Instituto Alberto Luiz Coimbra de
Pós-Graduação e Pesquisa de Engenharia
COPPE UFRJ

ANÁLISE DE CONVERGÊNCIA DOS MÉTODOS DE GAUSS-NEWTON DO PONTO DE VISTA DO PRINCÍPIO MAJORANTE

Max Leandro Nobre Gonçalves

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientadores: Paulo Roberto Oliveira
Orizon Pereira Ferreira

Rio de Janeiro
Junho de 2011

ANÁLISE DE CONVERGÊNCIA DOS MÉTODOS DE GAUSS-NEWTON DO PONTO DE VISTA DO PRINCÍPIO MAJORANTE

Max Leandro Nobre Gonçalves

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Paulo Roberto Oliveira, D.Ing.

Prof. Orizon Pereira Ferreira, D.Sc.

Prof. Gregório Malajovich Muñoz, D.Sc.

Prof. João Xavier da Cruz Neto, D.Sc.

Prof. Susana Scheimberg de Makler, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
JUNHO DE 2011

Gonçalves, Max Leandro Nobre

Análise de Convergência dos Métodos de Gauss-Newton do Ponto de Vista do Princípio Majorante /Max Leandro Nobre Gonçalves. – Rio de Janeiro: UFRJ/COPPE, 2011.

IX, 77 p. 29, 7cm.

Orientadores: Paulo Roberto Oliveira

Orizon Pereira Ferreira

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2011.

Referências Bibliográficas: p. 74 – 77.

1. condição majorante. 2. problemas de mínimos quadrado não-lineares. 3. problemas de otimização de composição convexa. 4. método de Gauss-Newton. 5. método quase Gauss-Newton inexato. 6. algoritmo de Gauss-Newton. 7. convergence local e semi-local. I. Oliveira, Paulo Roberto *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

À minha família pelo carinho e amparo ao longo desses anos.

Agradecimentos

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

Aos Profs. Paulo Roberto Oliveira e Orizon Pereira Ferreira, pela amizade, paciência e dedicação que foram indispensáveis para a concretização deste trabalho.

À minha esposa Ana Paula, pela amizade, compreensão e pelas constantes ajudas.

À todos os meus familiares e amigos, que apoiaram-me em mais um degrau de minha vida. Em especial, aos colegas de doutorado, que ajudaram-me nos momentos de dificuldades.

À (CAPES), pela Bolsa de Estudos Concedida, sem a qual seria difícil a realização desta dissertação

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

ANÁLISE DE CONVERGÊNCIA DOS MÉTODOS DE GAUSS-NEWTON DO PONTO DE VISTA DO PRINCÍPIO MAJORANTE

Max Leandro Nobre Gonçalves

Junho/2011

Orientadores: Paulo Roberto Oliveira
Orizon Pereira Ferreira

Programa: Engenharia de Sistemas e Computação

A busca de soluções dos problemas de mínimos quadrados não-lineares e de otimização de composição convexa é objeto de interesse em várias áreas da ciência e das engenharias. Devido a sua velocidade de convergência e eficiência computacional, os métodos de Gauss-Newton têm sido bastante utilizados para o propósito de obter estas soluções. Neste trabalho apresentamos análises de convergência dos métodos de Gauss-Newton, usando o princípio majorante. Em cada caso, nossa análise deixa clara a relação entre a função majorante e a função não-linear associada ao problema, o que torna as condições e demonstrações de convergência mais simples.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

LOCAL CONVERGENCE ANALYSIS OF GAUSS-NEWTON'S METHOD UNDER MAJORANT CONDITION

Max Leandro Nobre Gonçalves

June/2011

Advisors: Paulo Roberto Oliveira
Orizon Pereira Ferreira

Department: Systems Engineering and Computer Science

The search for solutions of the problems of nonlinear least squares and convex composite optimization, is object of interest in some areas of science and engineering. Due the speed of convergence and computational efficiency, the Gauss-Newton methods have been sufficiently use to obtain these solutions. In this work we present convergence analysis of the Gauss-Newton methods, using the majorant principle. In each case, our analysis makes clear the relationship between the majorant function and nonlinear function associated with the problem, which became the conditions and proof of convergence easier.

Sumário

Introdução

O método de Gauss-Newton e suas variações são os mais eficientes métodos conhecidos para resolver problemas de mínimos quadrados não-lineares e de otimização de composição convexa. A convergência dos métodos de Gauss-Newton pode falhar, ou mesmo deixar de gerar uma sequência infinita. Para assegurar a convergência dos métodos para a solução dos respectivos problemas, algumas condições devem ser impostas. Neste trabalho, determinaremos condições que garantam a convergência dos métodos de Gauss-Newton para as soluções desejadas.

Sejam X e Y espaços reais ou complexos de Hilbert , $\Omega \subset X$ um conjunto aberto e $F : \Omega \to Y$ Fréchet diferenciável. Considere o *problema de mínimos quadrados não-linear*

$$\min \ \|F(x)\|^2. \tag{1}$$

Este problema tem sido uma frutífera área de estudo nos últimos 30 anos, principalmente pelo seu grande número de aplicações em problemas práticos, ver [1]. Essas aplicações, buscam encontrar os parâmetros de um modelo matemático, que melhor descreva um conjunto de dados numéricos de um experimento químico, físico, estatístico ou econômico, usando uma função da forma (1) para medir a discrepância entre as saídas do modelo e o conjunto de dados.

Se $F'(x)$ é injetiva e tem imagem fechada para todo $x \in \Omega$, o método de Gauss-Newton encontra pontos estacionários para o problema (1), i.e., soluções para o sistema de equações não-lineares

$$F'(x)^* F(x) = 0,$$

onde A^* denota a matriz adjunta do operador A. Formalmente o *método de Gauss-Newton* é descrito como: Dado $x_0 \in \Omega$ defina

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

É importante ressaltar que se x_* é a solução de (1), $F(x_*) = 0$ e $F'(x_*)$ é invertível os resultados para o método de Gauss-Newton se reduzem aos resultados para o método de Newton. Por outro lado, note que cada passo do método de

Gauss-Newton consiste em resolver um sistema de equações lineares, o que pode ser computacionalmente muito "caro"se o número de incógnitas for muito grande. Uma alternativa é resolver aproxidamente o sistema linear, o que dá origem a uma nova classe de processos iterativos. Estes processos são conhecidos como métodos inexatos de Gauss-Newton. Os *métodos inexatos de Gauss-Newton* usados para resolver (1) são formalmente descritos como: Dado $x_0 \in \Omega$ defina

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

onde $B(x_k) : X \to Y$ é um operador linear e $r_k \in Y$ é um resíduo apropriado. Em particular, o processo acima é o *método de Gauss-Newton inexato* se $B_k = F'(x_k)^*F'(x_k)$, é o *método de Gauss-Newton modificado inexato* se $B_k = F'(x_0)^*F'(x_0)$, e é o *método de quase Gauss-Newton inexato* se B_k é uma aproximação para $F'(x_k)^*F'(x_k)$.

O *método quase Gauss-Newton inexato*, onde controle residual relativo escalado é feito em cada iteração, é formalmente descrito como: Dado $x_0 \in \Omega$ defina

$$x_{k+1} = x_k + S_k, \qquad B(x_k)S_k = -F'(x_k)^*F(x_k) + r_k, \qquad k = 0, 1, \dots, \tag{2}$$

onde $B(x_k)$ é uma aproximação invertível para $F'(x_k)^*F'(x_k)$ e o resíduo r_k satisfaz

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^*F(x_k)\|,$$

para as dadas sequências $\{\theta_k\}$ e $\{P_k\}$ de números reais e matrizes invertíveis, respectivamente. A sequência $\{\theta_k\}$ é usualmente chamada de sequência "forcing"e é tomada uniformemente limitada por 1. Além disso, a vantagem de se introduzir a matriz P_k (considerado pela primeira vez em [2]) é que, se o sistema linear em (2) for mal condicionado, esta matriz pode fazer uma correção.

Outro problema que consideraremos em nosso trabalho, é o *problema de otimização de composição convexa*

$$\min\ h(F(x)), \tag{3}$$

onde $h : \mathbb{R}^m \to \mathbb{R}$ é uma função convexa de valores reais e $F : \mathbb{R}^n \to \mathbb{R}^m$ é continuamente diferenciável. É bem conhecido, ver [3–5] e as suas referências, que uma ampla variedade de aplicações para esta formulações pode ser encontrada dentro da programação matemática, especificamente, em inclusão convexa, problemas de minimax e métodos de penalização. Além de suas aplicações práticas, este modelo também provê ferramentas poderosas para estudar condições de otimalidade de primeira e segunda ordem para otimização restrita.

Note que o estudo de (3) está relacionado com o problema de inclusão convexa

$$F(x) \in C := \{z \in \mathbb{R}^m : h(z) \leq h(y),\ \forall y \in \mathbb{R}^m\}, \tag{4}$$

pois se $x_* \in \mathbb{R}^n$ satisfaz a inclusão convexa (4) então x_* é uma solução de (3), mas se $x_* \in \mathbb{R}^n$ é uma solução de (3) não necessariamente satisfaz a inclusão convexa (4).

Para cada $\Delta \in (0, +\infty)$ e $x \in \mathbb{R}^n$ seja

$$D_\Delta(x) := \operatorname{argmin}\{h(F(x) + F'(x)d) : \ d \in \mathbb{R}^n,\ \|d\| \leq \Delta\}. \tag{5}$$

Assim, dados $\Delta \in (0, +\infty]$, $\eta \in [1, +\infty)$ e um ponto $x_0 \in \mathbb{R}^n$, *o algoritmo tipo Gauss-Newton* associado a (Δ, η, x_0) como definido em [3] (ver também, [4, 5]) é como segue:

Algoritmo 0.1

Inicialização. *Dados* $\Delta \in (0, +\infty]$, $\eta \in [1, +\infty)$ *e* $x_0 \in \mathbb{R}^n$. *Faça* $k = 0$.
critério de Parada. *Calcule* $D_\Delta(x_k)$. *Se* $0 \in D_\Delta(x_k)$, *Pare. Caso contrário.*
Passo Iterativo . *Calcule* d_k *satisfazendo*

$$d_k \in D_\Delta(x_k), \qquad \|d_k\| \leq \eta d(0, D_\Delta(x_k)),$$

e faça

$$x_{k+1} = x_k + d_k,$$

$k = k + 1$ *e vá para* Critério de Parada.

A análise de convergência local clássica para os métodos de Gauss-Newton requer, entre outras hipóteses, que F' satisfaça a condição Lipschitz (ver, por exemplo [3, 5–7]). Nos últimos anos, tem surgido um grande número de artigos tratando da questão da convergência dos métodos de Newton, incluindo os métodos de Gauss-Newton por relaxar a hipótese de continuidade Lipschitz para a derivada de F (ver, por exemplo, [4, 8–24]). Estas novas análises, além de provar a convergência dos métodos, têm permitido unificar resultados até então sem relação e em alguns casos estimar os raios de convergência ótima e unicidade de solução.

Usando o princípio majorante introduzido por Kantorovich em [8] e usado por Ferreira em [15] e [23], Ferreira, Gonçalves em [16] e Ferreira, Svaiter em [17], o presente trabalho tem por objetivos apresentar:

1. uma nova análise de convergência local do método de Gauss-Newton;

2. uma nova análise de convergência local do método quase Gauss-Newton inexato, onde controle residual relativo escalado é feito em cada iteração;

3. uma nova análise de convergência semi-local da sequência gerada pelo algoritmo de Gauss-Newton.

Em nossa análise, a condição Lipschitz para F' é relaxada usando uma função majorante. Nosso tratamento tem a vantagem de deixar clara a relação entre a função majorante e a função F associada aos problemas em consideração, ver por exemplo os Lemas 2.1, 2.2, 2.3 e 2.4, o que não está explícito em outras análises de convergência dos métodos que seguem a mesma idéia. Com isso, os resultados apresentados aqui, tornaram as condições e demonstrações de convergência mais simples e mais didáticas. Além disso, nossa condição majorante permite unificar resultados previamente não relacionados pertencentes aos métodos.

Outras vantagens da nossa análise são que no caso de convergência local dos métodos de Gauss-Newton o contexto foi estendido para espaços de Hilbert e no caso quase Gauss-Newton inexato adicionamos resultados de convergência para a condição de Smale. Ademais, na prova de convergência semi-local da sequência gerada pelo algoritmo de Gauss-Newton utilizamos uma técnica que no lugar de olhar apenas a sequência gerada, identificamos regiões onde a sequência de Gauss-Newton para o problema de otimização convexa é bem definido quando comparado com método de Newton aplicado a uma função auxiliar associada a função majorante, esta técnica foi introduzida em [25].

A análise de convergência para os métodos de Gauss-Newton também foram estudados em [4, 10, 11, 13, 18, 19, 24]. A principal diferença destas análises com as nossas análises é que em lugar de nossa condição majorante é usada a condição de Wang, introduzida em [22]. De fato, pode ser mostrado que elas são equivalentes. Porém, a formulação como uma condição majorante é mais natural e como já dito deixa as condições e provas de convergência mais simples.

Gostaríamos de destacar também, que nossa análise de convergência local para o método de Gauss-Newton sob condição majorante, já se encontra publicado em periódico especializado, ver [26].

Esta dissertação está organizada da seguinte forma. No capítulo 1 damos as notações e alguns resultados preliminares sobre operadores lineares, derivadas, funções convexas e analíticas e regularidade. No capítulo 2 dedicamos a um breve estudo das funções majorantes, provando as principais relações com a função F. No capítulo 3 concentra-se a discussão sobre a convergência local do método de Gauss-Newton. Mostraremos que sob certas condições, a sequência gerada pelo método está bem definida e converge para uma solução de (1). Além disso, determinaremos os maiores raios de convergência ótimo e unicidade de solução. Por fim, aplicaremos resultados obtidos para as condições Lipschitz e Smale. No capítulo 4, mostraremos que sob certas condições, a sequência gerada pelo método quase Gauss-Newton inexato está bem definida e converge para uma solução de (1). Por fim, aplicaremos resultados obtidos para as condições Lipschitz e Smale. No capítulo 5, mostraremos que sob certas condições, a sequência gerada pelo algoritmo de Gauss-Newton converge para um ponto $x_* \in \mathbb{R}^n$ tal que $F(x_*) \in C$, em particular, x_* resolve (3). Além disso, aplicaremos os resultados obtidos para o caso onde o ponto inicial é um ponto regular e o caso onde o ponto inicial satisfaz a condição de Robinson. Por fim, apresentaremos resultados de convergência do algoritmo de Gauss-Newton sob condições Lipschitz e Smale.

