos

$$\langle \mathsf{T}(y) - \mathsf{T}(x), y - x\rangle \geqslant 0.$$

Um subconjunto de $\mathbb{R}^n \times \mathbb{R}^n$ é monótono se for o gráfico de um operador monótono.

Exemplo 1.5.1. *Sendo* $\psi : \mathbb{R}^n \to \mathbb{R}$ *convexa, então* $\partial\psi$ *é monótono.*

Demonstração. Dados $x,\ y \in \mathbb{R}^n$, tal que $z \in \partial\psi(x)$ e $w \in \partial\psi(y)$, temos, em particular, que

$$\begin{aligned} \psi(y) &\geqslant \psi(x) + \langle z, y - x\rangle \\ \psi(x) &\geqslant \psi(y) + \langle w, x - y\rangle. \end{aligned}$$

Então $\langle w - z, y - x\rangle \geqslant 0$. Portanto $\partial\psi$ é um operador monótono. □

Definição 1.5.6. *Seja* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador ponto conjunto. Então* T *é monótono maximal se*

$$\langle v - u, y - x\rangle \geqslant 0,\ \forall\ y \in \mathrm{Dom}(T),\ v \in T(y) \Rightarrow u \in T(x), \tag{1.6}$$

sendo $x \in \mathrm{Dom}(T)$. *Equivalente a dizer que o gráfico de* T *não está contido no gráfico de qualquer outro operador monótono.*

Exemplo 1.5.2. *Seja* $\psi : \mathbb{R}^n \to \mathbb{R}$ *uma função própria, semicontínua inferiormente e convexa então* $\partial\psi$ *é um operador monótono maximal.*

Demonstração. Ver o Corolário 31.5.2 em [28], página 341. □

Proposição 1.5.1. *Seja* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *operador maximal monótono. Então:*

a) $T(x)$ *é fechado e convexo para qualquer* $x \in \mathrm{Dom}(T)$.

b) O gráfico de T *é fechado.*

Demonstração. a) Pela definição de operador monótono maximal podemos escrever

$$T(x) = \bigcap_{(y,v)\in G(T)} \{u \in \mathbb{R}^n : \langle u - v, x - y\rangle \geqslant 0\}$$

Ou seja, $T(x)$ é a interseção de semi-espaços fechados, e portanto, é fechado e convexo.

b) Sejam $x^k \to x$, $u^k \in T(x^k)$, tal que $u^k \to u$. Consideremos $(y, v) \in G(T)$, então passando o limite em $\langle u^k - v, x^k - y\rangle \geqslant 0$, obtemos $\langle u - v, x - y\rangle \geqslant 0$, e portanto, $u \in T(x)$.

□

Tal resultado nos fornece que qualquer operador ponto conjunto monótono maximal tem o gráfico fechado e, portanto, é um operador contínuo. Finalizamos as considerações referentes a operadores ponto conjunto.

1.6 Funções Analíticas

Abordamos, nesta seção, conceitos de funções analíticas em $\mathbb{R}$ que são utilizados adiante para demonstrar alguns resultados do Capítulo 4, quando voltaremos a atenção para a análise da convergência sob a condição de Smale. Como referência para a elaboração desta nos embasamos em [17].

Definição 1.6.1. *Sejam* $I \subset \mathbb{R}$ *um intervalo aberto. Dizemos que uma função* $f : I \to \mathbb{R}$ *é analítica quando pode ser localmente expandida em séries de Taylor, ou seja, para cada* $a \in I$*, existe* $\varepsilon > 0$ *tal que a série de Taylor*

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

converge para $f(a+h)$ *quando* $|h| < \varepsilon$.

Exemplo 1.6.1. *A função* $f : (-1, 1] \to \mathbb{R}^n$ *definida por* $f(x) = \ln(1+x)$ *é analítica, conseguimos a expansão em séries de Taylor tomando* $a = 0$

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}.$$

Observação 1.6.1. *Uma condição necessária e suficiente para que a série* $\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} h^n$ *convirja para* $f(a+h)$ *é que* $\lim_{n\to\infty} r_n(h) = 0$*, onde* $r_n = \frac{f^{(n)}(a+\theta_n h)}{n!} h^n$*, com* $0 < \theta_n < 1$.

Proposição 1.6.1. *Se* $0 \leqslant t < 1$*, então* $\sum_{i=0}^{\infty} (i+2)(i+1)t^i = \frac{2}{(1-t)^3}$.

Demonstração. Consideremos a função analítica $g : (-1, 1) \to \mathbb{R}$, definida por

$$g(t) = (1-t)^{-1},$$

com derivadas

$$g'(t) = (1-t)^{-2}, \quad g''(t) = 2(1-t)^{-3}, \quad \ldots, \quad g^{(i)}(t) = i!(1-t)^{-(i+1)}. \tag{1.7}$$

Pela Definição 1.6.1, podemos escrever g da forma

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

Segue de (1.7) e da igualdade anterior, que $g(t) = \sum_{i=0}^{\infty} t^i$. Derivando duas vezes, temos

$$g''(t) = \sum_{i=0}^{\infty} (i+2)(i+1)t^i. \tag{1.8}$$

Combinando a igualdade (1.8) com a segunda equação em (1.7), concluímos a demonstração. □

Capítulo 2

O Método de Newton para Sistemas de Equações

Neste capítulo, desenvolvido a partir das referências [10], [7], [22], [23] e [25], enunciamos e demonstramos o teorema de Kantorovich, resultado elegante com diversas aplicações teóricas e práticas.

2.1 O Problema e o Método

Iniciamos esta seção com uma pequena discussão onde apresentamos o método de Newton, no qual a ideia básica para encontrar um zero de uma função diferenciável não linear consiste em aproximar o zero da função não linear por zeros de aproximações lineares sucessivas. Dito de outra forma, buscamos resolver a equação

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

sendo que $\Omega \subset \mathbb{R}^n$ é um conjunto aberto e $F : \Omega \to \mathbb{R}^n$ é uma função continuamente diferenciável não-linear. Construímos a aproximação linear de F, em torno do ponto inicial $x^0 \in \Omega$, ou seja,

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

Assim, resolvemos a equação linear

$$F(x^0) + F'(x^0)(x - x^0) = 0. \tag{2.1}$$

Note que não podemos garantir que $F'(x^0) \neq 0$. Se $F'(x^0)$ for não singular a equação (2.1) possui uma única solução, denotaremos por $x^1 = x^0 - F'(x^0)^{-1}F(x^0)$. Percebemos que há outro porém, x^1 pode não pertencer a Ω e ainda $F'(x^1)$ pode ser singular, para que possamos repetir o procedimento necessitamos de hipóteses apropriadas sobre o ponto inicial x^0 e F. Suponhamos que tal processo possa ser repetido infinitamente, podemos definir uma sequência $\{x^k\}$ contida em Ω da forma

$$x^{k+1} = x^k - F'(x^k)^{-1}F(x^k), \quad k = 0,\ 1, \dots . \tag{2.2}$$

Aqui, encontramos outro dilema, o limite da sequência $\{x^k\}$, se existir, pode não pertencer ao conjunto Ω.

Para contornar estes problemas e garantirmos a convergência, o ponto inicial deve ser tomado em uma vizinhança (adequada) da solução do problema em questão e a derivada do operador não linear em consideração deve ser não singular em tal solução.

Como podemos perceber o método de Newton tem a desvantagem de exigir, *a priori*, o conhecimento de algum zero do operador e hipóteses a respeito do comportamento de tal operador nesse zero. A seguir, estudamos o comportamento da sequência $\{t_k\}$, obtida pelo método de Newton aplicado a um polinômio particular.

2.2 O Método de Newton para uma função polinomial

Tomando como base a referência [10], analisamos o método de Newton aplicado para resolver $P(t) = 0$, sendo $P : \mathbb{R} \longrightarrow \mathbb{R}$, definida por

$$P(t) = \frac{\beta K}{2}t^2 - t + \eta. \tag{2.3}$$

Estabelecemos $\{t_k\}$ como a sequência gerada pelo método de Newton que tem como ponto inicial $t_0 = 0$. Denotemos as raízes do polinômio por

$$t_* = \frac{1}{\beta K}(1 - \sqrt{1 - 2h}), \quad t_{**} = \frac{1}{\beta K}(1 + \sqrt{1 - 2h}), \tag{2.4}$$

onde $h \equiv \beta K \eta$. Sendo $\eta,\ \beta, K > 0$, temos que

$$0 < t_* \leqslant t_{**},$$

sendo a desigualdade estrita entre t_* e t_{**} se, e somente se, $h < 1/2$.

Proposição 2.2.1. *A função escalar* P *tem uma menor raíz não negativa* $t_* \in (0, 1/(\beta K)]$. *Além disso, para cada* $t \in [0, t_*)$

$$P(t) > 0, \qquad P'(t) \leqslant \beta K(t - t_*) < 0.$$

Demonstração. A primeira parte da proposição segue do fato de $t_* \leqslant 1/(\beta K)$, que é uma consequência imediata das hipóteses sobre η e βK. Como $P(0) = \eta > 0$, P é estritamente positiva no intervalo $[0, t_*)$. Por (2.3) e $t_* \leqslant 1/(\beta K)$, segue que

$$P'(t) = \beta K t - 1 = \beta K(t - 1/(\beta K)) \leqslant \beta K(t - t_*).$$

Como $t < t_*$, obtemos o resultado. □

De acordo com a Proposição 2.2.1, temos que $P'(t) \neq 0$ para todo $t \in [0, t_*)$. Portanto, as iteradas de Newton estão bem definidas em $[0, t_*)$. Denotamos por n_P,

$$\begin{array}{rcccc} n_P & : & [0, t_*) & \longrightarrow & \mathbb{R} \\ & & t & \longmapsto & t - P(t)/P'(t). \end{array}$$

Perceba que até agora temos apenas uma única iteração do método de Newton bem definida em $[0, t_*)$. Mostraremos adiante que a iterada de Newton pode ser repetida infinitamente e partindo de qualquer ponto em $[0, t_*)$.

Proposição 2.2.2. *Para qualquer* $t \in [0, t_*)$

$$t_* - n_P(t) = -\frac{\beta K}{2P'(t)}(t_* - t)^2, \quad t < n_P(t) < t_*. \tag{2.5}$$

Em particular, n_P *aplica* $[0, t_*)$ *em* $[0, t_*)$.

Demonstração. Suponha $t \in [0, t_*)$. Como P é um polinômio de segundo grau e $P(t_*) = 0$

$$\begin{aligned} 0 &= \frac{\beta K t_*^2}{2} - t_* + \eta \\ &= \frac{\beta K t_*^2}{2} - t_* + \eta + t - t + \beta K t^2 - \beta K t^2 + \beta K t t_* - \beta K t t_* \\ &= P(t) + P'(t)(t_* - t) + \frac{\beta K}{2}(t_* - t)^2. \end{aligned}$$

Dividindo por $P'(t) \neq 0$, obtemos

$$t_* - t + \frac{P(t)}{P'(t)} = -\frac{\beta K}{2P'(t)}(t_* - t)^2.$$

Segue a igualdade em (2.5). Pela Proposição 2.2.1, temos que $P(t) > 0$ e $P'(t) < 0$. Combinando com a definição de n_P e a igualdade em (2.5), seguem as desigualdades

$t < n_P(t) < t_*$ e a última afirmação da proposição segue diretamente dessas desigualdades. Segue a igualdade em (2.5). Pela Proposição 2.2.1, temos que $P(t) > 0$ e $P'(t) < 0$. Combinando com a definição de n_P e a igualdade em (2.5), seguem as desigualdades $t < n_P(t) < t_*$ e a última afirmação da proposição segue diretamente dessas desigualdades. □

A Proposição 2.2.2 mostra que dado $t \in [0, t_*)$ qualquer, a sequência $\{n_P^k(t)\}$,

$$n_P^0(t) = t, \quad , n_P^{k+1}(t) = n_P\big((n_P^k(t))\big), \quad , k = 0,\ 1, \ldots.$$

está bem definida, estritamente crescente em $[0, t_*)$, e então convergente. Portanto, o método de Newton para resolver $P(t) = 0$, dado o ponto inicial $t_0 = 0$, gera uma sequência infinita $t_k = n_P^k(t_0)$, que pode ser também definida como

$$t_0 = 0, \quad , t_{k+1} = n_P(t_k), \ \ k = 0,\ 1, \ldots. \tag{2.6}$$

Agora, vamos estabelecer uma fórmula para o termo t_k, tal reformulação consiste em algumas manipulações e de fatos que conhecemos a respeito da sequência $\{t_k\}$. Como t_* e t_{**} são raízes do polinômio $P(t) = \dfrac{\beta K}{2}(t - t_*)(t - t_{**})$, e $P'(t) = \dfrac{\beta K}{2}[(t - t_*) + (t - t_{**})]$. Pelas equações anteriores e por (2.6), temos que

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

Analogamente, obtemos

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

Combinando as duas equações acima

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

Suponha que $h < 1/2$, nesse caso $t_* < t_{**}$. Definimos $\theta = t_*/t_{**} < 1$, assim, aplicando indução sobre k, temos que

$$\frac{t_k - t_*}{t_k - t_{**}} = \theta^{2^k}.$$

Então $t_k - t_* = \theta^{2^k}(t_k - t_{**})$, assim

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

Desenvolvendo a subtração entre parênteses, obtemos

$$t_k = t_* - \frac{\theta^{2^k}}{1-\theta^{2^k}} \frac{2\sqrt{1-2h}}{\beta K}, \quad k = 0,\ 1, \ldots. \tag{2.7}$$

Usamos a igualdade anterior no próximo resultado, onde mostramos a convergência Q-linear para t_*, além de obtermos a convergência Q-quadrática, a partir de uma restrição adicional.

Corolário 2.2.1. *A sequência $\{t_k\}$ é bem definida, estritamente crescente e está contida em $[0, t_*)$. Além disso, converge Q-linearmente para t_*, como segue*

$$t_* - t_{k+1} = \frac{\beta K}{-2P'(t_k)} (t_* - t_k)^2 \leqslant \frac{1}{2}(t_* - t_k), \quad k = 0,\ 1, \ldots. \tag{2.8}$$

Se $h < 1/2$, então a sequência $\{t_k\}$ converge Q-quadraticamente.

$$t_* - t_{k+1} = \frac{1-\theta^{2^k}}{1+\theta^{2^k}} \frac{\beta K}{2\sqrt{1-2h}} (t_* - t_k)^2 \leqslant \frac{\beta K}{\sqrt{1-2h}} (t_* - t_k)^2, \quad k = 0,\ 1, \ldots, \tag{2.9}$$

onde $\theta = t_*/t_{**} < 1$.

Demonstração. A primeira parte do Corolário segue da proposição anterior. Pela Proposição 2.2.2 temos que para qualquer k

$$t_* - n_P(t_k) = \frac{\beta K}{-2P'(t_k)} (t_* - t_k)^2.$$

Da definição de $n_P(t_k)$, em (2.6), segue a igualdade em (2.8). Como $t_k \in [0, t_*)$, decorre da Proposição 2.2.1 que

$$P'(t_k) \leqslant \beta K (t_k - t_*) < 0.$$

Multiplicando a inequação pelo termo $(t_* - t_k)/(2P'(t_k)) < 0$, obtemos a inequação em (2.8). Agora, supondo que $h < 1/2$, equivalentemente $t_* < t_{**}$. E lembrando que podemos expressar t_k por

$$t_k = t_* - \frac{\theta^{2^k}}{1-\theta^{2^k}} \frac{2\sqrt{1-2h}}{\beta K}, \ \ k = 0,\ 1, \ldots.$$

Segue de (2.4) que

$$\begin{aligned}
t_k &= \frac{1}{\beta K}(1 - \sqrt{1-2h}) - \frac{\theta^{2^k}}{1-\theta^{2^k}} \frac{2\sqrt{1-2h}}{\beta K} \\
&= \frac{1}{\beta K (1-\theta^{2^k})} (1 - \sqrt{1-2h} - \theta^{2^k} + \theta^{2^k}\sqrt{1-2h} - 2\theta^{2^k}\sqrt{1-2h}) \\
&= -\frac{1+\theta^{2^k}}{1-\theta^{2^k}} \frac{\sqrt{1-2h}}{\beta K} + \frac{1}{\beta K}.
\end{aligned}$$

Substituindo t_k em $P'(t_k) = \beta K t_k - 1$, obtemos a igualdade

$$P'(t_k) = -\frac{1+\theta^{2^k}}{1-\theta^{2^k}}(\sqrt{1-2h}). \tag{2.10}$$

Combinando (2.10) e (2.8)

$$\begin{aligned} t_* - t_{k+1} &= \frac{\beta K}{-2P'(t_k)}(t_{*-t_k})^2 \\ &= \frac{\beta K}{-2}\left(-\frac{1-\theta^{2^k}}{1+\theta^{2^k}}\frac{1}{\sqrt{1-2h}}\right)(t_* - t_k)^2 \\ &= \frac{1-\theta^{2^k}}{1+\theta^{2^k}}\frac{\beta K}{\sqrt{1-2h}}(t_* - t_k)^2. \end{aligned}$$

Assim, segue a igualdade de (2.9). Como $(1-\theta^{2^k})/(1+\theta^{2^k}) \leqslant 1$ temos a desigualdade em (2.9). □

2.3 O Teorema de Newton-Kantorovich para Equações Não Lineares

O teorema de Kantorovich atribui-se de certas condições semilocais para garantir a existência de uma única solução da equação não-linear $F(x) = 0$, sem a necessidade de conhecer zeros do operador, pois tais condições são impostas apenas sobre o ponto inicial. Empregando de modo construtivo o método de Newton, o teorema garante a convergência para uma solução do processo iterativo. Apresentamos uma demonstração obtida por meio de funções majorantes.

Teorema 2.3.1. *Sejam* $D \subset \mathbb{R}^n$ *aberto e convexo e* $F : D \to \mathbb{R}^n$ *diferenciável, tal que*

$$\|F'(x) - F'(y)\| \leqslant K\|x-y\|, \quad \forall\ x,\ y \in D.$$

Para algum $x^0 \in D$, *suponha que* $[F'(x^0)]^{-1}$ *é definida em todo* $\mathbb{R}^n$ *e que* $h \equiv \beta K \eta \leqslant 1/2$, *onde* $\|[F'(x^0)]^{-1}\| \leqslant \beta$ *e* $\|[F'(x^0)]^{-1}F(x^0)\| \leqslant \eta$. *Defina*

$$t_* = \frac{1}{\beta K}(1-\sqrt{1-2h}), \quad t_{**} = \frac{1}{\beta K}(1+\sqrt{1-2h}) \tag{2.11}$$

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

A prova segue como consequência dos seguintes Lemas.

Lema 2.3.1. *Sejam* $\{x^k\}$ *uma sequência em* $\mathbb{R}^n$ *e* $\{t_k\}$ *uma sequência não negativa de números reais tais que*

$$\|x^{k+1} - x^k\| \leqslant t_{k+1} - t_k, \quad k = 0,\ 1, \ldots, \tag{2.12}$$

e $t_k \to t_* < \infty$. *Então existe um* $x^* \in \mathbb{R}^n$ *tal que* $x^k \to x^*$ *e*

$$\|x^* - x^k\| \leqslant t_* - t_k, \quad k = 0,\ 1, \ldots. \tag{2.13}$$

Demonstração. Usando a desigualdade triangular e (2.12), temos que

$$\begin{aligned} \|x^{k+p} - x^k\| &\leqslant \sum_{n=1}^{p} \|x^{k+n} - x^{k+(n-1)}\| \\ &\leqslant \sum_{n=1}^{p} (t_{k+n} - t_{k+(n-1)}) \\ &= t_{k+p} - t_k. \end{aligned}$$

Como $t_k \to t_*$, segue que $\{x^k\}$ é uma sequência de Cauchy, logo existe $x^* \in \mathbb{R}^n$ tal que $x^k \to x^*$. Fazendo $p \to \infty$ em $\|x^{k+p} - x^k\| \leqslant t_{k+p} - t_k$, segue (2.13). □

Lema 2.3.2. *Para todo* $x \in Q \equiv \{x;\ \|x - x^0\| < 1/(\beta K)\} \cap D$, $[F'(x)]^{-1}$ *é definida em todo* $\mathbb{R}^n$ *e*

$$\|[F'(x)]^{-1}\| \leqslant \frac{\beta}{1 - \beta K \|x - x^0\|}. \tag{2.14}$$

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

$$\|N(N(x)) - N(x)\| \leqslant \frac{1}{2}\left(\frac{\beta K \|x - N(x)\|^2}{1 - \beta K \|x^0 - N(x)\|}\right). \tag{2.15}$$

Demonstração. A primeira afirmação segue como consequência do Lema de Banach. Para a segunda parte do Lema, usando (2.14), temos que

$$\begin{aligned} \|N(N(x) - N(x))\| &= \|[F'(N(x))]^{-1}F'(N(x))\| \\ &\leqslant \frac{\beta}{1 - \beta K \|x^0 - N(x)\|}\|F(N(x))\|. \end{aligned}$$

Vamos estimar $\|F(N(x))\|$, para tal usaremos a igualdade $F(x) + F'(x)(N(x) - x) = 0$, o

Teorema do Valor Médio e o fato de F' ser Lipschitz

$$\begin{aligned}
\|F(N(x))\| &= \|F(N(x)) - F(x) - F'(x)(N(x)-x)\| \\
&= \left\| \int_0^1 F'(\theta N(x) + (1-\theta)x)(N(x)-x)d\theta - F'(x)(N(x)-x) \right\| \\
&\leqslant \|N(x)-x\| \int_0^1 \|F'(\theta N(x) + (1-\theta)x) - F'(x)\| d\theta \\
&\leqslant \|N(x)-x\| \int_0^1 K\theta \|N(x)-x\| d\theta \\
&= \frac{K}{2}\|N(x)-x\|^2.
\end{aligned}$$

Segue que $\|N(N(x)) - N(x)\| \leqslant \dfrac{1}{2}\left(\dfrac{\beta K\|x - N(x)\|^2}{1 - \beta K\|x^0 - N(x)\|}\right)$. □

Lema 2.3.3. *A sequência de Newton* $\{x^k\}$ *está bem definida e é majorada pela sequência definida por*

$$t_{k+1} = t_k - \frac{(\beta K/2)t_k^2 - t_k + \eta}{\beta K t_k - 1}, \quad t_0 = 0 \tag{2.16}$$

Além disso, $t_k \to t_*$, *onde* t_* *é definido em (2.4).*

Demonstração. Já mostramos, na Seção 2.2, que t_k são as iteradas de Newton para o polinômio $P(t) = \dfrac{\beta K}{2}t^2 - t + \eta$ com raízes t_*e t_{**}, então pelo método de Newton $t_k \to t_*$. Suponha que existem $x^1, \ldots, x^k$ e $\|x^i - x^{i-1}\| \leqslant t_i - t_{i-1}$ (para $i = 1, \ldots, k$). É imediato que a desigualdade vale para $i = 1$, pois

$$\|x^1 - x^0\| = \|[F'(x^0)]^{-1}F(x^0)\| \leqslant \eta = \|t_1 - t_0\|.$$

Por outro lado, depois de algumas manipulações, vale a desigualdade

$$\|x^k - x^0\| \leqslant \sum_{i=1}^{k} \|x^i - x^{i-1}\| \leqslant \sum_{i=1}^{k} (t_i - t_{i-1}) = t_k - t_0 \leqslant t_*.$$