Capítulo 1

Notações e Resultados Preliminares

Nosso objetivo neste capítulo é introduzir algumas notações e revisar alguns tópicos de operadores lineares, incluindo resultados da inversa generalizada de Moore-Penrose. Revisaremos também alguns resultados de derivadas, análise convexa, funções analíticas e regularidade. Estes assuntos serão necessários ao desenvolvimento dos capítulos seguintes.

Em todo lugar X e Y indicam espaços de Hilbert. A *bola aberta* e *fechada* de centro $a \in X$ e raio $\delta > 0$ são denotados, respectivamente, por

$$B(a, \delta) := \{x \in X;\ \|x - a\| < \delta\}, \qquad B[a, \delta] := \{x \in X;\ \|x - a\| \leqslant \delta\}.$$

A distância de um ponto $x \in X$ a um conjunto $W \subset X$ é dado por

$$d(x, W) := \inf\{\|x - w\| : w \in W\}.$$

O *polar* de um conjunto convexo fechado $W \subset X$ é o conjunto

$$W^o := \{z \in X :< z, w >\leq 0, \forall w \in W\}.$$

O *núcleo* de um operador linear $A : X \to Y$, denotado por $Ker(A)$, é o conjunto

$$Ker(A) = \{x \in X : Ax = 0)\}.$$

Denotamos $P(X)$ o conjunto de todos subconjuntos de X. Assim, a soma de um ponto $x \in X$ a um conjunto $Z \in P(X)$ é o conjunto dado por

$$y + Z = \{y + x : x \in Z\}.$$

A seguir definiremos norma de operadores, antes denotemos por $\mathcal{L}(X,Y)$ o espaço dos operadores lineares contínuos de X em Y.

Definição 1.1 *Seja $T \in \mathcal{L}(X,Y)$. Definimos a norma de operadores $\|.\|$ como sendo o número*

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

Note que valem as seguintes desigualdades:

i) $\|Au\| \leq \|A\|\|u\|$ para todo $A \in \mathcal{L}(X,Y)$ e $u \in X$.

ii) $\|A\,B\| \leq \|A\|\|B\|$ e $\|A+B\| \leq \|A\| + \|B\|$ para todo $A, B \in \mathcal{L}(X,Y)$.

Um resultado bem conhecido da teoria de operadores, o qual enunciaremos a seguir, é o lema de Banach.

Lema 1.1 *(Lema de Banach) Sejam $B \in \mathcal{L}(X,X)$ e I o operador identidade em X. Se $\|B-I\| < 1$, então B é invertível e vale $\|B^{-1}\| \leq 1/(1-\|B-I\|)$.*

Demonstração: Ver Lema 1, p.p. 189 de Smale [27] com $A = I$ e $c = \|B-I\|$. ∎

Lembramos do Teorema do Gráfico fechado que se $A \in \mathcal{L}(X,Y)$ é injetiva e tem imagem fechada, então A é isomorfo à sua imagem, i.e., $A^{-1} : ImA \to X$ é contínuo.

Dado $A \in \mathcal{L}(X,Y)$, denotemos a inversa generalizada de Moore-Penrose de A por $A^{\dagger}$. Se $A \in \mathcal{L}(X,Y)$ é injetiva e tem imagem fechada, então

$$A^{\dagger} := (A^*A)^{-1}A^*,$$

onde A^* denota a adjunta do operador linear A.

A seguir, daremos dois resultados a respeito da inversa generalizada de Moore-Penrose que serão necessários mais tarde, para garantir a boa definição do método de Gauss-Newton e do método quase Gauss-Newton inexato.

Lema 1.2 *Sejam $A, B \in \mathcal{L}(X,Y)$ tais que A e B têm imagens fechadas. Se A é injetiva, $E = B - A$ e $\|EA^{\dagger}\| < 1$, então B é injetiva.*

Demonstração: De fato, como $B = A + E = (I + EA^{\dagger})A$ e de hipótese $\|EA^{\dagger}\| < 1$, segue do Lema 1.1 que $I + EA^{\dagger}$ é invertível. Portanto, B é injetiva. ∎

O próximo lema é provado em Stewart [28] (ver também, Wedin [29]) para matrizes A e B de ordem $m \times n$ com $m \geq n$ e $posto(A) = posto(B) = n$. Porém, o resultado também é válido em um contexto mais geral conforme afirmamos abaixo.

Lema 1.3 *Sejam $A, B \in \mathcal{L}(X,Y)$ tais que A e B são injetivos e têm imagens fechadas. Assuma que $E = B - A$ e $\|A^\dagger\|\|E\| < 1$, então*

$$\|B^\dagger\| \leq \frac{\|A^\dagger\|}{1 - \|A^\dagger\|\|E\|}, \qquad \|B^\dagger - A^\dagger\| \leq \frac{\sqrt{2}\|A^\dagger\|^2\|E\|}{1 - \|A^\dagger\|\|E\|}.$$

Faremos agora uma breve revisão da derivada de uma função $F : X \to Y$.

Definição 1.2 *Seja $F : \Omega \to Y$ uma função definida no conjunto aberto $\Omega \subset X$. Um operador $T \in \mathcal{L}(X,Y)$ é a derivada de Fréchet de F em $x_0 \in \Omega$ se para todo $w \in X$ tal que $(w + x_0) \in \Omega$ vale a igualdade*

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

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

Se a função F' é diferenciável em Ω, a sua derivada é chamada de segunda derivada e será denotada por F''. Assim, $F''(x)$ é um elemento do espaço $\mathcal{L}(X, \mathcal{L}(X,Y))$ dos operadores lineares que levam X em $\mathcal{L}(X,Y)$. Para simplificar a notação identificaremos o espaço $\mathcal{L}(X, \mathcal{L}(X,Y))$ por $\mathcal{L}_2(X,Y)$.

Podemos generalizar o conceito acima, supondo que a n-ésima derivada de uma função F de X em Y é a derivada da derivada de ordem $n-1$. Assim, a n-ésima derivada, denotada por $F^{(n)}(x)$, será obviamente um elemento do espaço $\mathcal{L}_n(X,Y)$ dos operadores n-lineares de X em Y. Novamente indetificamos o espaço $\mathcal{L}(X, \mathcal{L}(X, \ldots, Y), \ldots))$ com o espaço dos operadores n-lineares $\mathcal{L}_n(X,Y)$.

Utilizando as notações acima temos a seguinte observação:

Observação 1.1 *Seja $F : \Omega \to Y$ uma função diferenciável no conjunto aberto $\Omega \subset X$. Então $F'(x) \in \mathcal{L}(X,Y)$ e*

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

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

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

Analogamente, se F é uma função n-vezes diferenciável em Ω, introduzimos a norma do operador n-linear $F^{(n)}(x) \in \mathcal{L}_n(X,Y)$ como sendo

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

Agora, estamos interessados em revisar alguns resultados da análise convexa e um resultado da teoria de funções analíticas.

Proposição 1.1 *Sejam $R > 0$ e $\varphi : [0, R) \to \mathbb{R}$ uma função convexa.*

i) *Então*

$$D^+\varphi(0) = \lim_{u \to 0^+} \frac{\varphi(u) - \varphi(0)}{u} = \inf_{0<u} \frac{\varphi(u) - \varphi(0)}{u},$$

onde $D^+\varphi(0)$ denota a derivada à direita de φ em zero;

ii) *Se $u, v, w \in [0, R)$, $u < w$ and $u \leq v \leq w$ então*

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

iii) *Se $\tau \in [0, 1]$ então a função $l : (0, R) \to \mathbb{R}$ definida por*

$$l(t) = \frac{\varphi(t) - \varphi(\tau t)}{t},$$

é crescente.

Demonstração: Ver Teorema 4.1.1 e observação 4.1.2, pp. 21 de Hiriart-Urruty and Lemaréchal [30]. ∎

Lembremos que uma função $F : \Omega \to Y$, onde $\Omega \subset X$ é *analítica* se, para cada ponto $a \in \Omega$ existe uma série de potências

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

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

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

Proposição 1.2 *Se $0 \leq t < 1$, então $\sum_{i=0}^{\infty}(i + 2)(i + 1)t^i = 2/(1 - t)^3$.*

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

$$g'(t) = (1 - t)^{-2}, \quad g''(t) = 2(1 - t)^{-3}, \ldots, \quad g^{(i)}(t) = i!(1 - t)^{-(i+1)}. \tag{1.1}$$

Agora usando a definição de função analítica e as inequações anteriores obtemos que

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

Derivando duas vezes a equação acima temos que

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

a qual combinada com a segunda equação em (1.1) prova o resultado desejada. ∎

Encerramos este capítulo, com um breve estudo de regularidade. O estudo de regularidade fornecerá ferramentas que auxiliarão na prova de convergência da sequência gerada pelo **Algoritmo 0.1** para uma solução de (3). Para mais detalhes a respeito de regularidade ver [3, 4].

Considere C como definido em (4), i.e., C é o conjunto de todos os pontos mínimos da função convexa h. Para cada $x \in \mathbb{R}^n$, defina o conjunto $D_C(x)$ associado a C como

$$D_C(x) := \{d \in \mathbb{R}^n : F(x) + F'(x)d \in C\}.$$

A seguir, provaremos uma relação entre os conjuntos $D_\Delta(x)$ (definido em (5)) e $D_C(x)$.

Proposição 1.3 *Seja $x \in \mathbb{R}^n$. Se $D_C(x) \neq \emptyset$ e $d(0, D_C(x)) \leq \Delta$, então*

$$D_\Delta(x) = \{d \in \mathbb{R}^n : \|d\| \leq \Delta, F(x) + F'(x)d \in C\} \subset D_C(x).$$

Como consequência, $d(0, D_\Delta(x)) = d(0, D_C(x))$.

Demonstração: Pela definição de C em (4) e $D_\Delta(x)$ em (5) é imediato que

$$\{d \in \mathbb{R}^n : \|d\| \leq \Delta, F(x) + F'(x)d \in C\} \subset D_\Delta(x). \tag{1.2}$$

Agora, seja $d \in D_\Delta(x)$. Como $D_C(x) \neq \emptyset$ e $d(0, D_C(x)) \leq \Delta$, existe um $\bar{d} \in D_C(x)$ tal que $\|\bar{d}\| \leq \Delta$ e $F(x) + F'(x)\bar{d} \in C$. Daí, pela definição de C em (4) e $D_\Delta(x)$ em (5) obtemos $\bar{d} \in D_\Delta(x)$. Portanto, como $\bar{d}, d \in D_\Delta(x)$, usando novamente a definição de $D_\Delta(x)$ em (5), temos que

$$h(F(x) + F'(x)d) = h(F(x) + F'(x)\bar{d}).$$

Assim, usando $F(x) + F'(x)\bar{d} \in C$, a última equação e definição de C segue que $F(x) + F'(x)d \in C$, o que juntamente com (1.2) prova a primeira afirmação. A segunda afirmação, i.e., $D_\Delta(x) \subset D_C(x)$ é imediata da definição de $D_C(x)$. Para concluir a prova, primeiro note que a inclusão $D_\Delta(x) \subset D_C(x)$ implica que

$$d(0, D_\Delta(x)) \geq d(0, D_C(x)). \tag{1.3}$$

Como $D_C(x) \neq \emptyset$ e $d(0, D_C(x)) \leq \Delta$, existe $\bar{d} \in D_C(x)$ tal que

$$\|\bar{d}\| = d(0, D_C(x)) \leq \Delta.$$

Daí, pela definição de C em (4) e $D_\Delta(x)$ em (5) concluímos que $\bar{d} \in D_\Delta(x)$. Portanto,

$$d(0, D_\Delta(x)) \leq \|\bar{d}\| = d(0, D_C(x)),$$

o que juntamente com (1.3) conclui a prova. ∎

Definição 1.3 *Sejam $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função continuamente diferenciável e $h : \mathbb{R}^m \to \mathbb{R}$ uma função convexa de valores-reais. Um ponto $x_0 \in \mathbb{R}^n$ é chamado um ponto quase-regular da inclusão (4), i.e., da inclusão*

$$F(x) \in C := \{z \in \mathbb{R}^m : h(z) \leq h(y),\ \forall y \in \mathbb{R}^m\},$$

se existe um $r \in (0, +\infty)$ e uma função crescente de valores-positivos $\beta : [0, r) \to (0, +\infty)$ tais que

$$D_C(x) \neq \emptyset, \qquad d(0, D_C(x)) \leq \beta(\|x - x_0\|) d(F(x), C), \qquad \forall x \in B(x_0, r). \tag{1.4}$$

Seja $x_0 \in \mathbb{R}^n$ um ponto quase-regular da inclusão (4). Denotamos r_{x_0} o supremo dos r tais que (1.4) vale para alguma função crescente de valores-positivos β em $[0, r)$, i.e.,

$$r_{x_0} := \sup \{r : \exists\, \beta : [0, r) \to (0, +\infty) \text{ satisfazendo } (1.4)\}. \tag{1.5}$$

Seja $r \in [0, r_{x_0})$. O conjunto $\mathcal{B}_r(x_0)$ denota o conjunto de todas as funções crescentes de valores-positivos β em $[0, r)$ tais que (1.4) vale, i.e.,

$$\mathcal{B}_r(x_0) := \{\beta : [0, r) \to (0, +\infty) : \beta \text{ satisfaz } (1.4)\}.$$

Definimos

$$\beta_{x_0}(t) := \inf \left\{\beta(t) : \beta \in \mathcal{B}_{r_{x_0}}(x_0)\right\}, \qquad t \in [0, r_{x_0}). \tag{1.6}$$

O número r_{x_0} e a função β_{x_0} são chamados, respectivamente, o raio quase-regular e a função limitante quase-regular para um ponto quase-regular x_0.