Assim, pelo Lema 2.3.2, $\{x^k\}$ está bem definida e temos:

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

Mostremos, agora, a última igualdade, que segue de algumas manipulações.

$$
\begin{aligned}
\frac{((\beta K)/2)(t_k - t_{k-1})^2}{1 - \beta K t_k} &= \frac{((\beta K)/2){t_k}^2 - t_k + \eta + t_k - \eta + ((\beta K)/2){t^2}_{k-1} - \beta K t_k t_{k-1}}{1 - \beta K t_k} \\
&= \frac{-P(t_k)}{P'(k_k)} + \frac{t_k - \eta + ((\beta K)/2){t^2}_{k-1} - \beta K t_k t_{k-1}}{-P'(t_k)} \\
&= \frac{-P(t_k)}{P'(t_k)} + \frac{t_k(1 - \beta K t_{k-1}) - \eta + ((\beta K)/2){t^2}_{k-1}}{-P'(t_k)} \\
&= \frac{-P(t_k)}{P'(t_k)} + \frac{-t_k(P'(t_k)) + ((\beta K)/2)t^2_{k-1} - \eta}{-P'(t_k)} \\
&= \frac{-P(t_k)}{P'(t_k)} + \frac{-\left(t_{k-1} - \dfrac{P(t_{k-1})}{P'(t_{k-1})}\right)P'(t_{k-1}) + ((\beta K)/2){t^2}_{k-1} - \eta}{-P'(t_k)} \\
&= \frac{-P(t_k)}{P'(t_k)} + \frac{-t_{k-1}P'(t_{k-1}) + P(t_{k-1}) + ((\beta K)/2){t^2}_{k-1} - \eta}{-P'(t_k)} \\
&= \frac{-P(t_k)}{P'(t_k)} + \frac{-t_{k-1}(\beta K t_{k-1} - 1) + P(t_{k-1}) + ((\beta K)/2){t^2}_{k-1} - \eta}{-P'(t_k)} \\
&= \frac{-P(t_k)}{P'(t_k)} + \frac{-\beta K {t^2}_{k-1} + t_{k-1} + P(t_{k-1}) + ((\beta K)/2){t^2}_{k-1}\eta}{-P'(t_k)} \\
&= \frac{-P(t_k)}{P'(t_k)} + \frac{-(\beta K/2){t^2}_{k-1} + t_{k-1} - \eta + P(t_{k-1})}{-P'(t_k)} \\
&= \frac{-P(t_k)}{P'(t_k)} \\
&= -\frac{((\beta K)/2)t_k^2 - t_k + \eta}{\beta K t_k - 1} \\
&= t_{k+1} - t_k.
\end{aligned}
$$

□

A partir dos lemas anteriores vamos demonstrar o Teorema 2.3.1.

Demonstração. (Teorema 2.3.1) Os Lemas 2.3.1 e 2.3.3 garantem a existência de x^* em S, tal que $x^k \to x^*$. Que x^* é a solução segue a partir de:

$$
\begin{aligned}
\|F(x^k)\| &= \|F'(x^k)(x^{k+1} - x^k)\| \\
&= \|F'(x^k)(x^{k+1} - x^k) - F'(x^0)(x^{k+1} - x^k) + F'(x^0)(x^{k+1} - x^k)\| \\
&\leqslant \left(\|F'(x^0)\| + \|F'(x^0) - F'(x^k)\|\right)\|x^{k+1} - x^k\| \\
&\leqslant (\|F'(x^0)\| + Kt_*)\,\|x^{k+1} - x^k\|.
\end{aligned}
$$

Como $\|x^{k+1} - x^k\| \to 0$ e F é contínua em S concluímos que $\|F(x^k)\| \to 0$, quando $k \to \infty$. Desse modo, $F(x^*) = 0$. Se $h < \dfrac{1}{2}$ então as raízes t_* e t_{**} são distintas e, da

Seção 2.2, temos que a ordem de convergência de t_k para t_* é pelo menos quadrática, consequentemente a ordem de convergência de x^k para x^* é pelo menos quadrática. $\square$

É importante salientar que existem outras versões do Teorema 2.3.1, que diferem em hipóteses e resultados obtidos. Neste trabalho mencionamos apenas uma delas, com base em [10].

Capítulo 3

O Método de Newton para Equações Generalizadas

Neste capítulo generalizamos os resultados obtidos no capítulo anterior, no qual trabalhamos a convergência da solução, pelo método de Newton, de uma equação não linear. Apresentamos resultados obtidos em [30] no qual a resolução da equação generalizada é baseada na linearização parcial. Relacionamos a função majorante com o operador F, mostramos a boa definição e convergência do método de Newton em uma determinada região, verificamos taxas de convergência para a solução, mostramos o raio ótimo de convergência e, por fim, provamos a unicidade da convergência numa região.

3.1 O Problema e o Método

Iniciamos esta seção apresentando a definição de linearização parcial, importante ferramenta na demonstração do próximo teorema.

Definição 3.1.1. *Seja* $\Omega \subset \mathbb{R}^n$*aberto e não vazio,* $\mathsf{F} : \Omega \to \mathbb{R}^n$ *com derivada* F' *e* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$*. A linearização parcial da aplicação* $\mathsf{F} + \mathsf{T}$ *em* $x \in \mathbb{R}^n$ *é a aplicação ponto conjunto* $\mathsf{L}_\mathsf{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *dada por*

$$\mathsf{L}_\mathsf{F}(x, y) := \mathsf{F}(x) + \mathsf{F}'(x)(y - x) + \mathsf{T}(y). \tag{3.1}$$

Para cada $x \in \mathbb{R}^n$*, a inversa* $\mathsf{L}_\mathsf{F}(x, \cdot)^{-1} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *da aplicação* $\mathsf{L}_\mathsf{F}(x, \cdot)$ *em* $z \in \mathbb{R}^n$ *é denotada por*

$$\mathsf{L}_\mathsf{F}(x, z)^{-1} := \{y \in \mathbb{R}^n : z \in \mathsf{F}(x) + \mathsf{F}'(x)(y - x) + \mathsf{T}(y)\}. \tag{3.2}$$

Observação 3.1.1. *Se na Definição 3.1.1 tivéssemos* $T \equiv 0$, $z = 0$ *e* $F'(x)$ *inversível, então* $L_F(x,0)^{-1} = x - F'(x)^{-1}F(x)$ *seria a iterada de Newton para resolver a equação* $F(x) = 0$.

Retomamos, aqui, o problema inicial no qual buscamos a convergência do método de Newton para a solução da equação generalizada

$$F(x) + T(x) \ni 0.$$

Semelhante ao que vimos para equações, reformulamos a equação majorante para analisar a convergência semilocal e propriedades do método de Newton, que consiste em gerar uma sequência $F(x^k) + F'(x^k)(x^{k+1} - x^k) + T(x^{k+1})$, partindo de um ponto inicial x^0. Notemos que tal método pode ser pensado como um método tipo-Newton baseado na linearização parcial de $F + T$.

3.2 Análise Semilocal do Método

Nesta seção apresentamos o nosso principal resultado, estabelecemos algumas condições iniciais para garantir a boa definição e a convergência do método de Newton para a solução da equação (1) em uma determinada região. Como se trata de uma generalização do Teorema de Newton-Kantorovich, percebemos a notória relação entre as hipóteses estabelecidas. Sua demonstração segue de uma série de lemas que são apresentados nas seções subsequentes, onde a ideia central para mostrar a convergência do método é o princípio majorante, adaptado ao problema.

Teorema 3.2.1. *Consideremos o aberto* Ω *subconjunto não vazio de* $\mathbb{R}^n$, $F : \Omega \to \mathbb{R}^n$ *contínua com derivada* F' *também contínua,* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *um operador monótono maximal e* $x^0 \in \Omega$. *Suponha que* $F'(x^0)$ *seja um operador positivo e* $\widehat{F'(x^0)}^{-1}$ *existe. Seja* $R > 0$ *e* $\kappa := \sup\{t \in [0, R) : \ B(x^0, t) \subset \Omega\}$. *Suponha que existe* $f : [0, R) \to \mathbb{R}$ *duas vezes continuamente diferenciável, tal que*

$$\|\widehat{F'(x^0)}^{-1}\| \, \|F'(y) - F'(x)\| \leqslant f'\left(\|x - y\| + \|x - x^0\|\right) - f'\left(\|x - x^0\|\right), \tag{3.3}$$

para quaisquer $x, \ y \in B(x^0, \kappa)$. *Além disso, vamos supor que*

$$\|x^1 - x^0\| \leqslant f(0), \tag{3.4}$$

e que as seguintes condições ocorrem,

(H1) $f(0) > 0$ *e* $f'(0) = -1$.

(H2) f' *é convexa e estritamente crescente.*

(H3) $f(t) = 0$, *para algum* $t \in (0, R)$.

Então, f *tem um menor zero* $t_* \in (0, R)$, *as sequências geradas pelo método de Newton para resolver a equação generalizada* $F(x) + T(x) \ni 0$ *e a equação* $f(t) = 0$, *com ponto inicial* x^0 *e* $t_0 = 0$

$$0 \in F(x^k) + F'(x^k)(x^{k+1} - x^k) + T(x^{k+1}), \quad t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)}, \quad k = 0,\ 1, \ldots \tag{3.5}$$

estão bem definidas, $\{t_k\}$ *é estritamente crescente, contida em* $(0, t_*)$ *e converge para* t_*, $\{x^k\}$ *está contida em* $B(x^0, t_*)$ *e converge para* $x^* \in B[x^0, t_*]$. *Além disso, as sequências* $\{x^k\}$ *e* $\{t_k\}$ *satisfazem,*

$$\|x^* - x^k\| \leqslant t_* - t_k, \quad \|x^* - x^{k+1}\| \leqslant \frac{t_* - t_{k+1}}{(t_* - t_k)^2}\|x^* - x^k\|^2, \tag{3.6}$$

para todo $k = 0,\ 1, \ldots,$ *e as sequências* $\{x^k\}$ *e* $\{t_k\}$ *convergem Q-linearmente, como segue*

$$\|x^* - x^{k+1}\| \leqslant \frac{1}{2}\|x^* - x^k\|, \quad t_* - t_{k+1} \leqslant \frac{1}{2}(t_* - t_k), \quad k = 0,\ 1, \ldots \tag{3.7}$$

Se adicionarmos a seguinte hipótese

(H4) $f'(t_*) < 0$.

Então as sequências $\{x^k\}$ *e* $\{t_k\}$ *convergem Q-quadraticamente, como segue*

$$\|x^* - x^{k+1}\| \leqslant \frac{f''_-(t_*)}{-2f'(t_*)}\|x^* - x^k\|, \quad t_* - t_{k+1} \leqslant \frac{f''_-(t_*)}{-2f'(t_*)}(t_* - t_k)^2, \tag{3.8}$$

para todo $k = 0,\ 1, \ldots$.