Observação 1.2 *Note que, pela definição de r_{x_0} e β_{x_0} é imediato concluir que para* *todo* $r \leq r_{x_0}$ *tal que* $\lim_{t \to r^-} \beta(t) < +\infty$ *temos*

$$\beta_{x_0}(t) = \inf \{\beta(t) : \beta \in \mathcal{B}_r(x_0)\}, \qquad t \in [0, r).$$

Capítulo 2

Princípio Majorante

Para provarmos a convergência dos métodos de Gauss-Newton, usaremos uma condição majorante similar à usada em [15–17, 23]. Nosso objetivo neste capítulo é definir as funções majorantes, estudar algumas de suas propriedades e obter suas relações com a função não-linear F associada aos problemas (1) e (3).

2.1 A Função Majorante Radial

Nesta seção, definiremos a função majorante radial e estudaremos algumas propriedades relacionadas a ela. Os resultados obtidos aqui são os principais instrumentos no estudo de convergência local do método de Gauss-Newton e do método quase Gauss-Newton inexato para problemas de mínimos quadrados não-lineares.

Definição 2.1 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continuamente diferenciável tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$, $R > 0$,*

$$c := \|F(x_*)\|, \qquad \beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\left\{t \in [0, R) : B(x_*, t) \subset \Omega\right\}.$$

Suponha que $F'(x_*)^* F(x_*) = 0$ e $F'(x_*)$ seja injetiva. Então, uma função continuamente diferenciável $f_r : [0,\ R) \to \mathbb{R}$ é uma função majorante radial de F em $B(x_*, \kappa)$ se satisfaz

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq f_r'\left(\|x - x_*\|\right) - f_r'\left(\tau\|x - x_*\|\right), \tag{2.1}$$

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

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

h2) *f_r' é convexa e estritamente crescente;*

h3) $\mu := \sqrt{2}\, c\, \beta^2 D^+ f_r'(0) < 1$.

A seguir faremos uma análise da função definida acima.

2.1.1 Propriedades da Função Majorante radial (f_r)

Nesta subseção, estaremos interessados em estudar algumas propriedades da função majorante radial. Destacamos que para os resultados desta subseção não se faz necessária a condição (2.1), parte da definição de função majorante radial.

Proposição 2.1 *São válidas as seguintes desigualdades*

$$t f_r'(t) - f_r(t) > 0 \qquad e \qquad f_r(t) + t > 0, \qquad \forall\, t \in (0,\, R).$$

Demonstração: De **h2** temos que f' é estritamente crescente em $[0, R)$, então f_r é estritamente convexa em $[0, R)$. Assim,

$$f_r(0) > f_r(t) - t f_r'(t) \qquad \text{e} \qquad f_r(t) > f_r(0) + t f_r'(0), \qquad \forall\, t \in (0,\, R).$$

Com simples manipulações algébricas e usando **h1**, as desigualdades da proposição seguem das desigualdades acima. ∎

Proposição 2.2 *Defina a constante* $\nu =: \sup\{t \in [0, R) : \beta[f_r'(t) + 1] < 1\}$ *. Então, ν é positiva e é válida a seguinte desigualdade*

$$0 < \beta[f_r'(t) + 1] < 1, \qquad \forall\, t \in (0,\, \nu). \tag{2.2}$$

Demonstração: Primeiro, segue da continuidade de f_r' em $[0, R)$ e **h1** que

$$\lim_{t \to 0} \beta[f_r'(t) + 1] = 0.$$

Daí, existe $\delta > 0$ tal que $\beta[f_r'(t) + 1] < 1$ para todo $t \in [0,\, \delta)$. Portanto, $\delta \leq \nu$, o que estabelece a primeira afirmação.

Para concluir a prova, note que as desigualdades seguem da hipótese que f_r' é estritamente crescente, **h1** e definição de ν. ∎

Proposição 2.3 *As seguintes funções são crescentes:*

i) $[0,\, R) \ni t \mapsto 1/[1 - \beta(f_r'(t) + 1)]$;

ii) $(0,\, R) \ni t \mapsto [t f_r'(t) - f_r(t)]/t^2$;

iii) $(0,\, R) \ni t \mapsto [f_r'(t)+1]/t$;

iv) $(0,\, R) \ni t \mapsto f_r(t)/t$.

Como consequência, as funções

$$(0,\, R) \ni t \mapsto \frac{tf_r'(t)-f_r(t)}{t^2[1-\beta(f_r'(t)+1)]}, \qquad\qquad (0,\, R) \ni t \mapsto \frac{f_r'(t)+1}{t[1-\beta(f_r'(t)+1)]}.$$

também são crescentes.

Demonstração: O item **i** é imediato, pois f_r' é estritamente crescente em $[0, R)$.

Para provar o item **ii**, note que com algumas manipulações algébricas e **h1**, obtemos

$$\frac{tf_r'(t)-f_r(t)}{t^2} = \int_0^1 \frac{f_r'(t)-f_r'(\tau t)}{t}\, d\tau.$$

Daí, usando a Proposição 1.1 com $f_r' = \varphi$ segue que a função em consideração é crescente.

Para estabelecer o item **iii**, use $f_r'(0) = -1$ e Proposição 1.1 com $f_r' = \varphi$ e $\tau = 0$.

Prova do item **iv**. De **h2** temos que f_r' é estritamente crescente, então f_r é estritamente convexa em $[0, R)$. Usando que $f_r(0) = 0$, temos $f_r(t)/t = [f_r(t) - f_r(0)]/[t-0]$. Daí, o item **iv** segue da Proposição 1.1 com $f_r = \varphi$ e $\tau = 0$.

Para provar a última parte da proposição, combine os itens **i** e **ii** para a primeira função e o itens **i** e **iii** para a segunda função. ∎

Proposição 2.4 *Defina a constante*

$$\rho := \sup \left\{ t \in (0, \nu) : \frac{\beta[tf_r'(t)-f_r(t)] + \sqrt{2}c\beta^2[f_r'(t)+1]}{t[1-\beta(f_r'(t)+1)]} < 1 \right\}.$$

Então, ρ é positivo e são válidas as seguintes desigualdades

$$0 < \frac{\beta[tf_r'(t)-f_r(t)] + \sqrt{2}c\beta^2[f_r'(t)+1]}{t[1-\beta(f_r'(t)+1)]} < 1, \qquad \forall\, t \in (0,\, \rho). \tag{2.3}$$

Demonstração: Primeiro, note que com algumas manipulações algébricas e usando **h1** temos

$$\frac{\beta[tf_r'(t)-f_r(t)] + \sqrt{2}c\beta^2[f_r'(t)+1]}{t[1-\beta(f_r(t)+1)]} = \frac{\beta\left[f_r'(t) - \frac{f_r(t)-f_r(0)}{t-0}\right] + \sqrt{2}\,c\,\beta^2\,\frac{f_r'(t)-f_r'(0)}{t-0}}{1-\beta(f_r'(t)+1)}.$$

Segue da última equação, continuidade de f_r' e Proposição 1.1 com $\varphi = f_r'$ que

$$\lim_{t\to 0} \frac{\beta[tf_r'(t) - f_r(t)] + \sqrt{2}c\beta^2[f_r'(t)+1]}{t[1-\beta(f_r'(t)+1)]} = \sqrt{2}c\beta^2 D^+ f_r'(0).$$

Agora, como $\mu = \sqrt{2}c\beta^2 D^+ f_r'(0) < 1$ de **h3**, concluímos que existe $\delta > 0$ tal que

$$\frac{\beta[tf_r'(t) - f_r(t)] + \sqrt{2}c\beta^2[f_r'(t)+1]}{t[1-\beta(f_r'(t)+1)]} < 1, \qquad t \in (0, \delta),$$

Daí, segue que $\delta \leq \rho$, o que prova a primeira afirmação.

Para provar (2.3), use a primeira desigualdade da Proposição 2.1, f_r' estritamente crescente e desigualdade (2.2) para a primeira desigualdade e definição de ρ e última parte da Proposição 2.3 para a segunda desigualdade. ∎

Proposição 2.5 *Dadas as constantes ϑ, ω_1 e ω_2 satisfazendo $0 \leq \vartheta < 1$, $0 \leq \omega_2 < \omega_1$ e $\omega_1(\mu + \mu\vartheta + \vartheta) + \omega_2 < 1$, defina a constante*

$$\bar{\rho} := \sup\left\{t \in (0,\nu) : (1+\vartheta)\omega_1\beta\frac{tf_r'(t) - f_r(t) + \sqrt{2}c\beta[f_r'(t)+1]}{t[1-\beta(f_r'(t)+1)]} + \omega_1\vartheta + \omega_2 < 1\right\}.$$

Então a constante $\bar{\rho}$ é positiva e são válidas as seguintes desigualdades

$$0 < (1+\vartheta)\omega_1\beta\frac{tf_r'(t) - f_r(t) + \sqrt{2}c\beta[f_r'(t)+1]}{t[1-\beta(f_r'(t)+1)]} + \omega_1\vartheta + \omega_2 < 1, \quad \forall\, t \in (0, \bar{\rho}). \tag{2.4}$$

Demonstração: Primeiro, note que com simples manipulações algébricas e usando **h1**, temos

$$\lim_{t\to 0} \frac{tf_r'(t) - f_r(t)}{t[1-\beta(f_r'(t)+1)]} = \lim_{t\to 0} \frac{f_r'(t) - (f_r(t) - f_r(0))/t}{1-\beta(f_r'(t)+1)} = 0.$$

Agora, usando **h1**, simples manipulações algébricas e que f_r' é convexa, segue da Proposição 1.1 com $\varphi = f_r'$ que

$$\lim_{t\to 0} \frac{f_r'(t)+1}{t[1-\beta(f_r'(t)+1)]} = \lim_{t\to 0} \frac{(f_r'(t) - f_r'(0))/t}{1-\beta(f_r'(t)+1)} = D^+ f_r'(0).$$

Daí, combinando as duas equações acima é fácil concluir que

$$\lim_{t\to 0}(1+\vartheta)\omega_1\beta\frac{tf_r'(t) - f_r(t) + \sqrt{2}c\beta[f_r'(t)+1]}{t[1-\beta(f_r'(t)+1)]} + \omega_1\vartheta + \omega_2$$
$$= (1+\vartheta)\omega_1\sqrt{2}c\beta^2 D^+ f_r'(0) + \omega_1\vartheta + \omega_2.$$

Como $\mu = \sqrt{2}c\beta^2 D^+ f_r'(0)$ e $\omega_1(\mu + \mu\vartheta + \vartheta) + \omega_2 < 1$, obtemos que existe um $\delta > 0$

tal que

$$(1+\vartheta)\omega_1\beta\frac{tf_r'(t)-f_r(t)+\sqrt{2}c\beta[f_r'(t)+1]}{t[1-\beta(f_r'(t)+1)]}+\omega_1\vartheta+\omega_2<1, \qquad t\in(0,\delta),$$

Portanto, $\delta\le\rho$, o que prova a primeira afirmação.

Para provar (2.4), use a primeira desigualdade da Proposição 2.1, f_r' estritamente crescente e desigualdade (2.2) para a primeira desigualdade e definição de ρ e última parte da Proposição 2.3 para a segunda desigualdade. ∎

Proposição 2.6 *Seja $\beta_0 := \|[F'(x_*)^*F'(x_*)]^{-1}\|$. Defina a constante*

$$\sigma := \sup\left\{t\in(0,\kappa) : \frac{\beta(f_r(t)+t)+c\beta_0(f_r'(t)+1)}{t}<1\right\}.$$

Adicionalmente, se $c\beta_0 D^+f_r'(0)<1$, então σ é positiva e são válidas as seguintes desigualdades

$$0<\frac{\beta(f_r(t)+t)+c\beta_0(f_r'(t)+1)}{t}<1, \qquad \forall\, t\in(0,\,\sigma). \tag{2.5}$$

Demonstração: Primeiro, note que com algumas manipulações algébricas e usando **h1** temos

$$\frac{\beta(f_r(t)+t)+c\beta_0(f_r'(t)+1)}{t}=\beta\left[\frac{f_r(t)-f_r(0)}{t-0}-f_r'(0)\right]+c\beta_0\frac{f_r'(t)-f_r'(0)}{t-0}.$$

Segue da última equação, continuidade de f_r' e Proposição 1.1 com $\varphi=f_r'$ que

$$\lim_{t\to 0}\frac{\beta(f_r(t)+t)+c\beta_0(f_r'(t)+1)}{t}\;=\;c\,\beta_0\,D^+f_r'(0).$$

Agora, como $c\,\beta_0\,D^+f_r'(0)<1$, concluímos que existe $\delta>0$ tal que

$$\frac{\beta(f_r(t)+t)+c\beta_0(f_r'(t)+1)}{t}<1, \qquad t\in(0,\delta),$$

Daí, segue que $\delta\le\sigma$, o que prova a primeira afirmação.

Para provar (2.5), use a segunda desigualdade da Proposição 2.1 e que f_r' é estritamente crescente para a primeira desigualdade e definição de σ os itens **iii** e **iv** da Proposição 2.3 para a segunda desigualdade. ∎

2.1.2 Principais Relações entre f_r e F

Nesta subseção, apresentaremos as principais relações entre a função majorante radial f_r e a função não-linear F associada ao problema (1).

Lema 2.1 *Seja $x \in \Omega$. Se $\|x - x_*\| < \min\{\nu, \kappa\}$, então $F'(x)^*F'(x)$ é invertível e são válidas as seguintes desigualdades*

$$\left\|F'(x)^\dagger\right\| \leq \frac{\beta}{1 - \beta[f_r'(\|x - x_*\|) + 1]}, \quad \left\|F'(x)^\dagger - F'(x_*)^\dagger\right\| < \frac{\sqrt{2}\beta^2[f_r'(\|x - x_*\|) + 1]}{1 - \beta[f_r'(\|x - x_*\|) + 1]}.$$

Em particular, se $\|x - x_*\| < r$, onde $r = \min\{\kappa, \rho\}$ ou $r = \min\{\kappa, \bar{\rho}\}$, então $F'(x)^*F'(x)$ é invertível em $B(x_*, r)$.

Demonstração: Primeiro, para simplificar a demonstração definiremos as matrizes

$$A = F'(x_*), \qquad B = F'(x), \qquad E = F'(x) - F'(x_*). \tag{2.6}$$

Seja $x \in \Omega$ tal que $\|x - x_*\| < \min\{\nu, \kappa\}$. Daí, usando a definição de β, desigualdade (2.1) com $\tau = 0$ e desigualdade (2.2), obtemos que

$$\|F'(x) - F'(x_*)\|\|[F'(x_*)^*F'(x_*)]^{-1}F'(x_*)^*\| \leq \beta[f_r'(\|x - x_*\|) - f_r'(0)] < 1.$$

Combinando a última desigualdade e as definições em (2.6), temos

$$\|EA^\dagger\| \leq \|E\|\|A^\dagger\| < 1. \tag{2.7}$$

Considerando a inequação anterior e que $F'(x_*)$ é injetiva segue do Lema 1.2 que $F'(x)$ é injetiva. Então, $F'(x)^*F'(x)$ é invertível e pela definição de r também podemos concluir que $F'(x)^*F'(x)$ é invertível para todo $x \in B(x_*, r)$.

Agora, como já sabemos $F'(x)$ e $F'(x_*)$ são injetivas. Daí, usando as definições em (2.6) e inequação (2.7) as desigualdades do lema seguem das desigualdades do Lema 1.3. ∎

É conveniente estudar o erro linear de F para cada ponto em Ω, por isso definimos

$$E_F(x, y) := F(y) - [F(x) + F'(x)(y - x)], \qquad y,\, x \in \Omega. \tag{2.8}$$

Iremos limitar este erro pelo erro da linearização da função majorante radial f_r

$$e_{f_r}(t, u) := f_r(u) - [f_r(t) + f_r'(t)(u - t)], \qquad t,\, u \in [0, R). \tag{2.9}$$

Lema 2.2 *Se $\|x_* - x\| < \kappa$, então vale $\|E_F(x, x_*)\| \leq e_{f_r}(\|x - x_*\|, 0)$.*

Demonstração: Como $B(x_*, \kappa)$ é um conjunto convexo, segue $x_* + \tau(x - x_*) \in B(x_*, \kappa)$, para $0 \leq \tau \leq 1$. Daí, usando definição de E_F, F é continuamente diferenciável em Ω e algumas manipulações algébricas, obtemos que

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

Agora, considerando a desigualdade acima junto com (2.1) é fácil ver que

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

Portanto, combinando a última desigualdade com a definição (2.9) segue a desigualdade desejada. ∎

Denominamos como o passo de Gauss-Newton para a função F a seguinte igualdade:

$$S_F(x) := -F'(x)^{\dagger} F(x). \tag{2.10}$$

Lema 2.3 *Se* $\|x - x_*\| < \min\{\nu, \kappa\}$, *então*

$$\|S_F(x)\| \leq \frac{\beta e_{f_r}(\|x - x_*\|, 0) + \sqrt{2}c\beta^2[f_r'(\|x - x_*\|) + 1]}{1 - \beta[f_r'(\|x - x_*\|) + 1]} + \|x - x_*\|.$$

Demonstração: Usando (2.10), definição da inversa de Moore-Penrose, $F'(x_*)^* F(x_*) = 0$ e (2.8), segue com algumas manipulações algébricas que

$$\begin{aligned}
\|S_F(x)\| &= \|F'(x)^{\dagger} (F(x_*) - [F(x) + F'(x)(x_* - x)]) \\
&\quad + (F'(x)^{\dagger} - F'(x_*)^{\dagger})F(x_*) + x_* - x\| \\
&\leq \|F'(x)^{\dagger}\|\|E_F(x, x_*)\| + \|F'(x)^{\dagger} - F'(x_*)^{\dagger}\|\|F(x_*)\| + \|x - x_*\|.
\end{aligned}$$

Agora, a última inequação junto com os Lemas 2.1 e 2.2 e definição de c, implica que

$$\|S_F(x)\| \leq \frac{\beta e_{f_r}(\|x - x_*\|, 0)}{1 - \beta[f_r'(\|x - x_*\|) + 1]} + \frac{\sqrt{2}c\beta^2[f_r'(\|x - x_*\|) + 1]}{1 - \beta[f_r'(\|x - x_*\|) + 1]} + \|x - x_*\|,$$

a qual é equivalente à inequação desejada. ∎

2.2 A função Majorante Esférica

Nesta seção, definiremos a função majorante esférica e estudaremos algumas propriedades de uma determinada função auxiliar associada a ela. Os resultados aqui são os principais instrumentos no estudo de convergência da sequência gerada pelo algoritmo de Gauss-Newton para problemas de otimização de composição convexa.

Definição 2.2 *Sejam $R > 0$, $x_0 \in \mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função continuamente diferenciável. Então, uma função duas vezes diferenciável $f_e : [0,\ R) \to \mathbb{R}$ é uma função majorante esférica para a função F em $B(x_0, R)$ se satisfaz*

$$\|F'(y) - F'(x)\| \leq f_e'(\|y-x\| + \|x-x_0\|) - f_e'(\|x-x_0\|), \tag{2.11}$$

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

a1) *$f_e(0) = 0$, $f_e'(0) = -1$;*

a2) *f_e' é convexa e estritamente crescente.*

Neste caso também é conveniente limitar o erro linear de F pelo erro linear de f_e.

Lema 2.4 *Tome*

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

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

$$\|E_F(x,y)\| \leqslant e_{f_e}(t,v)\left(\frac{\|y-x\|}{v-t}\right)^2.$$

Demonstração: Como $B(x_0, R)$ é um conjunto convexo, segue que

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

Daí, como F é continuamente diferenciável em $B(x_0, R)$, (2.8) é equivalente a

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

Considerando a equação anterior e (2.11) obtemos que

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

Agora, usando a convexidade de f', as hipóteses $\|x-x_0\| < t$, $\|y-x\| < v-t$, $v < R$ e Proposição 1.1 segue que, para qualquer $u \in [0,1]$

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

Combinando as duas equações acima obtemos que

$$\|E_F(x,y)\| \leq \int_0^1 [f_e'(t+u(v-t)) - f_e'(t)]\frac{\|y-x\|^2}{v-t}\,du,$$

a qual, depois de calculada a integral e usando a definição em (2.9), dá-nos a inequação desejada. ∎

2.2.1 Propriedades da Função auxiliar $f_{\xi,\alpha}$

Nesta subseção, estaremos interessados em definir e estudar algumas propriedades de uma determinada função auxiliar associada com a função majorante esférica. Destacamos que para os resultados desta subseção não se faz necessária a condição (2.11), parte da definição de função majorante esférica.

Inicialmente, seja $f_e : [0,\ R) \to \mathbb{R}$ uma função majorante esférica para a função F em $B(x_0, R)$. Tome $\xi > 0$, $\alpha > 0$ e defina a *função auxiliar* $f_{\xi,\alpha} : [0, R) \to \mathbb{R}$,

$$f_{\xi,\alpha}(t) = \xi + (\alpha - 1)t + \alpha f_e(t). \tag{2.12}$$

Considere as seguintes condições para $f_{\xi,\alpha}$:

a3) existe $t_* \in (0, R)$ tal que $f_{\xi,\alpha}(t) > 0$ para todo $t \in (0, t_*)$ e $f_{\xi,\alpha}(t_*) = 0$;

a4) $f'_{\xi,\alpha}(t_*) < 0$.

Daqui em diante, assumimos que $f_e : [0,\ R) \to \mathbb{R}$ é uma função majorante esférica para a função F em $B(x_0, t_*)$ e **a3** vale. A condição **a4** será válida somente quando explicitamente afirmada.

No capítulo 6, veremos que a sequência gerada pelo algoritmo de Gauss-Newton para resolver problemas de otimização de composição convexa será "majorada" pela sequência de Newton associada a esta função auxiliar.

Proposição 2.7 *As seguintes afirmações são válidas:*

i) $f_{\xi,\alpha}(0) = \xi > 0$, $f'_{\xi,\alpha}(0) = -1$;

ii) $f'_{\xi,\alpha}$ *é convexa e estritamente crescente.*

Demonstração: Imediato, simplesmente use a definição em (2.12), **a1** e **a2**. ∎

Proposição 2.8 *A função* $f_{\xi,\alpha}$ *é estritamente convexa e*

$$f_{\xi,\alpha}(t) > 0, \quad f'_{\xi,\alpha}(t) < 0, \qquad t < t - f_{\xi,\alpha}(t)/f'_{\xi,\alpha}(t) < t_*, \qquad \forall\, t \in [0, t_*). \quad (2.13)$$

Além disso, $f'_{\xi,\alpha}(t_*) \leq 0$.

Demonstração: Como $f'_{\xi,\alpha}$ é estritamente crescente (Proposição 2.7), segue que $f_{\xi,\alpha}$ é estritamente convexa.

A primeira inequação em (2.13) segue imediato de **a3**. Agora, como $f_{\xi,\alpha}$ é estritamente convexa,

$$0 = f_{\xi,\alpha}(t_*) > f_{\xi,\alpha}(t) + f'_{\xi,\alpha}(t)(t_* - t), \qquad t \in [0, t_*). \quad (2.14)$$

A segunda inequação em (2.13) segue da inequação acima e **a3**. A terceira inequação em (2.13) segue da primeira e da segunda desigualdade. A última inequação em (2.13) é obtida por dividir a inequação em (2.14) por $-f'_{\xi,\alpha}(t)$ (o qual é positivo) e manipulações algébricas na inequação resultante.

Agora, como $f_{\xi,\alpha} > 0$ em $[0, t_*)$ e $f_{\xi,\alpha}(t_*) = 0$, devemos ter $f'_{\xi,\alpha}(t_*) \leq 0$. ∎

Seja $n_{f_{\xi,\alpha}}$ a função iteração de Newton associada à função auxiliar $f_{\xi,\alpha}$, definida por

$$\begin{array}{rcl} n_{f_{\xi,\alpha}} : [0, t_*) & \to & \mathbb{R} \\ t & \mapsto & t - f_{\xi,\alpha}(t)/f'_{\xi,\alpha}(t). \end{array} \quad (2.15)$$

Da segunda desigualdade em (2.13), temos que a função iteração de Newton $n_{f_{\xi,\alpha}}$ está bem definida em $[0, t_*)$.

Proposição 2.9 *Para cada* $t \in [0, t_*)$ *é válido que* $\xi \leq n_{f_{\xi,\alpha}}(t)$.

Demonstração: A Proposição 2.8 implica que $f_{\xi,\alpha}$ é estritamente convexa. Daí usando o item **i** da Proposição 2.7 é fácil ver $t-\xi \geq -f_{\xi,\alpha}(t)$. Então, segue da definição acima que

$$n_{f_{\xi,\alpha}}(t)-\xi = t-\frac{f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)}-\xi \geq -f_{\xi,\alpha}(t)-\frac{f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)} = \frac{f_{\xi,\alpha}(t)}{-f'_{\xi,\alpha}(t)}[f'_{\xi,\alpha}(t)+1], \quad \forall\, t \in [0,t_*).$$

A Proposição 2.7 implica que $f'_{\xi,\alpha}(0) = -1$ e $f'_{\xi,\alpha}$ é estritamente crescente. Deste modo, segue que $f'_{\xi,\alpha}(t)+1 \geq 0$, para todo $t \in [0,t_*)$. Portanto, combininando a equação acima com duas primeiras inequações na Propositição 2.8 a desigualdade segue. ∎

Proposição 2.10 *A função iteração de Newton $n_{f_{\xi,\alpha}}$ leva $[0,t_*)$ em $[0,t_*)$ e*

$$t < n_{f_{\xi,\alpha}}(t), \qquad t_* - n_{f_{\xi,\alpha}}(t) \leqslant \frac{1}{2}(t_*-t), \qquad \forall\, t \in [0,t_*). \tag{2.16}$$

Se $f_{\xi,\alpha}$ também satisfaz **a4**, *i.e., $f'_{\xi,\alpha}(t_*) < 0$, então*

$$t_* - n_{f_{\xi,\alpha}}(t) \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)}(t_*-t)^2, \qquad \forall\, t \in [0,t_*). \tag{2.17}$$

Demonstração: As primeiras duas afirmações da proposição seguem trivialmente da última inequação em (2.13).

Prova da segunda inequação em (2.16). Tomando $t \in [0,t_*)$ e usando (2.15) e a continuidade de $f'_{\xi,\alpha}$ obtemos que

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

Como $f'_{\xi,\alpha}$ é convexa (Proposição 2.7) e $t < t_*$, segue da Proposição 1.1 que

$$f'_{\xi,\alpha}(u) - f'_{\xi,\alpha}(t) \leqslant \left[f'_{\xi,\alpha}(t_*) - f'_{\xi,\alpha}(t)\right]\frac{u-t}{t_*-t}, \qquad \forall\, u \in [t,t^*].$$

Levando em conta a positividade de $-1/f'_\alpha(t)$ (segunda inequação em (2.13)) e combinando as duas últimas equações temos que

$$t_* - n_{f_{\xi,\alpha}}(t) \leqslant (-1/f'_{\xi,\alpha}(t)) \int_t^{t_*} \left[f'_{\xi,\alpha}(t_*) - f'_{\xi,\alpha}(t)\right]\frac{u-t}{t_*-t}\, du.$$

Resolvendo a integral acima obtemos que

$$t_* - n_{f'_{\xi,\alpha}}(t) \leqslant \frac{1}{2}\left(\frac{f'_{\xi,\alpha}(t_*) - f'_{\xi,\alpha}(t)}{-f'_{\xi,\alpha}(t)}\right)(t_* - t). \tag{2.18}$$

Portanto, a inequação acima junto com $f'_{\xi,\alpha}(t_*) \leq 0$ e $f'_{\xi,\alpha}(t) < 0$ implica a segunda inequação em (2.16).