Exemplo 3.2.1. *Como exemplo de função que satisfaz as hipóteses (H1), (H2) e (H3) podemos citar* $f : [0, +\infty) \to \mathbb{R}$ *definida por*

$$f(t) = t^{1+p} - t + b,\ b > 0 \text{ e } p = 2n,\ n \in \mathbb{N}.$$

Observação 3.2.1. *A condição* $\widehat{F'(x^0)}$ *é um operador positivo associado a existência de* $\widehat{F'(x^0)}$, *implica, pelo Lema 3.3.3, que*

$$\langle F'(x^0)x, x\rangle \geqslant \|x\|^2,\ \forall x \in \mathbb{R}^n \tag{3.9}$$

Além disso, se $T \equiv \{0\}$ então a função $F^*(x) := F(x^0) + F(x^0)(x - x^0)$ é fortemente monótona para cada x em uma vizinhança de x^*. De fato, para cada $x,\ y \in \mathbb{R}^n$,

$$\langle F^*(x) - F^*(y), x - y\rangle = \langle F'(x^0)(x - y), x - y\rangle \geqslant c\|x - y\|^2,$$

onde a última desigualdade segue a partir de (3.9). Isto nos fornece, em particular, que o problema

$$F(x^k) + F'(x^k)(x^{k+1} - x^k) = 0,\ k = 0,\ 1, \ldots$$

tem uma única solução para cada $k = 0,\ 1, \ldots$ (ver [5]), em particular implica que $F'(x^0)^{-1}$ existe, portanto, quando $T \equiv \{0\}$, as hipóteses $F'(x^0)$ ser um operador positivo e $\widehat{F'(x^0)}^{-1}$ existe, reduzindo a condição clássica $F'(x^0)^{-1}$ existe.

3.3 O Princípio Majorante

Nesta seção, embasados em [11] e [30], estabelecemos alguns resultados a respeito da função majorante $f : [0, R) \to \mathbb{R}$, e algumas relações entre a função majorante e a aplicação ponto conjunto $F + T$. Para tal, começamos com algumas proposições.

Proposição 3.3.1. *A função* f *tem uma menor raíz* $t_* \in (0, R)$*, é estritamente convexa e*

$$f(t) > 0, \quad f'(t) < 0, \quad t < t - \frac{f(t)}{f'(t)} < t_*, \forall\ t \in [0, t_*). \tag{3.10}$$

Além disso, $f'(t_*) \leqslant 0$ *e*

$$f'(t_*) < 0 \Longleftrightarrow \exists\ t \in (t_*, R),\ f(t) < 0. \tag{3.11}$$

Demonstração. Como f é contínua em $[0, R)$ e tem um zero no aberto $(0, R)$ *(H3)*, deve ter um menor zero $t_* > 0$, uma vez que $f(0) > 0$ *(H1)*. Desde que f' é estritamente crescente *(H2)*, temos que f é estritamente convexa.

A primeira desigualdade em (3.10) segue da suposição $f(0) > 0$ e da definição de t_* como a menor raíz de f. Desde que f é estritamente convexa, temos da Proposição 1.4.2 que

$$0 = f(t_*) > f(t) + f'(t)(t_* - t),\ t \in [0, R),\ t \neq t_*. \tag{3.12}$$

Se $t \in [0, t_*)$ então $f(t) > 0$ e $t_* - t > 0$, que, combinado com (3.12) produz a segunda desigualdade em (3.10). A terceira desigualdade em (3.10) segue das duas anteriores.

Obtemos a última desigualdade em (3.10) dividindo a inequação em (3.12) por $-f'(t)$, que é estritamente positivo,

$$0 = \frac{f(t_*)}{-f'(t)} > \frac{f(t)}{-f'(t)} - t_* + t.$$

Para a última parte da proposição, como $f > 0$ em $[0, t_*)$ e $f(t_*) = 0$, devemos ter $f'(t_*) \leqslant 0$. Em (3.11), a ida segue imediato. Para demonstrar a outra implicação, invertemos t e t_* em (3.12) e note que $f(t_*) = 0$. $\square$

Iremos, estabelecer o maior intervalo em que o método de Newton para a função majorante está bem definido. Consideremos

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

Claramente temos que $t_* \leqslant \bar{t}$. Desde que $f'(t) < 0$ para todo $t \in [0, \bar{t})$, a iterada de Newton da função majorante f está bem definida em $[0, \bar{t})$. Vamos definir a aplicação $n_f : [0, \bar{t}) \to \mathbb{R}$, tal que

$$n_f(t) = t - \frac{f(t)}{f'(t)}. \tag{3.13}$$

Mostramos no próximo resultado que a sequência gerada pelo método de Newton para resolver $f(t) = 0$ é bem comportada em determinado intervalo e é estritamente crescente, além disso, estabelecemos taxas de convergência.

Proposição 3.3.2. *Para todo* $t \in [0, t_*)$*, temos que* $n_f(t) \in [0, t_*)$*, e também*

$$t < n_f(t), \quad t_* - n_f(t) \leqslant \frac{1}{2}(t_* - t), \quad \forall\, t \in [0, t_*). \tag{3.14}$$

Se f *satisfaz* $f'(t_*) < 0$ *(H4), então*

$$t_* - n_f(t) \leqslant \frac{f''_-(t_*)}{-2f'(t_*)}(t_* - t)^2, \quad \forall\, t \in [0, t_*). \tag{3.15}$$

Demonstração. As duas primeiras afirmações seguem da última desigualdade em (3.10).

Para mostrar a segunda desigualdade em (3.14) consideremos algum $t \in [0, t_*)$, sendo $f(t_*) = 0$. Segue de (3.13) e da continuidade de f' que

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

Como f' é convexa e $t < t_*$, segue do Corolário 1.4.1 que

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

Lembrando que $-f'(t)$ é positivo, combinando as equações anteriores, segue

$$\begin{aligned} t_* - n_f(t) &\leqslant \frac{1}{f'(t)}) \int_t^{t_*} [f'(t_*) - f'(t)] \frac{u-t}{t_*-t} du \\ &= \left[\frac{f'(t_*) - f'(t)}{-f'(t)}\right] \frac{1}{t_* - t} \int_t^{t_*} (u-t) du \\ &= \frac{1}{2}\left[\frac{f'(t_*) - f'(t)}{-f'(t)}\right] \frac{1}{t_* - t}(t_* - t)^2. \end{aligned}$$

A partir disso, obtemos a seguinte desigualdade

$$t_* - n_f(t) \leqslant \frac{1}{2}\left[\frac{f'(t_*) - f'(t)}{-f'(t)}\right](t_* - t). \tag{3.16}$$

Sendo $f'(t_*) \leqslant 0$, f' estritamente crescente e lembrando que $t \in [0, t_*)$, segue que

$$0 < f'(t_*) - f'(t) \leqslant -f'(t) \implies 0 < \frac{f'(t_*) - f'(t)}{-f'(t)} \leqslant 1.$$

Assim,

$$t_* - n_f(t) \leqslant \frac{1}{2}(t_* - t).$$

Para finalizar, vamos assumir que f satisfaz *(H4)*, tome $t \in [0, t_*)$. Como f' é crescente, $f'(t_*) < 0$ e $f'(t) < 0$, temos que

$$\frac{1}{-f'(t)} \leqslant \frac{1}{f'(t_*)}$$

Logo,

$$\begin{aligned} \frac{f'(t_*) - f'(t)}{-f'(t)} &\leqslant \frac{f'(t_*) - f'(t)}{-f'(t_*)} \\ &= \frac{1}{f'(t_*)} \frac{f'(t_*) - f'(t)}{t_* - t}(t_* - t) \\ &\leqslant \frac{f''_-(t_*)}{-f'(t_*)}(t_* - t). \end{aligned}$$

Combinando a inequação anterior com (3.16)

$$\frac{2}{(t_* - t)}[t_* - n_f(t)] \leqslant \frac{f'(t_*) - f'(t)}{-f'(t)} \leqslant \frac{f''_-(t_*)}{-f'(t_*)}(t_* - t).$$

Dividindo a inequação pelo termo positivo $2/(t_* - t)$ obtemos

$$t_* - n_f(t) \leqslant \frac{f''_-(t_*)}{-2f'(t_*)}(t_* - t)^2, \quad \forall\, t \in [0, t_*).$$

Concluímos, assim, a demonstração. □

A definição sobre t_k em (3.5) é equivalente a

$$t_0 = 0, \quad t_{k+1} = n_f(t_k), \quad k = 0,\ 1, \dots \tag{3.17}$$

Temos que a sequência $\{t_k\}$ é monótona crescente em $[0, t_*)$ e t_* é o zero de f neste intervalo, portanto $t_k \to t_*$. O próximo resultado contém as principais propriedades de convergência da sequência $\{t_k\}$, que é consequência direta da Proposição 3.3.2.

Corolário 3.3.1. *A sequência $\{t_k\}$ está bem definida, é estritamente crescente e está contida em $[0, t_*)$. Além disso, $\{t_k\}$ satisfaz a segunda desigualdade em (3.7) e converge Q-linearmente para t_*. Se também satisfaz a condição (H4), então $\{t_k\}$ satisfaz a segunda desigualdade em (3.8) e converge Q-quadraticamente.*

Portanto, obtemos todos os resultados necessários acerca da sequência majorante $\{t_k\}$ no Teorema 3.2.1. Agora, vamos estabelecer algumas relações entre a função majorante e a aplicação ponto conjunto $F + T$. Na sequência vamos mostrar que a linearização parcial do operador $F + T$ tem uma aplicação ponto conjunto inversa, que é Lipschitz em uma determinada vizinhança de x^0. Desde que as iteradas de Newton iniciando em um ponto dessa vizinhança convergem para um zero da linearização parcial de $F + T$ em tal ponto, é conveniente primeiro estudar o erro de linearização de F em um ponto de Ω.

$$E_F(x, y) := F(y) - [F(x) + F'(x)(y - x)]\,; \quad x,\ y \in \Omega. \tag{3.18}$$

No próximo resultado limitamos tal erro pelo erro de linearização da função majorante f, denotado por

$$e_f(t, u) := f(u) - [f(t) + f'(t)(u - t)]; \quad t,\ u \in [0, R). \tag{3.19}$$

Lema 3.3.1. *Consideremos $x,\ y \in B(x^0, R)$ e $0 \leqslant t < v < R$. Se $\|x - x^0\| \leqslant t$ e $\|y - x\| \leqslant v - t$ então*

$$\|(F'(x^0))^{-1} E_F(x, y)\| \leqslant e_f(t, v)\frac{\|y - x\|^2}{(v - t)^2}. \tag{3.20}$$

Demonstração. Primeiramente, temos que a combinação convexa $x + \lambda(y - x) \in B(x^0, R)$ para todo $\lambda \in [0, 1]$. Desde que F é continuamente diferenciável em Ω, o erro de linearização de F em (3.18) é equivalente a

$$E_F(x, y) = \int_0^1 [F'(x + \lambda(y - x)) - F'(x)](y - x)d\lambda.$$

Que combinado com a suposição em (3.3), nos fornece

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

Agora, usando a convexidade de f' e as hipóteses $\|x-x^0\| \leqslant t$ e $\|y-x\| \leqslant \nu - t$, $\nu < R$ temos

$$\begin{aligned}f'(\|x-x^0\|+\lambda\|y-x\|)-f'(\|x-x^0\|) &\leqslant f'(t+\lambda\|y-x\|)-f'(t)\\ &\leqslant [f'(t+\lambda(\nu-t))-f'(t)]\frac{\|y-x\|}{\nu-t},\end{aligned}$$

para todo $\lambda \in [0,1]$. Combinando tais desigualdades, concluímos que

$$\begin{aligned}\|(F'(x^0))^{-1}E_F(x,y)\| &\leqslant \int_0^1 [f'(t+\lambda(\nu-t))-f'(t)]\frac{\|y-x\|^2}{\nu-t}d\lambda\\ &= \left[\frac{1}{\nu-t}(f(\nu)-f(t))-f'(t)\right]\frac{\|y-x\|^2}{\nu-t}\\ &= [f(\nu)-(f(t)+f'(t)(\nu-t))]\frac{\|y-x\|^2}{(\nu-t)^2}\\ &= e_f(t,\nu)\frac{\|y-x\|^2}{(\nu-t)^2}.\end{aligned}$$

□

A seguir mostramos $\widehat{F'(x)}^{-1}$ existe para todo $x \in B(x^0, t_0)$, partindo do pressuposto que $\widehat{F'(x^0)}^{-1}$ existe.