Finalmente, assuma que $f_{\xi,\alpha}$ satisfaz **a4**. Tome $t \in [0, t_*)$. Como $f'_{\xi,\alpha}$ é crescente, $f'_{\xi,\alpha}(t_*) < 0$ e $f'_{\xi,\alpha}(t) < 0$, obtemos que

$$\begin{aligned}\frac{f'_{\xi,\alpha}(t_*) - f'_{\xi,\alpha}(t)}{-f'_{\xi,\alpha}(t)} &\leq \frac{f'_{\xi,\alpha}(t_*) - f'_{\xi,\alpha}(t)}{-f'_{\xi,\alpha}(t_*)} \\ &= \frac{1}{-f'_{\xi,\alpha}(t_*)} \frac{f'_{\xi,\alpha}(t_*) - f'_{\xi,\alpha}(t)}{t_* - t}(t_* - t) \leq \frac{f''_{\xi,\alpha}(t_*)}{-f'_{\xi,\alpha}(t_*)}(t_* - t),\end{aligned}$$

onde a última inequação segue do fato que $f_{\xi,\alpha}$ é duas vezes diferenciável em $[0, t_*)$. Combinando a inequação acima com (2.18), segue a desigualdade em (2.17). ∎

A sequência de Newton $\{t_k\}$ para resolver a equação $f_{\xi,\alpha}(t) = 0$ com ponto inicial $t_0 = 0$ é definida como

$$t_0 = 0, \quad t_{k+1} = n_{f_{\xi,\alpha}}(t_k), \qquad k = 0, 1, \ldots . \tag{2.19}$$

Daí, segue da Proposição 2.10 que

Corolário 2.1 *A sequência $\{t_k\}$ está bem definida, é estritamente crescente, está contida em $[0, t_*)$ e converge Q-linearmente para t_* como segue*

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

Se $f_{\xi,\alpha}$ também satisfaz **a4**, *então $\{t_k\}$ converge Q-quadraticamente para t_* como segue*

$$t_* - t_{k+1} \leq \frac{f''_{\xi,\alpha}(t_*)}{-2f'_{\xi,\alpha}(t_*)}(t_* - t_k)^2, \quad k = 0, 1, \ldots .$$

Proposição 2.11 *A função $[0, t_*) \ni t \to -(f_{\xi,\alpha}(t))/(f'_{\xi,\alpha}(t))$ é decrescente.*

Demonstração: Usando o fato de que $f_{\xi,\alpha}$ é duas vezes diferenciável e $f'_{\xi,\alpha} < 0$ em $[0, t_*)$, é fácil ver que

$$\left(\frac{-f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)}\right)' = \frac{f_{\xi,\alpha}(t) f''_{\xi,\alpha}(t) - (f'_{\xi,\alpha}(t))^2}{(f'_{\xi,\alpha}(t))^2}, \qquad \forall\, t \in [0, t_*).$$

Daí, é suficiente mostrar que

$$f_{\xi,\alpha}(t)f''_{\xi,\alpha}(t) - (f'_{\xi,\alpha}(t))^2 < 0, \qquad \forall\, t \in [0, t_*). \tag{2.20}$$

Como $f_{\xi,\alpha}$ é estritamente convexa (Proposição 2.8) e $f'_{\xi,\alpha}$ é convexa (Proposição 2.7), temos que

$$0 > f_{\xi,\alpha}(t)+f'_{\xi,\alpha}(t)(t_*-t),\ f''_{\xi,\alpha}(t) > 0,\ f'_{\xi,\alpha}(t_*) \geq f'_{\xi,\alpha}(t)+f''_{\xi,\alpha}(t)(t_*-t), \quad \forall\, t \in [0, t_*).$$

Usando as inequações acima e a segunda inequação em (2.13), obtemos que

$$f_{\xi,\alpha}(t)f''_{\xi,\alpha}(t) - (f'_{\xi,\alpha}(t))^2 < f'_{\xi,\alpha}(t)(t-t_*)f''_{\xi,\alpha}(t) - (f'_{\xi,\alpha}(t))^2 \leq -f'_{\xi,\alpha}(t)f'_{\xi,\alpha}(t_*),$$

a qual combinado com a Proposição 2.8 implica na inequação em (2.20). Portanto, a proposição segue. ∎

Proposição 2.12 *A desigualdade $\xi < t_*$ é válida. Sejam $\eta \geq 1$ e β_{x_0} uma função crescente de valores positivos em $[0, t_*)$. Se*

$$\alpha \geq \frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)(f'_e(t)+1)+1}, \qquad \forall\, \xi \leq t < t_*, \tag{2.21}$$

então

$$\eta\beta_{x_0}(t)/\alpha \leq -1/f'_{\xi,\alpha}(t), \qquad \forall\, \xi \leq t < t_*.$$

Demonstração: Primeiro, usando que $f_{\xi,\alpha}$ é estritamente convexa (Proposição 2.8), definição de t_* em **a3** e Proposição (2.7) segue que

$$0 = f_{\xi,\alpha}(t_*) > f_{\xi,\alpha}(0) + f'_{\xi,\alpha}(0)(t_* - 0) = \xi - t_*,$$

o que prova a primeira afirmação.

Agora, combinando a inequação em (2.21), **a1** e **a2**, obtemos com simples cálculos que

$$\alpha\eta\beta_{x_0}(t)(f'(t)+1) + \alpha \geq \eta\beta_{x_0}(t), \qquad \forall\, \xi \leq t < t_*.$$

Daí, usando $f'_{\xi,\alpha}(t) = (\alpha-1)+\alpha f'(t)$ e algumas manipulações algébricas, temos que

$$\eta\beta_{x_0}(t)f'_{\xi,\alpha}(t) \geq -\alpha, \qquad \forall\, \xi \leq t < t_*,$$

a qual combinada com a segunda inequação em (2.13) prova a inequação desejada. ∎

Proposição 2.13 *Seja $0 < \bar{\alpha} < \alpha$ com as correspondentes funções auxiliares $f_{\xi,\bar{\alpha}}$ e $f_{\xi,\alpha}$ e seus menores zeros $t_{\bar{*}}$ e t_*, respectivamente. Então as seguintes afirmações são válidas:*

i) $f_{\xi,\bar{\alpha}} < f_{\xi,\alpha}$ *em* $(0, R)$*;*

ii) $f'_{\xi,\bar{\alpha}} < f'_{\xi,\alpha}$ *em* $(0, R)$*;*

iii) $t_{\bar{*}} < t_*$.

Demonstração: Como f'_e é estritamente crescente (**a2**) segue que f_e é estritamente convexa. Daí usando **a1**, temos que $f_e(t) + t > 0$, para todo $t \in (0, t_*)$. Portanto, como $\alpha > \bar{\alpha} > 0$ e fácil ver que

$$\bar{\alpha}(t + f_e(t)) < \alpha(t + f_e(t)), \qquad \forall\, t \in [0, t_*),$$

Para concluir a prova do item **i**, some $\xi - t$ em ambos os lados da inequação anterior e use a definição em (2.12).

Para provar o item **ii**, note que de **a2** temos que f'_e é estritamente crescente, isto junto com $\alpha > \bar{\alpha}$ implica que $(\alpha - \bar{\alpha})(f'_e(t) - f'_e(0)) > 0$, para todo $t \in (0, t_*)$. Daí, usando **a1** e algumas manipulações algébricas, obtemos que

$$(\bar{\alpha} - 1) + \bar{\alpha} f'_e(t) < (\alpha - 1) + \alpha f'_e(t), \qquad \forall\, t \in [0, t_*).$$

Portanto, a afirmação segue da inequação acima e definição em (2.12).

Para estabelecer o item **iii**, use item **i** e definições de $t_{\bar{*}}$ e t_* em **a3**. ∎

Capítulo 3

Análise Local do Método de Gauss-Newton

Neste capítulo, apresentaremos uma análise de convergência local do método de Gauss-Newton para resolver problemas de mínimos quadrados não-lineares. Comprovaremos que sob certas condições a sequência gerada pelo método de Gauss-Newton está bem definida e converge linearmente para uma solução do problema em consideração. Encerraremos, aplicando os resultados obtidos para a condição Lipschitz clássica e condição de Smale para funções analíticas. Os resultados deste capítulo já se encontram publicados em [26].

Inicialmente, sejam X e Y espaços de Hilbert real ou complexo, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continualmente diferenciável. Considere o seguinte *problema dos mínimos quadrados não-linear*

$$\min \ \|F(x)\|^2. \tag{3.1}$$

Se $F'(x)$ é injetiva e tem imagem fechada para todo $x \in \Omega$, o método de Gauss-Newton usado para encontrar pontos estacionários do problema acima é formalmente descrito como: Dado $x_0 \in \Omega$ defina

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

A seguir, faremos uma análise de convergência local do método definido acima.

3.1 Convergência do Método de Gauss-Newton

Nesta seção, provaremos a convergência local do método de Gauss-Newton para resolver (3.1). Com algumas exigências para a função F e admitindo que o ponto $x_* \in \Omega$ é um ponto estacionário de (3.1), mostraremos que tomando o ponto inicial

x_0 numa vizinhança apropriada de x_*, a sequência gerada pelo método de Gauss-Newton está bem definida e converge linearmente para x_*. Além disso, determinaremos os raios de convergência ótima e de unicidade do ponto estacionário. Estas afirmações se encontram no teorema:

Teorema 3.1 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continuamente diferenciável tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$, $R > 0$,*

$$c := \|F(x_*)\|, \qquad \beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

Suponha que $F'(x_*)^* F(x_*) = 0$, $F'(x_*)$ é injetiva e que existe uma função continuamente diferenciável $f_r : [0,\ R) \to \mathbb{R}$ tal que

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq f_r'(\|x - x_*\|) - f_r'(\tau\|x - x_*\|), \tag{3.2}$$

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

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

h2) *f_r' é convexa e estritamente crescente;*

h3) *$\mu := \sqrt{2}\, c\, \beta^2 D^+ f_r'(0) < 1$.*

Sejam as constantes positivas $\nu := \sup\{t \in [0, R) : \beta[f_r'(t) + 1] < 1\}$,

$$\rho := \sup\left\{t \in (0, \nu) : \frac{\beta[t f_r'(t) - f_r(t)] + \sqrt{2}c\beta^2[f_r'(t) + 1]}{t[1 - \beta(f_r'(t) + 1)]} < 1\right\}, \quad r := \min\{\kappa,\ \rho\}.$$

Então, o método de Gauss-Newton para resolver (3.1), com ponto inicial $x_0 \in B(x_*, r)/\{x_*\}$

$$x_{k+1} = x_k - F'(x_k)^\dagger F(x_k), \qquad k = 0, 1 \ldots, \tag{3.3}$$

está bem definido, a sequência $\{x_k\}$ está contida em $B(x_*, r)$, converge para x_* e vale

$$\begin{aligned} \|x_{k+1} - x_*\| \leq{}& \frac{\beta[f_r'(\|x_0 - x_*\|)\|x_0 - x_*\| - f_r(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]}\|x_k - x_*\|^2 \\ &+ \frac{\sqrt{2}c\beta^2[f_r'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\|[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]}\|x_k - x_*\|, \qquad k = 0, 1, \ldots. \end{aligned} \tag{3.4}$$

Além disso, se $[\beta(\rho f_r'(\rho) - f_r(\rho)) + \sqrt{2}c\beta^2(f_r'(\rho) + 1)]/[\rho(1 - \beta(f_r'(\rho) + 1))] = 1$ e $\rho < \kappa$, então $r = \rho$ é o melhor raio de convergência possível.
Adicionalmente, se

h4) $c\,\beta_0\, D^+ f_r'(0) < 1$, *então* x_* *é o único ponto estacionário de* $F(x)^*F(x)$ *em* $B(x_*, \sigma)$, *onde*

$$\sigma := \sup\{t \in (0,\kappa) : [\beta(f_r(t)/t{+}1){+}c\beta_0(f_r'(t){+}1)/t] < 1\}, \beta_0 := \|[F'(x_*)^*F'(x_*)]^{-1}\|.$$

Observação 3.1 *Com o objetivo de deixar o Teorema acima "auto-contido", enunciamos novamente, mas desta vez de maneira implícita, a definição de função majorante radial. Portanto, todas as afirmações que faremos neste capítulo envolvendo a função majorante radial* f_r *e suas relações com a função não-linear* F *foram provadas nas Seções 2.1.1 e 2.1.2, respectivamente.*

Observação 3.2 *Se* $c = 0$ *(o chamado caso residual-zero), a inequação* (3.4) *implica que o método de Gauss-Newton converge Q-quadraticamente para* x_*. *Daí, seu comportamento é similar ao método de Newton (ver [15] e [22]). Se* c *é relativamente pequeno (o chamado caso residual-pequeno), a inequação* (3.4) *implica que o método de Gauss-Newton converge Q-linearmente para* x_*. *Porém, se* c *é relativamente grande (o chamado caso residual-grande), o método de Gauss-Newton pode não ser localmente convergente, ver condição* **h3** *e exemplo 10.2.4 na pp.225 de [6]. Portanto, podemos concluir que o método de Gauss-Newton comporta-se melhor nos problemas de residual- zero-ou-pequeno que nos problemas de residual-grande, enquanto o método de Newton é igualmente eficiente em todos estes casos.*

Como Ω é aberto é imediato concluir que $\kappa > 0$. Agora, segue das Proposições 2.2, 2.4 e 2.6 que as constantes ν, ρ e σ são positivas. Consequentemente, $r > 0$.

Para facilitar a demonstração do Teorema 3.1, apresentaremos a seguir uma seção conténdo três lemas auxiliares.

3.1.1 Resultados Auxiliares

Nesta subseção nosso objetivo é demonstrar três lemas que nos auxiliarão na prova de convergência do Teorema 3.1.

Inicialmente, note que o Lema 2.1 garante, em particular, que $F'(x)^*F'(x)$ é invertível em $B(x_*, r)$, assim temos a boa definição da aplicação iteração de Gauss-Newton G_F nesta região

$$\begin{array}{rcl} G_F : B(x_*, r) & \to & Y \\ x & \mapsto & x - F'(x)^\dagger F(x). \end{array} \tag{3.5}$$

Podemos aplicar uma iteração de Gauss-Newton em qualquer $x \in B(x_*, r)$ para obter $G_F(x)$, o qual pode não pertencer a $B(x_*, r)$ ou até mesmo não pertencer ao domínio de F. Assim, isto é suficiente para garantir a boa definição de apenas uma iteração. Para assegurar que a iteração de Gauss-Newton possa ser repetida indefinidamente, precisamos de mais alguns resultados.

Lema 3.1 *Seja $x \in \Omega$. Se $\|x - x_*\| < r$, então*

$$\|G_F(x) - x_*\| \leq \frac{\beta[f_r'(\|x-x_*\|)\|x-x_*\| - f_r(\|x-x_*\|)]}{\|x-x_*\|^2[1-\beta(f_r'(\|x-x_*\|)+1)]}\|x-x_*\|^2 + \frac{\sqrt{2}c\beta^2[f_r'(\|x-x_*\|)+1]}{\|x-x_*\|[1-\beta(f_r'(\|x-x_*\|)+1)]}\|x-x_*\|.$$

Em particular,

$$\|G_F(x) - x_*\| < \|x - x_*\|.$$

Demonstração: Primeiro, como $\|x - x_*\| < r$, o Lema 2.1 implica que $F'(x)^*F'(x)$ é invertível; então $G_F(x)$ está bem definida. Agora, usando (3.5) e que $F'(x_*)^*F(x_*) = 0$ obtemos, após simples cálculos que

$$\begin{aligned} G_F(x) - x_* =&\ x - x_* - \left[F'(x)^*F'(x)\right]^{-1}F'(x)^*F(x) \\ =&\ \left[F'(x)^*F'(x)\right]^{-1}F'(x)^*[F'(x)(x-x_*) - F(x) + F(x_*)] \\ &+ \left[F'(x_*)^TF'(x_*)\right]^{-1}F'(x_*)^*F(x_*) - \left[F'(x)^*F'(x)\right]^{-1}F'(x)^*F(x_*). \end{aligned}$$

É fácil ver que reunido a última equação, definição da inversa de Moore-Penrose e (2.8), obtemos

$$\|G_F(x) - x_*\| \leq \left\|F'(x)^\dagger\right\| \|E_F(x, x_*)\| + \left\|F'(x_*)^\dagger - F'(x)^\dagger\right\| \|F(x_*)\|.$$

Como $c = \|F(x_*)\|$, combinando a última desigualdade com os Lemas 2.1 e 2.2 temos que

$$\|G_F(x) - x_*\| \leq \frac{\beta e_{f_r}(\|x-x_*\|, 0)}{1-\beta(f_r'(\|x-x_*\|)+1)} + \frac{\sqrt{2}c\beta^2(f_r'(\|x-x_*\|)+1)}{1-\beta(f_r'(\|x-x_*\|)+1)}.$$

Daí, usando a definição em (2.9) e **h1**, concluímos que

$$\|G_F(x)-x_*\| \leq \frac{\beta[f_r'(\|x-x_*\|)\|x-x_*\| - f_r(\|x-x_*\|)]}{1-\beta(f_r'(\|x-x_*\|)+1)} + \frac{\sqrt{2}c\beta^2[f_r'(\|x-x_*\|)+1]}{1-\beta(f_r'(\|x-x_*\|)+1)},$$

que é equivalente à primeira desigualdade do lema.

Para concluir a prova, note que o lado direito da desigualdade acima também é equivalente a

$$\left[\frac{\beta[f_r'(\|x-x_*\|)\|x-x_*\| - f_r(\|x-x_*\|)] + \sqrt{2}c\beta^2[f_r'(\|x-x_*\|)+1]}{\|x-x_*\|[1-\beta(f_r'(\|x-x_*\|)+1)]}\right]\|x-x_*\|.$$

Por outro lado, como $x \in B(x_*, r)/\{x_*\}$, i.e., $0 < \|x - x_*\| < r \leq \rho$ segue da desigualdade (2.3) com $t = \|x - x_*\|$, que o termo entre colchetes na equação acima é menor que um, o que conclui a prova. ∎

Nos próximos dois lemas, obtemos os raios de convergência ótimo e de unicidade do ponto estacionário.

Lema 3.2 *Se* $[\beta(\rho f_r'(\rho) - f_r(\rho)) + \sqrt{2}c\beta^2(f_r'(\rho)+1)]/[\rho(1-\beta(f_r'(\rho)+1))] = 1$ *e* $\rho < \kappa$, *então* $r = \rho$ *é o melhor raio de convergência possível.*

Demonstração: Seja a função $h : (-\kappa,\, \kappa) \to \mathbb{R}$, definida por

$$h(t) = \begin{cases} -t/\beta + t - f_r(-t), & t \in (-\kappa,\, 0], \\ -t/\beta + t + f_r(t), & t \in [0,\, \kappa). \end{cases} \tag{3.6}$$

É simples mostrar que $h(0) = 0$, $h'(0) = -1/\beta$, $h'(t) = -1/\beta + 1 + f_r'(|t|)$ e

$$|h'(t) - h'(\tau t)| \leq f_r'(|t|) - f_r'(\tau|t|), \quad \tau \in [0, 1], \quad t \in (-\kappa,\, \kappa).$$

Assim, $F = h$ satisfaz todas as hipóteses do Teorema 3.1 com $c = |h(0)| = 0$. Como $\rho < \kappa$, é suficiente mostrar que o método de Gauss-Newton para resolver (3.1), com $F = h$ e ponto inicial $x_0 = \rho$ não converge. Neste caso, como $c = 0$ nossa exigência adicional torna-se

$$(\beta(\rho f_r'(\rho) - f_r(\rho))/\rho(1 - \beta(f_r'(\rho) + 1)) = 1. \tag{3.7}$$

Agora, considerando a última equação, definição em (3.6) e manipulações algébricas, obtemos

$$x_1 = \rho - \frac{h'(\rho)^T h(\rho)}{h'(\rho)^T h'(\rho)} = \rho - \frac{-\rho/\beta + \rho + f_r(\rho)}{-1/\beta + 1 + f_r'(\rho)} = -\rho\left(\frac{\beta(\rho f_r'(\rho) - f_r(\rho))}{\rho(1 - \beta(f_r'(\rho) + 1))}\right) = -\rho.$$

Novamente, considerando nossa exigência adicional em (3.7), definição em (3.6), obtemos

$$x_2 = -\rho - \frac{h'(-\rho)^T h(-\rho)}{h'(-\rho)^T h'(-\rho)} = -\rho - \frac{\rho/\beta - \rho - f_r(\rho)}{-1/\beta + 1 + f_r'(\rho)} = \rho\left(\frac{\beta(\rho f_r'(\rho) - f_r(\rho))}{\rho(1 - \beta(f_r'(\rho) + 1))}\right) = \rho.$$

Portanto, o método de Gauss-Newton para resolver (3.1) com $F = h$ e ponto inicial $x_0 = \rho$, produz o ciclo

$$x_0 = \rho, \quad x_1 = -\rho, \quad x_2 = \rho, \; \ldots \; .$$

Consequentemente, a sequência não converge, o que prova o lema. ∎

Lema 3.3 *Se* **h4** *vale, então o ponto* x_* *é o único ponto estacionário de* $F'(x)^*F(x)$ *em* $B(x_*, \sigma)$.

Demonstração: Suponha que $y \in B(x_*, \sigma)$ é outro ponto estacionário de $F'(x)^*F(x)$. Daí, como $F'(y)^*F(y) = 0$, temos que

$$y - x_* = y - x_* - [F'(x_*)^*F'(x_*)]^{-1}F'(y)^*F(y).$$

Considerando a última equação, $F'(x_*)^*F(x_*) = 0$ e algumas manipulações algébricas obtemos

$$y - x_* = [F'(x_*)^*F'(x_*)]^{-1}F'(x_*)^*[F'(x_*)(y - x_*) - F(y) + F(x_*)] \\ + [F'(x_*)^*F'(x_*)]^{-1}(F'(x_*)^* - F'(y)^*)F(y).$$

Combinando a última equação com definições de c, β e β_0 obtemos, após alguns cálculos que

$$\|y - x_*\| \leq \beta \int_0^1 \|F'(x_*) - F'(x_* + u(y - x_*))\| \|y - x_*\| du + c\beta_0 \|F'(x_*)^* - F'(y)^*\|.$$

Agora, usando (3.2) com $x = x_* + u(y - x_*)$ e $\tau = 0$ no primeiro termo e $x = y$ e $\tau = 0$ no segundo termo do lado direito da última desigualdade, obtemos

$$\|y - x_*\| \leq \beta \int_0^1 [f_r'(u\|y - x_*\|) - f_r'(0)] \|y - x_*\| du + c\beta_0 [f_r'(\|y - x_*\|) - f_r'(0)].$$

Avaliando a integral acima e usando **h1**, segue da última desigualdade que

$$\|y - x_*\| \leq \left[\frac{\beta(f_r(\|y - x_*\|) + \|y - x_*\|) + c\beta_0(f_r'(\|y - x_*\|) + 1)}{\|y - x_*\|} \right] \|y - x_*\|.$$

Como $0 < \|y - x_*\| < \sigma$, podemos aplicar a desigualdade (2.5) com $t = \|x - x_*\|$, para concluir que o termo entre colchetes na última desigualdade é menor que um. Daí, $\|y - x_*\| < \|y - x_*\|$, o que é uma contradição. Portanto, $y = x_*$. ∎

Observação 3.3 *Note que, no Lema acima usamos que a condição* (3.2) *é satisfeita apenas para* $\tau = 0$.

3.1.2 Prova do Teorema 3.1

Primeiro, note que a equação em (3.3) junto com (3.5) implica que a sequência $\{x_k\}$ satisfaz

$$x_{k+1} = G_F(x_k), \qquad k = 0, 1, \ldots . \tag{3.8}$$

Demonstração: Desde que $x_0 \in B(x_*, r)/\{x_*\}$, i.e., $0 < \|x_0 - x_*\| < r$, usando um argumento de indução o Lema 2.1 e a última desigualdade do Lema 3.1, é fácil ver que $\{x_k\}$ está bem definida e contida em $B(x_*, r)$.

Agora, iremos provar que $\{x_k\}$ converge para x_*. Como $\{x_k\}$ está bem definida e contida em $B(x_*, r)$, combinando o Lema 3.1 com (3.8), temos que

$$\begin{aligned}\|x_{k+1} - x_*\| \leq &\frac{\beta[f_r'(\|x_k - x_*\|)\|x_k - x_*\| - f_r(\|x_k - x_*\|)]}{\|x_k - x_*\|^2[1 - \beta(f_r'(\|x_k - x_*\|) + 1)]}\|x_k - x_*\|^2 \\ &+ \frac{\sqrt{2}c\beta^2[f_r'(\|x_k - x_*\|) + 1]}{\|x_k - x_*\|[1 - \beta(f_r'(\|x_k - x_*\|) + 1)]}\|x_k - x_*\|,\end{aligned}$$

para todo $k = 0, 1, \ldots$. Por outro lado, usando (3.8) e a última parte do Lema 3.1, concluímos que

$$\|x_k - x_*\| < \|x_0 - x_*\|, \qquad k = 1, 2 \ldots. \tag{3.9}$$

Daí, combinando as duas últimas desigualdades e a última parte da Proposição 2.3, temos que

$$\begin{aligned}\|x_{k+1} - x_*\| \leq &\frac{\beta[f_r'(\|x_0 - x_*\|)\|x_0 - x_*\| - f_r(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]}\|x_k - x_*\|^2 \\ &+ \frac{\sqrt{2}c\beta^2[f_r'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\|[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]}\|x_k - x_*\|, \qquad k = 0, 1, \ldots,\end{aligned}$$

que é a desigualdade (3.4) do teorema. Considerando (3.9) e a última desigualdade temos que

$$\|x_{k+1} - x_*\| \leq$$
$$\left[\frac{\beta[f_r'(\|x_0 - x_*\|)\|x_0 - x_*\| - f_r(\|x_0 - x_*\|)] + \sqrt{2}c\beta^2[f_r'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\|[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]}\right]\|x_k - x_*\|,$$

para todo $k = 0, 1, \ldots$. Assim, segue da desigualdade (2.3) da Proposição 2.4 com $t = \|x_0 - x_*\|$ que o termo entre colchetes na última desigualdade é menor que um. Portanto, $\{\|x_k - x_*\|\}$ converge para zero, i.e., $\{x_k\}$ converge para x_*. Os raios de convergência ótima e de unicidade do ponto estacionário foram provados nos Lemas 3.2 e 3.3, respectivamente. ∎

No caso de problemas de resíduo-zero, i.e., $c = 0$, o Teorema 3.1 torna-se:

Corolário 3.1 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continuamente diferenciável tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$, $R > 0$,*

$$\beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

Suponha que $F(x_*) = 0$, $F'(x_*)$ é injetiva e que existe uma função continuamente diferenciável $f_r : [0,\ R) \to \mathbb{R}$ tal que

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq f_r'(\|x - x_*\|) - f_r'(\tau\|x - x_*\|),$$

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

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

h2) f_r' *é convexa e estritamente crescente.*

Sejam as constantes positivas $\nu := \sup\{t \in [0, R) : \beta[f_r'(t) + 1] < 1\}$,

$$\rho := \sup\left\{t \in (0,\nu) : \frac{\beta[tf_r'(t) - f_r(t)]}{t[1 - \beta(f_r'(t)+1)]} < 1\right\}, \quad r := \min\{\kappa, \rho\}.$$

Então, o método de Gauss-Newton para resolver (3.1)*, com ponto inicial* $x_0 \in B(x_*, r)/\{x_*\}$

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

está bem definido, a sequência $\{x_k\}$ *está contida em* $B(x_*, r)$*, converge para* x_* *o qual é o único zero de* $F(x)$ *em* $B(x_*, \sigma)$*, onde* $\sigma := \sup\{t \in (0,\kappa) : \beta[f_r(t)/t + 1] < 1\}$ *e vale*

$$\|x_{k+1} - x_*\| \leq \frac{\beta[f_r'(\|x_0 - x_*\|)\|x_0 - x_*\| - f_r(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]}\|x_k - x_*\|^2, \quad k = 0, 1, \ldots.$$

Além disso, se $[\beta(\rho f_r'(\rho) - f_r(\rho))]/[\rho(1 - \beta(f_r'(\rho) + 1))] = 1$ *e* $\rho < \kappa$*, então* $r = \rho$ *é o melhor raio de convergência possível.*

Observação 3.4 *O Corolário 3.1 é uma generalização dos resultados do método de Newton para resolver o sistema de equações não-lineares* $F(x) = 0$*, pois quando* $F'(x_*)$ *é invertível o Corolário 3.1 é similar ao Teorema 2.1 em [15].*

3.2 Casos Especiais

Nesta seção, apresentaremos dois casos especiais do Teorema 3.1, a saber, resultados de convergência sob condição Lipschitz clássica e resultados de convergência sob condição de Smale para funções analíticas.