Lema 3.3.2. *Seja* $x^0 \in \Omega$ *tal que* $\widehat{F'(x^0)}$ *é um operador positivo e* $\widehat{F'(x^0)}^{-1}$ *existe. Se* $\|x-x^0\| \leqslant t < t_*$, *então* $\widehat{F'(x)}$ *é um operador positivo e* $\widehat{F'(x)}^{-1}$ *existe. Além disso,*

$$\|\widehat{F'(x)}^{-1}\| \leqslant -\frac{\|\widehat{F'(x^0)}^{-1}\|}{f'(t)}.$$

Demonstração. Primeiramente, note que

$$\|\widehat{F'(x)}-\widehat{F'(x^0)}\| \leqslant \frac{1}{2}\|F'(x)-F'(x^0)\|+\frac{1}{2}\|(F'(x)-F'(x^0))^*\| = \|F'(x)-F'(x^0)\|. \tag{3.21}$$

Consideremos $x \in B[x^0, t]$, $0 \leqslant t < t_*$. Nessas condições, $f'(t) < 0$. Usando (3.21), (3.3) e levando em consideração *(H1)* e *(H2)*, obtemos

$$\begin{aligned}\|\widehat{F'(x^0)}^{-1}\|\,\|\widehat{F'(x)}-\widehat{F'(x^0)}\| &\leqslant \|\widehat{F'(x^0)}^{-1}\|\,\|F'(x)-F'(x^0)\|\\ &\leqslant f'(\|x-x^0\|)-f'(0)\\ &< f'(t)+1\\ &< 1.\end{aligned} \tag{3.22}$$

Pelo lema de Banach, concluímos que $\widehat{F'(x)}^{-1}$ existe. Além disso, que vale a desigualdade,

$$||\widehat{F'(x)}^{-1}|| \leqslant \frac{||\widehat{F'(x^0)}^{-1}||}{1 - ||\widehat{F'(x^0)}^{-1}|| \, ||F'(x) - F'(x^0)||} \leqslant \frac{||\widehat{F'(x^0)}^{-1}||}{1 - (f'(t) + 1)} \leqslant -\frac{||\widehat{F'(x^0)}^{-1}||}{f'(t)}.$$

Por outro lado, usando (3.22) temos que

$$||\widehat{F'(x)} - \widehat{F'(x^0)}|| \leqslant \frac{1}{||\widehat{F'(x^0)}^{-1}||}. \tag{3.23}$$

Consideremos $y \in \mathbb{R}^n$, segue da desigualdade anterior que

$$\langle (\widehat{F'(x^0)} - \widehat{F'(x)})y, y\rangle \leqslant ||\widehat{F'(x^0)} - \widehat{F'(x)}|| \, ||y||^2 \leqslant \frac{||y||^2}{||\widehat{F'(x^0)}^{-1}||}.$$

Que implica em

$$\langle \widehat{F'(x^0)}y, y\rangle - \frac{||y||^2}{||\widehat{F'(x^0)}^{-1}||} \leqslant \langle \widehat{F'(x)}y, y\rangle.$$

Desde que $\widehat{F'(x^0)}$ é um operador positivo e $\widehat{F'(x^0)}^{-1}$ existe, por hipótese, obtemos pelo Lema 1.5.1 que

$$\langle \widehat{F'(x^0)}y, y\rangle \geqslant \frac{||y||^2}{||\widehat{F'(x^0)}^{-1}||}.$$

Combinando as duas últimas desigualdades, concluímos que $\langle \widehat{F'(x)}y, y\rangle \geqslant 0$, isto é, $\widehat{F'(x)}$ é um operador positivo. □

A seguir apresentamos o Lema que garante a boa definição do método de Newton em uma determinada região.

Lema 3.3.3. *Desde que T seja um operador monótono maximal, se existe uma constante $c > 0$ tal que*

$$\langle F'(x^k)y, y\rangle \geqslant c||y||^2 \tag{3.24}$$

para cada $y \in \mathbb{R}^n$, então existe um único x^{k+1} tal que a primeira inclusão em (3.5) ocorre. A prova deste resultado pode ser encontrada em [32], Lema 2.2.

Observação 3.3.1. *Mostramos no Lema 3.3.2 que $\widehat{F'(x)}$ é um operador positivo e $\widehat{F'(x)}^{-1}$ existe, portanto segue diretamente do Lema 1.5.1 que para qualquer $y \in \mathbb{R}^n$*

$$\langle \widehat{F'(x)}y, y\rangle \geqslant \frac{||y||^2}{||\widehat{F'(x)}^{-1}||}.$$

Note que $\langle \widehat{F'(x)}y, y\rangle = \langle F'(x)y, y\rangle$, então pela Observação 3.3.1, nós concluímos que $F'(x)$ satisfaz (3.24) e consequentemente, a aplicação iterada de Newton é bem definida. Vamos chamar de N_{F+T}, a aplicação iterada de Newton para $F+T$ em uma determinada região, denotamos $N_{F+T} : B(x^0, t_*) \to \mathbb{R}^n$ é definido por

$$N_{F+T}(x) := L_F(x, 0)^{-1}. \tag{3.25}$$

Usando (3.2) concluímos que a definição da iterada de Newton em (3.25) é equivalente a

$$0 \in L_F(x, N_{F+T}(x)) := F(x) + F'(x)(N_{F+T}(x) - x) + T(N_{F+T}(x)), \forall\, x \in B(x^0, t_*). \tag{3.26}$$

Observação 3.3.2. *Desde que* $0 \in F(x^0) + F'(x^0)(x^1 - x^0) + T(x^1)$*,* $x \in \Omega$ *dizemos que* x^1 *é valor singular, segue de (3.2) que*

$$\{x^1\} = L_F(x^0, 0)^{-1}.$$

Portanto, temos garantia de poder aplicar a iterada de Newton uma única vez para todo ponto $x \in B(x^0, t_*)$ donde obtemos $N_{F+T}(x)$ o qual pode não pertencer a $B(x^0, t_*)$, ou mesmo nem pertencer ao domínio de F. Assim, isto é suficiente para garantir a boa definição de apenas uma iterada do método de Newton. Para assegurar que as iteradas de Newton podem ser repetidas indefinidamente, ou em particular, que sejam invariantes em subconjuntos de $B(x^0, t_*)$, necessitamos de resultados adicionais. Primeiramente, vamos definir alguns subconjuntos de $B(x^0, t_*)$ nos quais podemos aplicar o operador F. Definimos

$$K(t) := \left\{ x \in \Omega : \|x - x^0\| \leqslant t, \quad \|L_F(x, 0)^{-1} - x\| \leqslant -\frac{f(t)}{f'(t)} \right\}, \quad t \in [0, t_*), \tag{3.27}$$

$$K := \bigcup_{t \in [0, t_*)} K(t). \tag{3.28}$$

No próximo resultado mostramos que nesses subconjuntos de $B(x^0, t_*)$ temos que as iteradas de Newton estão "bem comportadas".

Lema 3.3.4. *Para cada* $0 \leqslant t < t_*$*, temos* $K(t) \subset B(x^0, t_*)$ *e* $N_{F+T}(K(t)) \subset K(n_f(t))$*. consequentemente,* $K \subseteq B(x^0, t_*)$ *e* $N_{F+T}(K) \subset K$*.*

Demonstração. A primeira inclusão segue diretamente da definição de $K(t)$. Consideremos $x \in K(t)$. Usando as definições (3.27) e (3.13) concluímos

$$\|x - x^0\| \leqslant t, \|L_F(x, 0)^{-1} - x\| \leqslant \frac{f(t)}{f'(t)}, \quad t < n_f(t) < t_*. \tag{3.29}$$

Da definição da iterada de Newton em (3.25) temos que para todo $x \in K(t)$

$$\begin{aligned} \|N_{F+T}(x) - x^0\| &\leqslant \|x - x^0\| + \|N_{F+T}(x) - x\| \\ &= \|x - x^0\| + \|L_F(x,0)^{-1} - x\|. \end{aligned}$$

Consequentemente, usando (3.13) e (3.29), segue que

$$\|N_{F+T}(x) - x^0\| \leqslant t - \frac{f(t)}{f'(t)} = n_f(t) < t_*. \tag{3.30}$$

Para simplificar notações vamos denotar $x^+ = N_{F+T}(x)$ e $y = L_F(x^+,0)^{-1}$. Assim, de (3.26) temos

$$0 \in F(x) + F'(x)(x^+ - x) + T(x^+).$$

Pelo que foi definido em (3.2)

$$0 \in F(x^+) + F'(x^+)(y - x^+) + T(y).$$

Como T é monótona maximal, segue que

$$\langle F(x) + F'(x)(x^+ - x) - F(x^+) - F'(x^+)(y - x^+), y - x^+\rangle \geqslant 0.$$

O qual implica

$$\langle F(x) - F(x^+) + F'(x)(x^+ - x), y - x^+\rangle \geqslant \langle F'(x^+)(y - x^+), y - x^+\rangle. \tag{3.31}$$

Desde que, pelo Lema 3.3.2, $\widehat{F'(x)}$ é um operador positivo e $\widehat{F'(x)^{-1}}$ existe, obtemos do Lema 1.5.1

$$\frac{\|y - x^+\|^2}{\|\widehat{F'(x^+)^{-1}}\|} \leqslant \langle \widehat{F'(x^+)}(y - x^+), y - x^+\rangle. \tag{3.32}$$

Temos $\langle \widehat{F'(x^+)}(y - x^+), y - x^+\rangle = \langle F'(x^+)(y - x^+), y - x^+\rangle$, tal igualdade combinada com (3.32) e (3.31) nos fornece

$$\begin{aligned} \|y - x^+\|^2 &\leqslant \|\widehat{F(x^+)}^{-1}\| \langle F'(x^+)(y - x^+), y - x^+\rangle \\ &\leqslant \|\widehat{F(x^+)}^{-1}\| \langle F(x) - F(x^+) + F'(x)(x^+ - x), y - x^+\rangle. \end{aligned}$$

Assim,

$$\|y - x^+\| \leqslant \|\widehat{F'(x^+)}^{-1}\| \, \|F(x) - F(x^+) + F'(x)(x^+ - x)\|. \tag{3.33}$$

Devido a $x^+ = N_{F+T}(x)$ temos de (3.30) que $\|x^+ - x^0\| \leqslant n_f(t)$. Então levando em consideração que f' é crescente e negativa, combinando (3.33), a segunda parte do Lema

3.3.2, (3.18), e o Lema 3.3.1, obtemos

$$
\begin{aligned}
\|y - x^{+}\| &\leqslant \|\widehat{F'(x^{+})}^{-1}\| \, \|F(x) - F(x^{+}) + F'(x)(x^{+} - x)\| \\
&\leqslant -\frac{\|\widehat{F'(x^{0})}^{-1}\|}{f'(n_f(t))}\|F(x^{+}) - (F(x) + F'(x))(x^{+} - x)\| \\
&= -\frac{\|\widehat{F'(x^{0})}^{-1}\|}{f'(n_f(t))}\|E_F(x, x^{+})\| \\
&\leqslant \frac{-1}{f'(n_f(t))} e_f(t, n_f(t)) \frac{\|x^{+} - x\|^2}{(n_f(t) - t)^2}.
\end{aligned} \tag{3.34}
$$