3.2.1 Resultados de Convergência Sob Condição Lipschitz Clássica

Nesta seção mostraremos um teorema correspondente ao Teorema 3.1, onde a condição geral (3.2) é substituida pela condição Lipschitz clássica (veja [6] e [7]).

Teorema 3.2 *Sejam* X, Y *espaços de Hilbert,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função continuamente diferenciável tal que* $F'(x)$ *tem imagem fechada* *em* Ω. *Tome* $x_* \in \Omega$,

$$c := \|F(x_*)\|, \qquad \beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}.$$

Suponha que $F'(x_*)^* F(x_*) = 0$, $F'(x_*)$ *é injetiva e existe um* $K > 0$ *tal que*

$$\sqrt{2}c\beta^2 K < 1, \qquad \|F'(x) - F'(y)\| \leq K\|x - y\|, \qquad \forall\, x, y \in B(x_*, \kappa).$$

Seja

$$r := \min\left\{\kappa,\ (2 - 2\sqrt{2}K\beta^2 c)/(3K\beta)\right\}.$$

Então, o método de Gauss-Newton para resolver (3.1), *com ponto inicial* $x_0 \in B(x_*, r)/\{x_*\}$

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

está bem definido, a sequência $\{x_k\}$ *está contida em* $B(x_*, r)$, *converge para* x_* *e vale*

$$\|x_{k+1} - x_*\| \leq \frac{\beta K}{2(1 - \beta K\|x_0 - x_*\|)}\|x_k - x_*\|^2 + \frac{\sqrt{2}c\beta^2 K}{1 - \beta K\|x_0 - x_*\|}\|x_k - x_*\|,$$

para todo $k = 0, 1, \ldots$. *Além disso, se* $(2 - 2\sqrt{2}K\beta^2 c)/(3K\beta) < \kappa$, *então* $r = (2 - 2\sqrt{2}K\beta^2 c)/(3K\beta)$ *é o melhor raio de convergência possível.*
Adicionalmente, se $c\beta_0 K < 1$, *então o ponto* x_* *é o único ponto estacionário de* $F(x)^* F(x)$ *em* $B(x_*, (2 - 2c\beta_0 K)/(\beta K))$, *onde* $\beta_0 = \|[F'(x_*)^T F'(x_*)]^{-1}\|$.

Demonstração: É imediato que F, x_* e $f_r : [0, \kappa) \to \mathbb{R}$ definida por $f_r(t) = Kt^2/2 - t$, satisfazem as hipóteses (3.2), **h1**, **h2**, **h3** e **h4** no Teorema 3.1. Neste caso, é fácil ver que as constantes ν e ρ como definidas no teorema 3.1, são

$$0 < \rho = (2 - 2\sqrt{2}K\beta^2 c)/(3\beta K)) \leq \nu = 1/\beta K.$$

Como consequência, $0 < r = \min\{\kappa, \rho\}$. Além disso, com simples cálculos podemos mostrar que

$$[\beta(\rho f_r'(\rho) - f_r(\rho)) + \sqrt{2}c\beta^2(f_r'(\rho) + 1)]/[\rho(1 - \beta(f_r'(\rho) + 1))] = 1,$$

e $[\beta(f_r(t)/t + 1) + c\beta_0(f_r'(t) + 1)/t] < 1$ para todo $t \in (0, (2 - 2c\beta_0 K)/(\beta K))$. Assim F, r, f_r e x_* satisfazem todas as hipóteses do Teorema 3.1. Portanto, as afirmações deste teorema seguem do Teorema 3.1. ∎

No caso de problemas de resíduo-zero, i.e., $c = 0$, o Teorema 3.2 torna-se:

Corolário 3.2 *Sejam* X, Y *espaços de Hilbert,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função continuamente diferenciável tal que* $F'(x)$ *tem imagem fechada* *em* Ω. *Tome* $x_* \in \Omega$,

$$\beta := \left\| F'(x_*)^\dagger \right\|, \qquad \kappa := \sup \{ t > 0 : B(x_*, t) \subset \Omega \}.$$

Suponha que $F(x_*) = 0$, $F'(x_*)$ *é injetiva e existe um* $K > 0$ *tal que*

$$\|F'(x) - F'(y)\| \leq K\|x - y\|, \qquad \forall\, x, y \in B(x_*, \kappa).$$

Seja

$$r := \min \{ \kappa,\, 2/(3K\beta) \}.$$

Então o método de Gauss-Newton para resolver (3.1)*, com ponto inicial* $x_0 \in B(x_*, r)/\{x_*\}$

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

está bem definido, a sequência $\{x_k\}$ *está contida em* $B(x_*, r)$ *e converge para* x_* *o qual é o único zero de* F *em* $B(x_*, 2/(\beta K))$ *e vale*

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

Além disso, se $2/(3K\beta) < \kappa$*, então* $r = 2/(3K\beta)$ *é o melhor raio de convergência possível.*

Observação 3.5 *Quando* $F'(x_*)$ *é invertível o Corolário 4.2 é similar ao Teorema 3.1 em [15].*

Para exemplificar, aplicaremos os resultados obtidos nesta seção em um exemplo numérico de Dedieu and Shub [14].

Exemplo 3.1 *Seja* $F : \mathbb{R} \to \mathbb{R}^2$ *tal que* $F(x) = (x, x^2 + c)^T$*, onde* $c \in \mathbb{R}$ *é dado. Daí,*

$$\min_{x \in \mathbb{R}} \|F(x)\|^2 = x^4 + (2c + 1)x^2 + c^2. \tag{3.10}$$

É fácil verificar que $F'(x) = (1, 2x)^T$, $\|F'(x) - F'(y)\| = 2|x - y|$*, para todo* $x, y \in \mathbb{R}$*, e* $x_* = 0$ *é um ponto estacionário de* $F(x)^T F(x)$*. Deste modo,* $K = 2$ *e* $\beta = 1$*. Portanto, concluímos do Teorema 3.2 que, se* $|c| < \sqrt{2}/4$*, então o método de Gauss-Newton, para resolver* (3.10)*, com ponto inicial* $x_0 \in B(0, (1 - 2\sqrt{2}|c|)/3)$ *está bem definido, a sequência* $\{x_k\}$ *está contida em* $B(0, (1 - 2\sqrt{2}|c|)/3)$*, converge para* 0 *e*

$$|x_{k+1}| \leq \frac{|x_k| + 2\sqrt{2}|c|}{1 - 2|x_0|} |x_k|, \qquad k = 0, 1, \ldots.$$

Além disso, se $c = 0$ *a sequência* $\{x_k\}$ *converge* Q*-quadraticamente para* $x_* = 0$.

3.2.2 Resultados de Convergência sob Condição de Smale

Nesta seção, mostraremos um teorema correspondente ao Teorema 3.1, onde a condição geral (3.2) é substituida pela condição de Smale (ver [14, 27]) para funções analíticas.

Teorema 3.3 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função analítica tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$,*

$$c := \|F(x_*)\|, \qquad \beta := \|F'(x_*)^\dagger\|, \qquad \kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}.$$

Suponha que $F'(x_*)^* F(x_*) = 0$, $F'(x_*)$ é injetiva e

$$\gamma := \sup_{n>1} \left\| \frac{F^{(n)}(x_*)}{n!} \right\|^{1/(n-1)} < +\infty, \qquad 2\sqrt{2} c \beta^2 \gamma < 1. \tag{3.11}$$

Sejam $a := (2 + 3\beta - \sqrt{2} c \beta^2 \gamma)$, $b := 4(1+\beta)(1 - 2\sqrt{2} c \beta^2 \gamma)$ e

$$r := \min\left\{\kappa, (a - \sqrt{a^2 - b})/(2\gamma(1+\beta))\right\}.$$

Então, o método de Gauss-Newton para resolver (3.1), com ponto inicial $x_0 \in B(x_*, r)/\{x_*\}$

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

está bem definido, a sequência $\{x_k\}$ está contida em $B(x_*, r)$, converge para x_* e vale

$$\begin{aligned} \|x_{k+1} - x_*\| \leq & \frac{\beta\gamma}{(1 - \gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} \|x_k - x_*\|^2 \\ & + \frac{\sqrt{2} c \beta^2 \gamma (2 - \gamma\|x_0 - x_*\|)}{(1 - \gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} \|x_k - x_*\|, \end{aligned}$$

para todo $k = 0, 1, \ldots$. Além disso, se $(a - \sqrt{a^2 - b})/(2\gamma(1+\beta)) < \kappa$, então $r = (a - \sqrt{a^2 - b})/(2\gamma(1+\beta))$ é o melhor raio de convergência possível.
Adicionalmente, se $2c\beta_0\gamma < 1$, então o ponto x_* é o único ponto estacionário de $F(x)^* F(x)$ em $B(x_*, \sigma)$, onde $\sigma := (\omega_1 - \sqrt{\omega_1^2 - \omega_2})/(2\gamma(1+\beta))$, $\omega_1 := (2 + \beta - c\beta_0)$, $\omega_2 := 4(1+\beta)(1 - 2c\beta_0\gamma)$, $\beta_0 := \|[F'(x_*)^T F'(x_*)]^{-1}\|$.

Observação 3.6 *Os resultados do Teorema acima apareceram pela primeira vez em [14], o qual também mostrou que neste caso um ponto estacionário de $F(x)^* F(x)$ é um minímo local estrito de (3.1).*

Os seguintes resultados serão necessários na prova do teorema acima.

Lema 3.4 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto, $F : \Omega \to Y$ uma função analítica. Suponha que $x_* \in \Omega$ e $B(x_*, 1/\gamma) \subset \Omega$, onde γ é definido em* (3.11). *Então, para todo $x \in B(x_*, 1/\gamma)$ vale*

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

Demonstração: Seja $x \in \Omega$. Como F é uma função analítica, temos que

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

Combinando (3.11) e a equação acima obtemos, após algumas manipulações algébricas que

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

Por outro lado, como $B(x_*, 1/\gamma) \subset \Omega$ temos $\gamma\|x - x_*\| < 1$. Daí, segue da Proposição 1.2 que

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

Combinando as duas últimas equações, obtemos o resultado desejado. ∎

O próximo resultado fornece uma condição mais fácil de ser verificada do que a condição (3.2), quando as funções em consideração são duas vezes continuamente diferenciáveis.

Lema 3.5 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto, $F : \Omega \to Y$ uma função duas vezes diferenciável. Se existe uma função $f : [0, R) \to \mathbb{R}$ duas vezes continuamente diferenciável tal que*

$$\|F''(x)\| \leqslant f''(\|x - x_*\|), \tag{3.12}$$

para todo $x \in \Omega$ e $\|x - x_*\| < R$, então F e f satisfaz (3.2).

Demonstração: Sejam $x \in B(x_*, \kappa)$ e $\tau \in [0, 1]$. É fácil ver que

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

Agora, como $\|x - x_*\| < \kappa$ e F satisfaz (3.12), obtemos da desigualdade acima que

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq \int_\tau^1 f''(t\|x - x_*\|)\|x - x_*\|dt.$$

Avaliando a integral acima, segue o resultado desejado. ∎

[Prova do Teorema 4.3]. *Considere a função real* $f_r : [0, 1/\gamma) \to \mathbb{R}$ *defina por*

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

É simples mostrar que f_r *é analítica e*

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

para $n \geq 2$. *Segue das equações acima que* f_r *satisfaz as hipóteses* **h1***,* **h2***,* **h3** *e* **h4** *no Teorema 3.1. Como* $f_r''(t) = (2\gamma)/(1 - \gamma t)^3$, *combinando os Lemas 3.4 e 3.5 concluimos que* F *e* f_r *satisfaz* (3.2) *com* $R = 1/\gamma$. *Neste caso, é fácil ver que as constantes* ν *e* ρ *como definidas no Teorema 3.1, são*

$$0 < \rho = (a - \sqrt{a^2 - b})/(2\gamma(1 + \beta)) < \nu = ((1 + \beta) - \sqrt{\beta(1 + \beta)})/(\gamma(1 + \beta)) < 1\gamma,$$

onde $a := (2 + 3\beta - \sqrt{2}c\beta^2\gamma)$, $b = 4(1 + \beta)(1 - 2\sqrt{2}c\beta^2\gamma)$. *Como consequência,* $0 < r = \min\{\kappa, \rho\}$. *Além disso, com simples cálculos podemos mostrar que*

$$[\beta(\rho f_r'(\rho) - f_r(\rho)) + \sqrt{2}c\beta^2(f_r'(\rho) + 1)]/[\rho(1 - \beta(f_r'(\rho) + 1))] = 1,$$

e $[\beta(f_r(t)/t + 1) + c\beta_0(f_r'(t) + 1)/t] < 1$ *para todo* $t \in (0, \sigma)$*, onde* $\sigma := (\omega_1 - \sqrt{\omega_1^2 - \omega_2})/(2\gamma(1 + \beta)$, $\omega_1 := (2 + \beta - c\beta_0)$, $\omega_2 := 4(1 + \beta)(1 - 2c\beta_0\gamma)$ *e* $\beta_0 := \|[F'(x_*)^T F'(x_*)]^{-1}\|$. *Assim,* F*,* σ*,* f_r *e* x_* *satisfazem todas as hipóteses do Teorema 3.1. Portanto, as afirmações deste teorema seguem do Teorema 3.1.*