Notemos que para a última desigualdade temos de (3.29) $\|x - x^0\| \leqslant t$ e $\|x^{+} - x\| \leqslant n_f(t) - t$. Por outro lado, usando a definição em (3.13) segue

$$
f(t) + f'(t)(n_f(t) - t) = 0.
$$

Assim,

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

Como foi tomado $x^{+} = N_{F+T}(x)$, segue de (3.13) e da segunda desigualdade em (3.29) que $\|x - x^{+}\| \leqslant n_f(t) - t$, assim, podemos considerar a desigualdade em (3.34) da forma

$$
\|y - x^{+}\| \leqslant -\frac{f(n_f(t))}{f'(n_f(t))}.
$$

Portanto, desde que (3.30) implique $\|x^{+} - x^0\| \leqslant n_f(t)$ a segunda inclusão está provada. A terceira inclusão $K \subseteq B(x^0, t_*)$ segue das definições de $K(t)$ e K, em (3.27) e (3.28). Finalmente chegamos a última inclusão, $N_{F+T}(K) \subset K$. Vamos supor $x \in K$, por (3.28) segue que $x \in K(t)$ para algum $t \in [0, t_*)$. Da segunda inclusão da proposição, temos $N_{F+T}(x) \in K(n_f(t))$. Assim, desde que $n_f(t) \in [0, t_*)$ e usando a definição de K em (3.28) concluímos a demonstração. □

3.4 Resultados Principais

Para provarmos o resultado de convergência, que é uma consequência dos resultados anteriores, primeiramente notemos que a sequência definida em (3.25) implica que a sequência $\{x^k\}$ definida em (3.5), pode ser formalizada por

$$
x^{k+1} = N_{F+T}(x^k), \quad k = 0, \ 1, \ldots, \tag{3.35}
$$

como $N_{F+T}(x) = L_F(x, 0)^{-1}$, temos que (3.35) é equivalente a

$$0 \in F(x^k) + F'(x^k)(x^{k+1} - x^k) + T(x^{k+1}), \quad k = 0,\ 1, \ldots.$$

No próximo resultado vamos mostrar que a sequência gerada pelo método de Newton é "bem comportada"em relação ao conjunto $K(t)$, definido em (3.27).

Corolário 3.4.1. *A sequência* $\{x^k\}$ *está bem definida, contida em* $B(x^0, t_*)$*, converge para* $x^* \in B[x^0, t_*]$ *satisfazendo* $0 \in F(x^*) + T(x^*)$*. Além disso,* $x^k \in K(t_k)$*, para* $k = 0,\ 1, \ldots$ *e*

$$||x^* - x^k|| \leqslant t_* - t_k, \quad k = 0,\ 1, \ldots.$$

Demonstração. Pelo que foi assumido em (3.4), temos $||x^1 - x^0|| \leqslant f(0)$. De *(H1)* segue $||x^1 - x^0|| \leqslant -\dfrac{f(0)}{f'(0)}$. Segue da Observação 3.3.2 que $\{x^1\} = L_F(x^0, 0)^{-1}$. Assim, obtemos $||L_F(x^0, 0)^{-1} - x^0|| \leqslant -\dfrac{f(0)}{f'(0)}$.
Pelas definições de $K(t)$ e K, concluímos que

$$\{x^0\} = K(0) \subset K \tag{3.36}$$

Sabemos pelo Lema 3.3.4 que $N_{F+T}(K) \subset K$. Por (3.35) temos $x^{k+1} = N_{F+T}(x^k) \subset K$. Assim, inferimos que $\{x^k\}$ está bem definida e contida em K.
A primeira inclusão da segunda parte do Lema 3.3.4 nos fornece $K \subset B(x^0, t_*)$. Logo, pelos resultados anteriores, segue que $\{x^k\} \subset B(x^0, t_*)$. Para provar a convergência de $\{x^k\}$, primeiramente vamos provar por indução o resultado

$$x^k \in K(t_k), \quad k = 0,\ 1, \ldots \tag{3.37}$$

Notemos que para $k = 0$, a prova segue diretamente de (3.36). Vamos supor $x^k \in K(t_k)$. Segue da primeira inclusão na segunda parte do Lema 3.3.4 que $N_{F+T}(K(t_k)) \subset K(n_f(t_k))$. Devido a hipótese, podemos considerar $x^k \in K(t_k)$, de modo que $N_{F+T}(x^k) \in K(n_f(t_k))$. Assim, combinando (3.35) e (3.17) temos $x^{k+1} \in K(t_{k+1})$. Com isso, concluímos a demonstração da indução.
Agora, usando o resultado obtido em (3.37) e a definição do conjunto K(t) em (3.27), temos

$$||L_F(x^k, 0)^{-1} - x^k|| \leqslant -\frac{f(t_k)}{f'(t_k)}, \quad k = 0,\ 1, \ldots,$$

que combinado com (3.35) e (3.5) nos fornece

$$\|x^{k+1} - x^k\| \leqslant t_{k+1} - t_k, \quad k = 0,\ 1, \ldots \tag{3.38}$$

então

$$\begin{aligned} \|x^{k+p} - x^k\| &\leqslant \sum_{n=1}^{p} \|x^{k+n} - x^{k+(n-1)}\| \\ &\leqslant \sum_{n=1}^{p} (t_{k+n} - t_{k+(n-1)}) \\ &= t_{k+p} - t_k. \end{aligned}$$

Como a sequência $\{t_k\}$ converge para t_*, concluímos que $\{x^k\}$ é de Cauchy em $B(x^0, t_*)$ e, assim, converge para algum $x^* \in B[x^0, t_*]$. Usando novamente a inequação em (3.38), agora com a certeza da convergência de $\{x^k\}$ deduzimos que a inequação $\|x^* - x^k\| \leqslant t_* - t_k$ para $k = 0,\ 1, \ldots,$ é válida.

Para finalizar, desde que $-F(x^k) - F'(x^k)(x^{k+1} - x^k) \in T(x^{k+1})$, F é uma aplicação continuamente diferenciável, x^k converge para x^* e T é monótono maximal que implica que T tem gráfico fechado, constatamos que $0 \in F(x^*) + T(x^*)$. □

Até o momento, provamos que a sequência $\{x^k\}$ converge para a solução x^* da equação generalizada $F(x) + T(x) \ni 0$ e que $x^* \in B[x^0, t_*]$. Agora, vamos demonstrar que $\{x^k\}$ converge Q-linearmente e que x^* é a única solução de $F(x) + T(x) \ni 0$ em $B[x^0, t_*]$. Além disso, assumindo que *(H4)* seja satisfeita, vamos mostrar que $\{x^k\}$ converge Q-quadraticamente para x^*. Para tal, necessitamos do resultado a seguir.

Lema 3.4.1. *Sejam* $x,\ y \in B(x^0, R)$ *e* $0 \leqslant t < v < R$. *Se*

$$t \leqslant t_*, \quad \|x - x^0\| \leqslant t, \quad \|y - x\| \leqslant v - t, \quad f(v) \leqslant 0, \quad 0 \in F(y) + T(y), \tag{3.39}$$

então vale a desigualdade

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

Demonstração. Para simplificar notações vamos definir $z = N_{F+T}(x)$. Desde que $0 \in F(y) + T(y)$, usando (3.26) e que T é monótono maximal temos que

$$\langle -F(x) + F'(x)(x - z) + F(y), z - y \rangle \geqslant 0,$$

donde obtemos

$$\langle -F(x) - F'(x)z + F'(x)x + F(y) - F'(x)y + F'(x)y, z - y \rangle \geqslant 0,$$

deduzimos então que

$$\langle F(y) - F(x) + F'(x)(x-y), z-y\rangle \geqslant \langle F'(x)(z-y), z-y\rangle. \tag{3.40}$$

Desde que $\|x - x^0\| \leqslant t < t_*$ obtemos do Lema 3.3.2 que $\widehat{F'(x)}$ é um operador positivo e $\widehat{F'(x)}^{-1}$ existe. Assim do Lema (1.5.1) temos que

$$\frac{\|z-y\|^2}{\|\widehat{F'(x)}^{-1}\|} \leqslant \langle \widehat{F'(x)}(z-y), z-y\rangle. \tag{3.41}$$

Como $\langle \widehat{F'(x)}(z-y), z-y\rangle = \langle F'(x)(z-y), z-y\rangle$, combinando esta igualdade com (3.41) e (3.40), respectivamente, segue

$$\begin{aligned}
\|y-z\|^2 &\leqslant \|\widehat{F'(x)}^{-1}\|\langle F'(x)(z-y), z-y\rangle \\
&\leqslant \|\widehat{F'(x)}^{-1}\|\langle F(y) - F(x) + F'(x)(z-y), z-y\rangle \\
&\leqslant \|\widehat{F'(x)}^{-1}\| \, \|F(x) - [F(y) + F'(x)(x-y)]\| \, \|z-y\|.
\end{aligned}$$

Assim

$$\|y-z\| \leqslant \|\widehat{F'(x)}^{-1}\| \, \|F(x) - [F(y) + F'(x)(x-y)]\|. \tag{3.42}$$

Da definição de $E_F(x,y)$ resulta que $\|y-z\| \leqslant \|\widehat{F'(x)}^{-1}\| \, \|E_F(x,y)\|$. Usando o Lema 3.3.2 e Lema 3.3.1 (os quais têm todas as suas hipóteses satisfeitas), respectivamente, temos

$$\begin{aligned}
\|z-y\| &\leqslant -\frac{\|\widehat{F'(x^0)}^{-1}\|}{f'(t)}\|E_F(x,y)\| \\
&\leqslant \frac{-1}{f'(t)} e_f(t,v) \frac{\|y-x\|^2}{(v-t)^2}
\end{aligned}$$

Como $f'(t) < 0$ para $0 \leqslant t < t_*$. Usando a definição de $e_f(t,v)$, (3.13) e a hipótese de que $f(v) \leqslant 0$ obtemos

$$-\frac{e_f(t,v)}{f'(t)} = v - t + \frac{f(t)}{f'(t)} - \frac{f(v)}{f'(t)} \leqslant v - t + \frac{f(t)}{f'(t)} = v - n_f(t).$$

Combinando as duas últimas desigualdades demonstramos o Lema. □

Corolário 3.4.2. *As sequências* $\{x^k\}$ *e* $\{t_k\}$ *satisfazem a seguinte desigualdade*

$$\|x^* - x^{k+1}\| \leqslant \frac{t_* - t_{k+1}}{(t_* - t_k)^2}\|x^* - x^k\|^2, \quad k = 0,\ 1, \ldots. \tag{3.43}$$

Como consequência, a sequência $\{x^k\}$ *converge Q-linearmente para a solução* x^**, como segue*

$$\|x^* - x^{k+1}\| \leqslant \frac{1}{2}\|x^* - x^k\|, \quad k = 0,\ 1, \ldots. \tag{3.44}$$

Se adicionarmos a hipótese (H4) então a sequência $\{x^k\}$ *converge Q-quadraticamente para* x^**, como segue*

$$\|x^* - x^{k+1}\| \leqslant \frac{f''_-(t_*)}{-2f'(t_*)}\|x^* - x^k\|^2, \quad k = 0,\ 1, \dots. \tag{3.45}$$

Demonstração. Para cada k, podemos aplicar o Lema 3.4.1 com $x = x^k,\ y = x^*,\ t = t_k$ e $v = t_*$, para obtermos

$$\|x^* - N_{F+T}(x^k)\| \leqslant [t_* - n_f(t_k)]\frac{\|x^* - x^k\|^2}{(t_* - t_k)^2}.$$

Assim, a desigualdade (3.43) segue de (3.35) e (3.17). Temos da primeira parte do Lema 3.3.2 e de (3.17) que

$$\frac{t_* - t_{k+1}}{t_* - t_k} \leqslant \frac{1}{2},$$

do Corolário 3.4.1, segue

$$\frac{\|x^* - x^k\|}{t_* - t_k} \leqslant 1.$$

Combinando estas duas desigualdades com (3.43), obtemos a desigualdade em (3.44).

$$\begin{aligned} \|x^* - x^{k+1}\| &\leqslant \frac{t_* - t_{k+1}}{t_* - t_k}\frac{\|x^* - x^k\|}{t_* - t_k}\|x^* - x^k\| \\ &\leqslant \frac{1}{2}\|x^* - x^k\|. \end{aligned}$$

Agora, vamos assumir que *(H4)* seja satisfeita, pelo Corolário 3.3.1, a segunda desigualdade em (3.8) ocorre e (3.43) implica em (3.45).

$$\begin{aligned} \|x^* - x^{k+1}\| &\leqslant \frac{t_* - t_{k+1}}{(t_* - t_k)^2}\|x^* - x^k\|^2 \\ &\leqslant \frac{f''_-(t_*)}{-2f'(t_*)}\|x^* - x^k\|^2. \end{aligned}$$

Com isto, concluímos a demonstração. □

Corolário 3.4.3. *O limite* x^* *da sequência* $\{x^k\}$ *é a única solução da equação generalizada* $F(x) + T(x) \ni 0$ *em* $B[x^0, t_*]$.

Demonstração. Vamos supor que existe $y^* \in B[x^0, t_*]$ tal que y^* é solução da equação generalizada $F(x) + T(x) \ni 0$. Provaremos por indução que

$$\|y^* - x^k\| \leqslant t_* - t_k, \quad k = 0,\ 1, \dots. \tag{3.46}$$

O caso $k = 0$ segue imediato, pois $t_0 = 0$ e $y^* \in B[x^0, t_*]$, assim, $\|y^* - x^0\| \leqslant t_*$. Vamos assumir que a desigualdade seja válida para algum k. Temos do Corolário 3.4.1 que

$x^k \in K(t_k)$, para $k = 0,\ 1, \ldots$, pela definição de $K(t_k)$ concluímos que $\|x^k - x^0\| \leqslant t_k$, para $k = 0,\ 1, \ldots$. Desde que $\|x^k - x^0\| \leqslant t_k$, podemos aplicar o Lema 3.4.1 com $x = x^k,\ y = y^*,\ t = t_k$ e $v = t_*$ para obtermos

$$\|y^* - N_{F+T}(x^k)\| \leqslant [t_* - n_f(t_k)] \frac{\|y^* - x^k\|^2}{(t_* - t_k)^2}.$$

Usando a hipótese de indução, (3.35) e (3.17) segue que

$$\|y^* - x^{k+1}\| \leqslant t_* - t_{k+1},$$

assim, demonstramos a indução.

Desde que as sequências $\{x^k\}$ e $\{t_k\}$ convirjam, respectivamente, para x^* e t_*, concluímos de (3.46) que $y^* = x^*$. Portanto, x^* é a única solução de $F(x) + T(x) \ni 0$ em $B[x^*, t_*]$. $\square$

Capítulo 4

Casos Especiais

Neste capítulo, elaborado a partir das referências [8] e [30], apresentamos alguns casos especiais do Teorema 3.2.1. Quando $\mathsf{F} \equiv \{0\}$ e f' satisfaz a condição de Lipschitz obtemos um caso particular do Teorema 3.2.1, baseado no método de Newton sob a condição Lipschitz. Uma versão do Teorema de Smale para o método de Newton para funções analíticas é provado no Teorema 4.2.1.

4.1 Sob a Condição Lipschitz

Apresentamos, agora, uma versão do método de Newton sob a condição Lipschitz para equações generalizadas.

Teorema 4.1.1. *Sejam o aberto, não vazio,* $\Omega \subset \mathbb{R}^n$*,* $\mathsf{F} : \Omega \xrightarrow{C^1} \mathbb{R}^n$*,* $\mathsf{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *monótono maximal e* $x^0 \in \Omega$*. Consideremos* $\mathsf{F}'(x^0)$ *positivo tal que* $\widehat{\mathsf{F}'(x^0)}^{-1}$ *exista. Suponhamos que há uma constante* $\mathsf{K} > 0$ *tal que* $\mathsf{B}(x^0, 1/\mathsf{K}) \subset \Omega$ *e que*

$$\|\widehat{\mathsf{F}'(x^0)}^{-1}\| \, \|\mathsf{F}'(x) - \mathsf{F}'(y)\| \leqslant \mathsf{K}\|x - y\|, \quad x,\ y \in \mathsf{B}(x^0, 1/\mathsf{K}). \tag{4.1}$$

Além disso, vamos considerar uma constante $\mathsf{b}\mathsf{K} < 1/2$*, tal que*

$$\|x^1 - x^0\| \leqslant \mathsf{b}. \tag{4.2}$$

Então a sequência $\{x^k\}$ *gerada pelo método de Newton para resolver* $\mathsf{F}(x) + \mathsf{T}(x) \ni 0$*, que tem* x^0 *como ponto inicial*

$$\mathsf{F}(x^k) + \mathsf{F}'(x^k)(x^{k+1} - x^k) + \mathsf{T}(x^{k+1}) \ni 0, \quad k = 0,\ 1, \ldots, \tag{4.3}$$

está bem definida, está contida em $B(x^0, t_*)$ *e converge para o ponto* x^* *que é a única solução de* $F(x) + T(x) \ni 0$ *em* $B[x^0, t_*]$*, onde* $t_* = (1 - \sqrt{1 - 2bK})/K$*. Além disso, a sequência* $\{x^k\}$ *satisfaz para qualquer* $k = 0,\ 1, \ldots$

$$\|x^* - x^{k+1}\| \leqslant \frac{K}{2\sqrt{1 - 2bK}} \|x^* - x^k\|^2.$$

Demonstração. Desde que $f : [0, 1/K) \to \mathbb{R}$, definida por $f(t) := (K/2)t^2 - t + b$ é uma função majorante para F no ponto x^0.

Mostremos que as hipóteses acerca do Teorema 3.2.1 são satisfeitas.

(H1) Temos $f(0) = b > 0$ e $f'(0) = -1$.

(H2) Claramente f' é convexa. Além disso, considerando $a,\ c \in [0, 1/K)$ e supondo, sem perda de generalidade, que $a < c$, temos $f'(a) = Ka - 1$ e $f'(c) = Kc - 1$, assim, $f'(a) < f'(c)$. Concluímos que f' é estritamente crescente.

(H3) Provaremos que $f(t) = 0$, para algum $t \in (0, 1/K)$. Como $f(t) = (K/2)t^2 - t + b$ tem raízes

$$\frac{1 - \sqrt{1 - 2Kb}}{K} \quad e \quad \frac{1 + \sqrt{1 - 2Kb}}{K},$$

onde apenas a primeira raíz, denotamos por t_*, está em $(0, 1/K)$.

Assim, a primeira parte do Teorema 4.1.1 segue invocando o Teorema 3.2.1, aplicado a este contexto particular. Para provarmos a convergência quadrática mostraremos que a hipótese adicional do Teorema 3.2.1 é satisfeita.

(H4) Como $b < 1/(2K)$, temos que

$$f'(t_*) = Kt_* - 1 = K\left(\frac{1 - \sqrt{1 - 2bK}}{K}\right) - 1 = -\sqrt{1 - 2bK} < 0.$$

Segue do Teorema 3.2.1

$$\begin{aligned} \|x^* - x^{k+1}\| &\leqslant \frac{f''_-(t_*)}{-2f'(t_*)} \|x^* - x^k\|^2 \\ &= \frac{K}{2\sqrt{1 - 2bK}} \|x^* - x^k\|^2. \end{aligned}$$

O que encerra a demonstração. $\square$

4.2 Sob a Condição de Smale

Mostramos uma versão da clássica convergência do método de Newton sob a condição de Smale para equações generalizadas.

Teorema 4.2.1. *Sejam* $\Omega \subset \mathbb{R}^n$, *aberto e não vazio,* $F : \Omega \to \mathbb{R}^n$ *uma função analítica,* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *monótono maximal e* $x^0 \in \Omega$. *Vamos supor* $F'(x^0)$ *positivo e que* $\widehat{F'(x^0)}^{-1}$ *exista. Consideremos:*

$$\gamma := \|\widehat{F'(x^0)}^{-1}\| \sup_{n>1} \left\| \frac{F^{(n)}(x^0)}{n!} \right\|^{\frac{1}{n-1}} < +\infty. \tag{4.4}$$

Além disso, suponhamos $B(x^0, 1/\gamma) \subset \Omega$ *e que exista* $b > 0$ *tal que*

$$\|x^1 - x^0\| \leqslant b \tag{4.5}$$

e $\alpha := b\gamma \leqslant 3-2\sqrt{2}$. *Então a sequência* $\{x^k\}$ *gerada pelo método de Newton para resolver* $F(x) + T(x) \ni 0$, *iniciando em* x^0

$$F(x^k) + F'(x^k)(x^{k+1} - x^k) + T(x^{k+1}) \ni 0, \quad k = 0,\ 1, \dots, \tag{4.6}$$

está bem definida, está contida em $B(x^0, t_*)$ *e converge para* x^* *que é a única solução de* $F(x) + T(x) \ni 0$ *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*

$$\|x^* - x^{k+1}\| \leqslant \frac{1}{2}\|x^* - x^k\|, \quad k = 0,\ 1, \dots.$$

Adicionalmente, se $\alpha < 3 - 2\sqrt{2}$, *então* $\{x^k\}$ *converge Q-quadraticamente.*

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

Antes de demonstrar o Teorema 4.2.1 precisamos de dois resultados importantes.

Lema 4.2.1. *Consideremos o aberto* $\Omega \subset \mathbb{R}^n$ *e seja* $F : \Omega \to \mathbb{R}^n$ *uma função analítica. Suponhamos que* x^0 *e* $B(x^0, 1/\gamma) \subset \Omega$, *onde* γ *foi definido em (4.4). Então para todo* $x \in B(x^0, 1/\gamma)$, *temos*

$$\|\widehat{F'(x^0)}^{-1}\|\ \|F''(x)\| \leqslant \frac{2\gamma}{(1 - \gamma\|x - x^0\|)^3}.$$

Demonstração. Seja $x \in \Omega$, desde que F é uma função analítica, podemos escrever $F''(x) = \sum_{n=0}^{\infty} \frac{1}{n!} F^{(n+2)}(x^0)(x - x^0)^n$. Assim,

$$\begin{aligned} \|F''(x)\| &= \left\| \sum_{n=0}^{\infty} \frac{F^{(n+2)}(x^0)}{n!}(x - x^0)^n \right\| \\ &\leqslant \sum_{n=0}^{\infty} (n+2)(n+1) \left\| \frac{F^{(n+2)}(x^0)}{(n+2)!} \right\| \|x - x^0\|^n \end{aligned} \tag{4.7}$$

Definimos anteriormente $\gamma = \|\widehat{F'(x^0)}^{-1}\| \sup_{n>1} \left\| \frac{F^{(n)}(x^0)}{n!} \right\|^{\frac{1}{n-1}}$. Então para $n+2$ segue que

$$\gamma = \|\widehat{F'(x^0)}^{-1}\| \sup_{n>1} \left\| \frac{F^{(n+2)}(x^0)}{(n+2)!} \right\|^{\frac{1}{n+1}} \geqslant \|\widehat{F'(x^0)}^{-1}\| \left\| \frac{F^{(n)}(x^0)}{n!} \right\|^{\frac{1}{n-1}}.$$

Logo,

$$\left(\frac{\gamma}{\|\widehat{F'(x^0)}^{-1}\|} \right)^{n+1} \geqslant \left\| \frac{F^{(n+2)}(x^0)}{(n+2)!} \right\|$$

Combinando a desigualdade anterior com (4.7)

$$\begin{aligned} \|F''(x)\| &\leqslant \sum_{n=0}^{\infty} (n+2)(n+1) \left(\frac{\gamma}{\|\widehat{F'(x^0)}^{-1}\|} \right)^{n+1} \|x - x^0\|^n \\ &= \gamma \sum_{n=0}^{\infty} (n+2)(n+1) \frac{1}{\|\widehat{F'(x^0)}^{-1}\|} (\gamma \|x - x^0\|)^n. \end{aligned}$$

Como $B(x^0, 1/\gamma) \subset \Omega$, segue que $\gamma\|x - x^0\| < 1$. Podemos, assim, usar a Proposição 1.6.1 para obtermos

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

Concluímos que

$$\|\widehat{F'(x^0)}^{-1}\| \, \|F''(x)\| \leqslant \frac{2\gamma}{(1 - \gamma\|x - x^0\|)^3}.$$

$\square$

Lema 4.2.2. *Sejam* $\Omega \subset \mathbb{R}^n$ *um conjunto aberto e* $F : \Omega \xrightarrow{C^2} \mathbb{R}^n$. *Seja* $x^0 \in \Omega$, $R > 0$ *e suponhamos* $\kappa = \sup\{t \in [0, R) : B(x^0, t) \subset \Omega\}$. *Seja* $f : [0, R) \xrightarrow{C^2} \mathbb{R}$ *tal que* $\|\widehat{F'(x^0)}^{-1}\| \, \|F''(x)\| \leqslant f''(\|x - x^0\|)$, *para todo* $x \in B(x^0, \kappa)$, *então* F *e* f *satisfazem (3.3).*

Demonstração. Tome $\lambda \in [0,1]$ e $x,\ y \in \Omega$, $\|x - x^0\| < R$, então:

$$\begin{aligned}
\|\widehat{F'(x^0)}^{-1}\| \, \|F'(y) - F'(x)\| &\leqslant \int_0^1 \|\widehat{F'(x^0)}^{-1} F''(x + \lambda(y-x))\| \, \|y - x\| d\lambda \\
&\leqslant \int_0^1 f''(\|x + \lambda(y-x) - x^0\|)\|y - x\| d\lambda \\
&= f'(\|x + \lambda(x-y) - x^0\|)\Big|_0^1 \\
&= f'(\|y - x^0\|) - f'(\|x - x^0\|) \\
&\leqslant f'(\|y - x\| + \|x - x^0\|) - f'(\|x - x^0\|).
\end{aligned}$$

Concluímos, assim, a demonstração do Lema. □

Com os resultados anteriores estamos aptos a demonstrar o Teorema 4.2.1.

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

$$f(t) := \frac{t}{1-\gamma t} - 2t + b$$

Combinando os Lemas 4.2.1 e 4.2.2, concluímos que F e f satisfazem (3.3). Mostraremos que as hipóteses do Teorema 3.2.1 são satisfeitas para este caso.

(H1) Temos que $f(0) = b > 0$ e que $f'(0) = -1$.

(H2) Notemos que f' é convexa, pois $(f'(t))'' = (6\gamma^2)/(1-\gamma t)^4 \geqslant 0$. Vamos supor $a,\ c \in (0, 1/\gamma]$, sem perda de generalidade vamos supor que $a < c$, como $\gamma > 0$ segue que

$$\gamma a < \gamma c \Rightarrow 1 - \gamma a > 1 - \gamma c \Rightarrow 1/[(1-\gamma a)^2 - 2] < 1/[(1-\gamma c)^2 - 2] \Rightarrow f'(a) < f'(c).$$

Concluímos que f' é estritamente crescente.

(H3) $f(t_*) = 0$ para $t_* = (\alpha + 1 - \sqrt{((1+\alpha)^2 - 8\alpha)})/(4\gamma)$.

(H4) Agora, considerando $\alpha < 3 - 2\sqrt{2}$ temos que as raízes de f diferem e, portanto, $f'(t_*) < 0$, como $f'(t) = 1/(1-\gamma t)^2 - 2$, e $f''(t) = (2\gamma)/(1-\gamma t)^3$, pelo Teorema 3.2.1, segue que

$$\begin{aligned}
\|x^* - x^{k+1}\| &\leqslant \frac{f''_-(t_*)}{-2f'(t_*)}\|x^* - x^k\|^2 \\
&= \frac{(2\gamma)/(1-\gamma t_*)^3}{-2(1/(1-\gamma t_*)^2 - 2)}\|x^* - x^k\|^2 \\
&= \frac{\gamma}{(1-\gamma t_*)[2(1-\gamma t_*)^2 - 1]}\|x^* - x^k\|^2.
\end{aligned}$$

□

Capítulo 5

Considerações Finais

Neste trabalho, apresentamos um método para resolver a equação generalizada $F(x) + T(x) \ni 0$, estendendo os resultados já obtidos para resolver equações não lineares. Baseados na linearização parcial do operador $F + T$, obtemos um resultado de convergência semilocal para o método de Newton, estabelecemos o raio ótimo, unicidade e taxas de convergência. Empregamos a abordagem do princípio majorante, o qual unifica vários resultados de convergência relacionados ao método de Newton. Ao final, consideramos também a análise sob a condição de Lipschitz e sob a condição de Smale.

Uma possibilidade de trabalho futuro é aplicar o método Quase-Newton para resolver a equação generalizada, além de tentar relaxar as hipóteses impostas aos operadores F, T e a função majorante f.

Referências Bibliográficas

[1] AUSLENDER, A.; TEBOULLE, M. *Asymptotic cones and functions in optimization and variational inequalities.* New York: Springer-Verlang, 2003.

[2] CHANG, D. C.; WANG, J.; YAO J. C. *Newton's method for variational inequality problems: smales point estimate theory under the γ-condition.* Applicable Analysis. **94** (1) (2015) 44-55.

[3] DONTCHEV, A. L. *Local analysis of a Newton-type method based on partial linearization.* In The Mathematics of Numerical Analysis, vol. 32, Lectures in Appl. Math., Providence RI, 1996, p.295-306.

[4] DONTCHEV, A. L.; ROCKAFELLAR, R. T. *Implicit functions and solution mappings.* Springer Monographs in Mathematics, Springer, Dordrecht, 2009. A view from variational analysis.

[5] FACCHINEI, F; PANG, J. S. *Finite-dimensional variational inequalities and complementarity problems.* New York: Springer-Verlang. 2003.

[6] FERREIRA, O. P. *A robust semi-local convergence analysis of Newton's method for cone inclusion problems in Banach spaces under invariant majorant condition.* J. Comput. Appl. Math. **279** (0) (2015) 318-335.

[7] FERREIRA, O. P. *Local convergence of Newton's method under majorant condition.* Journal of Computational and Applied Mathematics. **235** (1) (2011) 1515-1522.

[8] FERREIRA, O. P.; GONÇALVES, M. L. N.; OLIVEIRA, P. R. *Local convergence analysis of the Gauss-Newton method under a majorant condition.* J. Complexity. **27** (1) (2011) 111-125.

[9] FERREIRA, O. P.; SILVA, G. N. (2018): *Inexact Newton's Method to Nolinear Functions with values in a cone*, Applicable Analysis, DOI: 10.1080/00036811.2018.1430779.

[10] FERREIRA, O. P.; SVAITER, B. F. *Kantorovich's Theorem on Newton's Method.* arXiv preprint arXiv:1209.5704, 2012.

[11] FERREIRA, O. P.; SVAITER, B. F. *Kantorovich's majorants principle for Newton's method.* Computational Optimization and Applications. **42** (2) (2009) 213-229.

[12] IZMAILOV, A.; SOLODOV, M. *Otimização.* 3.ed. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2014, 1v.

[13] JOSEPHY, M. *Newton's method for generalized equations and the PIES energy model.* University of Wisconsin-Madison, 1979.

[14] KANTOROVICH, L. V. *On Newton's method for functional equations.* Dokl. Akad. Nauk SSSR. **59** (7) (1948) 1237-1240.

[15] LIMA, E. L. *Álgebra linear.* 8.ed. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2012. (Coleção Matemática Universitária).

[16] LIMA, E. L. *Análise real.* 11.ed. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2012, 1v. (Coleção Matemática Universitária).

[17] LIMA, E. L. *Curso de análise.* 14.ed. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2012, 1v. (Projeto Euclides).

[18] LIMA, E. L. *Análise real.* 6.ed. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2013, 2v. (Coleção Matemática Universitária).

[19] LIMA, E. L. *Curso de Análise.* 11.ed. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada, 2015, 2v. (Projeto Euclides).

[20] MENDES, R. M. N. *Álgebra Linear.* PUC Minas, Belo Horizonte, 2013, ISBN: 978-85-915683-0-7.

[21] MINTY, G. J. *On the monotonicity of the gradient of a convex function.* Pacific J. Math. **14** (1) (1964) 243-247.

[22] ORTEGA, J. M. *The Newton-Kantorovich theorem.* MAA The American Mathematical Monthly. **75** (6) (1968) 658-660.

[23] ORTEGA, J. M.; RHEINBOLDT, W. C. *Iterative solution of nonlinear equations in several variables.* New York-London: Academic Press, 1970.

[24] PIETRUS, A.; JEAN-ALEXIS, C. *Newton-secant method for functions with values in a cone.* Serdica Math. **39** (3-4) (2013) 271-286.

[25] POLYAK, B. T. *Newton's method and its use in optimization.* European Journal of Operational Research. **181** (2007) 1086-1096.

[26] POTRA, F. A. *The Kantorovich theorem and interior point methods.* Math. Program. **102** (1) (2005) 47-70.

[27] ROBINSON, S. M. *Extension of Newton's method to nonlinear functions with values in a cone.* Numer. Math. **19** (1972) 341-347.

[28] ROCKAFELLAR, R. T. *Convex analysis.* Princeton: Princeton University Press, 1970.

[29] ROCKAFELLAR, R. T. *Monotone operators and the proximal point algorithm.* SIAM Jornual on Control and Optimization. **14** (5) (1976) 877-898.

[30] SILVA, G. N. *Kantorovich's theorem on Newton's method for solving generalized equations under the majorant condition.* Journal Applied Mathematics and Computation. **286** (2016) 178-188.

[31] TIEL, J. V. *Convex analysis an introductory text.* Royal Netherlands Meteorological Institute, 1984.

[32] UKO, L. U. *Generalized equations and the generalized Newton method.* Mathematical Programming. **73** (3) (1996) 251-268.

[33] UKO, L. U.; ARGYROS, L. K. *Generalized equations, variational inequalities and a weak Kantorovich theorem.* Numer. Algorithms. **52** (3) (2009) 321-333.

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

[35] ZABREJKO, P. P.; NGUEN, D. F. *The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates.* Numerical Functional Analysis Optimmization. **9** (5-6) (1987) 671-684.