No caso de problemas de resíduo-zero, i.e., $c = 0$, o Teorema 3.1 torna-se:

Corolário 3.3 *Sejam* X, Y *espaços de Hilbert,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função analítica tal que* $F'(x)$ *tem imagem fechada em* Ω. *Tome* $x_* \in \Omega$,

$$\beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}.$$

Suponha que $F(x_*) = 0$, $F'(x_*)$ *é injetiva e*

$$\gamma := \sup_{n>1} \left\| \frac{F^{(n)}(x_*)}{n!} \right\|^{1/(n-1)} < +\infty.$$

Seja

$$r := \min \left\{ \kappa, (2 + 3\beta - \sqrt{\beta(8 + 9\beta)})/(2\gamma(1 + \beta)) \right\}.$$

Então, o método de Gauss-Newton para resolver (3.1), *com ponto inicial* $x_0 \in B(x_*, r)/\{x_*\}$

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

está bem definido, a sequência $\{x_k\}$ *está contida em* $B(x_*, r)$ *e converge para* x_* *o qual é o único zero de* F *em* $B(x_*, 1/(\gamma(1 + \beta)))$ *e vale*

$$\|x_{k+1} - x_*\| \leq \frac{\beta\gamma}{(1 - \gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} \|x_k - x_*\|^2,$$

para todo $k = 0, 1, \ldots$. *Além disso,* $(2 + 3\beta - \sqrt{\beta(8 + 9\beta)})/(2\gamma(1 + \beta)) < \kappa$, *então* $r = ((2 + 3\beta - \sqrt{\beta(8 + 9\beta)})/(2\gamma(1 + \beta))$ *é o melhor raio de convergência possível.*

Observação 3.7 *Quando* $F'(x_*)$ *é invertível o Corolário 4.3 é similar ao Teorema 3.4 em [15].*

Capítulo 4

Análise Local do Método Quase Gauss-Newton Inexato

Neste capítulo, apresentaremos uma análise de convergência local do método inexato quase Gauss-Newton para problemas de mínimos quadrados não-lineares. Comprovaremos que sob certas condições a sequência gerada pelo método inexato quase Gauss-Newton está bem definida e converge linearmente para uma solução do problema em consideração. Encerraremos, aplicando os resultados obtidos para a condição Lipschitz clássica e condição de Smale para funções analíticas. Os resultados deste capítulo se encontram submetidos em períodico especializado e podem ser vistos em [31].

Inicialmente, sejam X e Y espaços de Hilbert real ou complexo, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continuamente diferenciável. Considere o seguinte *problema dos mínimos quadrados não-linear*

$$\min \|F(x)\|^2. \tag{4.1}$$

Se $F'(x)$ é injetiva e tem imagem fechada para todo $x \in \Omega$, o método quase Gauss-Newton inexato, onde controle residual relativo escalado é feito em cada iteração, é formalmente descrito como: Dado $x_0 \in \Omega$ defina

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

onde $B(x_k)$ é uma aproximação invertível para $F'(x_k)^*F'(x_k)$ e o resíduo r_k satisfaz

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^* F(x_k)\|,$$

para as dadas sequências $\{\theta_k\}$ e $\{P_k\}$ de números reais e matrizes invertíveis, respectivamente.

A seguir, faremos uma análise de convergência local do método definido acima.

4.1 Convergência do Método Quase Gauss-Newton Inexato

Nesta seção, provaremos a convergência local do método quase Gauss-Newton inexato para resolver (4.1). Com algumas exigências para a função F e admitindo que o ponto $x_* \in \Omega$ é um ponto estacionário de (4.1), mostraremos que tomando o ponto inicial x_0 numa vizinhança apropriada de x_*, a sequência gerada pelo método quase Gauss-Newton inexato está bem definida e converge linearmente para x_*. Estas afirmações se encontram no teorema:

Teorema 4.1 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continuamente diferenciável tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$, $R > 0$,*

$$c := \|F(x_*)\|, \qquad \beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\left\{t \in [0, R) : B(x_*, t) \subset \Omega\right\}.$$

Suponha que $F'(x_*)^* F(x_*) = 0$, $F'(x_*)$ é injetiva e que existe uma função continuamente diferenciável $f_r : [0,\ R) \to \mathbb{R}$ tal que

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq f_r'\left(\|x - x_*\|\right) - f_r'\left(\tau\|x - x_*\|\right), \tag{4.2}$$

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

h1) *$f_r(0) = 0$ and $f_r'(0) = -1$;*

h2) *f_r' é convexa e estritamente crescente;*

h3) *$\mu := \sqrt{2}\, c\, \beta^2 D^+ f_r'(0) < 1$.*

São dadas as constantes não-negativas $0 \leq \vartheta < 1$, $0 \leq \omega_2 < \omega_1$ tais que $\omega_1(\mu + \mu\vartheta + \vartheta) + \omega_2 < 1$. Sejam as constantes positivas $\nu := \sup\left\{t \in [0, R) : \beta[f_r'(t) + 1] < 1\right\}$,

$$\bar{\rho} := \sup\left\{t \in (0, \nu) : (1 + \vartheta)\omega_1 \beta \frac{t f_r'(t) - f_r(t) + \sqrt{2} c \beta [f_r'(t) + 1]}{t[1 - \beta(f_r'(t) + 1)]} + \omega_1\vartheta + \omega_2 < 1\right\},$$

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

Considere $B(x_k)$ uma aproximação invertível de $F'(x_k)^* F'(x_k)$ satisfazendo

$$\|B(x_k)^{-1} F'(x_k)^* F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1} F'(x_k)^* F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \ldots.$$

Então, o método quase Gauss-Newton inexato para resolver (4.1), com ponto inicial $x_0 \in B(x_*, r) \backslash \{x_*\}$

$$x_{k+1} = x_k + S_k, \qquad B(x_k) S_k = -F'(x_k)^* F(x_k) + r_k, \qquad k = 0, 1, \ldots, \tag{4.3}$$

onde o resíduo r_k, a matriz invertível P_k (precondicionando o sistema linear em (4.3)*) e o termo "forcing"θ_k satisfazem as seguintes condições*

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^* F(x_k)\|, \quad 0 \leq \theta_k cond(P_k F'(x_k)^* F'(x_k)) \leq \vartheta, \quad k = 0, 1, \ldots,$$

está bem definido, a sequência $\{x_k\}$ está contida em $B(x_*, r)$, converge para x_* e vale

$$\begin{aligned} \|x_{k+1} - x_*\| \leq\ & (1+\vartheta)\omega_1 \beta \frac{[f_r'(\|x_0 - x_*\|)\|x_0 - x_*\| - f_r(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]}\|x_k - x_*\|^2 \\ & + \left(\frac{(1+\vartheta)\omega_1 \sqrt{2} c \beta^2 [f_r'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\|[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]} + \omega_1 \vartheta + \omega_2 \right) \|x_k - x_*\|, \quad (4.4) \end{aligned}$$

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

Observação 4.1 *Com o objetivo de deixar o Teorema acima "auto-contido", enunciamos novamente, mas desta vez de maneira implícita, a definição de função majorante radial. Portanto, todas as afirmações que faremos neste capítulo envolvendo a função majorante radial f_r e suas relações com a função não-linear F foram provadas nas Seções 2.1.1 e 2.1.2, respectivamente.*

Observação 4.2 *Em particular, se $\vartheta = 0$ (neste caso $\theta_k \equiv 0$ e $r_k \equiv 0$) no Teorema 4.1, obtemos a convergência do método quase Gauss-Newton sob condição majorante, o qual para $\omega_1 = 1$ e $\omega_2 = 0$, i.e., $B(x_k) = F'(x_k)^* F'(x_k)$, foi obtido no Teorema 3.1 do Capítulo 2. Agora, no caso de resíduo zero, i.e., $c = 0$ e $F'(x_*)$ invertível, obtemos a convergência do método quase-Newton inexato sob condição majorante, o qual foi obtido por Ferreira, Gonçalves [16] no Teorema 4. Finalmente, se $c = \vartheta = \omega_2 = 0$, $\omega_1 = 1$ e $F'(x_*)$ é invertível no Teorema 4.1, obtemos a convergência do método de Newton sob condição majorante, o qual foi obtido por Ferreira [15] no Teorema 2.1.*

Como Ω é aberto é imediato concluir que $\kappa > 0$. Agora, segue das Proposições 2.2 e 2.5 que as constantes ν e $\bar{\rho}$ são positivas. Consequentemente, $r > 0$.

Para facilitar a demonstração do Teorema 4.1, apresentaremos a seguir um resultado auxiliar.

Lema 4.1 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continuamente diferenciável tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$, $R > 0$ e sejam c, β, κ como definidos no Teorema 4.1. Suponha que $F'(x_*)^* F(x_*) = 0$, $F'(x_*)$ é injetiva e que existe uma $f_r : [0,\ R) \to \mathbb{R}$ continuamente diferenciável satisfazendo* (4.2), **h1, h2** *e* **h3**. *Sejam μ, ϑ, ω_1, ω_2, ν, $\bar{\rho}$ e r como no*

Teorema 4.1. Assuma que $x \in B(x_*, r)\backslash\{x_*\}$, i.e., $0 < \|x - x_*\| < r$. Defina

$$x_+ = x + S, \qquad B(x)S = -F'(x)^*F(x) + r, \tag{4.5}$$

onde $B(x)$ é uma aproximação invertível de $F'(x)^*F'(x)$ satisfazendo

$$\|B(x)^{-1}F'(x)^*F'(x)\| \leq \omega_1, \qquad \|B(x)^{-1}F'(x)^*F'(x) - I\| \leq \omega_2, \tag{4.6}$$

e o termo "forcing"θ e o resíduo r satisfazem

$$\theta \mathit{cond}(PF'(x)^*F'(x)) \leq \vartheta, \qquad \|Pr\| \leq \theta\|PF'(x)^*F(x)\|, \tag{4.7}$$

com P uma matriz invertível (precondicionando o sistema linear em (4.5)). Então x_+ está bem definido e vale

$$\begin{aligned} \|x_+ - x_*\| \leq{}& (1+\vartheta)\omega_1\beta\frac{[f_r'(\|x - x_*\|)\|x - x_*\| - f_r(\|x - x_*\|)]}{\|x - x_*\|^2[1 - \beta(f_r'(\|x - x_*\|) + 1)]}\|x - x_*\|^2 \\ &+ \left(\frac{(1+\vartheta)\omega_1\sqrt{2}c\beta^2[f_r'(\|x - x_*\|) + 1]}{\|x - x_*\|[1 - \beta(f_r'(\|x - x_*\|) + 1)]} + \omega_1\vartheta + \omega_2\right)\|x - x_*\|, \end{aligned} \tag{4.8}$$

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

$$\|x_+ - x_*\| < \|x - x_*\|.$$

Demonstração: Primeiro, como $\|x - x_*\| < r$, o Lema 2.1 implica que $F'(x)^*F'(x)$ é invertível. Daí, como $B(x)$ é uma aproximação invertível de $F'(x)^*F'(x)$ satisfazendo (4.6), segue que x_+ está bem definida. Agora, combinando (4.5), $F'(x_*)^*F(x_*) = 0$ e algumas manipulações algébricas temos que

$$\begin{aligned} x_+ - x_* = x - x_* - B(x)^{-1}F'(x)^*\big(F(x) - F(x_*)\big) + B(x)^{-1}r \\ + B(x)^{-1}F'(x)^*F'(x)\left[F'(x_*)^\dagger F(x_*) - F'(x)^\dagger F(x_*)\right]. \end{aligned}$$

Combinando a igualdade acima com algumas manipulações algébricas, obtemos que

$$\begin{aligned} x_+ - x_* = {}& B(x)^{-1}F'(x)^*F'(x)F'(x)^\dagger\big(F(x_*) - [F(x) + F'(x)(x_* - x)]\big) + B(x)^{-1}r \\ &+ B(x)^{-1}\left(F'(x)^*F'(x) - B(x)\right)(x - x_*) \\ &+ B(x)^{-1}F'(x)^*F'(x)[F'(x_*)^\dagger F(x_*) - F'(x)^\dagger F(x_*)]. \end{aligned}$$

A última equação, juntamente com (2.8) e (4.6), implica que

$$\|x_+ - x_*\| \leq \omega_1 \|F'(x)^\dagger\| \|E_F(x, x_*)\| + \|B(x)^{-1} r\| \\ + \omega_2 \|x - x_*\| + \omega_1 \|F'(x)^\dagger - F'(x_*)^\dagger\| \|F(x_*)\|.$$

Por outro lado, usando (2.10), (4.6) e (4.7) obtemos, com simples cálculos que

$$\begin{aligned} &\|B(x)^{-1} r\| \leq \|B(x)^{-1} P^{-1}\| \|Pr\| \\ &\leq \theta \|B(x)^{-1} F'(x)^* F'(x)\| \|(PF'(x)^* F'(x))^{-1}\| \|PF'(x)^* F'(x)\| \|F'(x)^\dagger F(x)\| \\ &\leq \omega_1 \vartheta \|S_F(x)\|. \end{aligned}$$

Assim, obtemos imediatamente das duas últimas equações que

$$\|x_+ - x_*\| \leq \omega_1 \|F'(x)^\dagger\| \|E_F(x, x_*)\| + \omega_1 \vartheta \|S_F(x)\| \\ + \omega_2 \|x - x_*\| + \omega_1 \|F'(x)^\dagger - F'(x_*)^\dagger\| \|F(x_*)\|.$$

Combinando a última equação com os Lemas 2.1, 2.2 e 2.3, obtemos

$$\|x_+ - x_*\| \leq (1 + \vartheta) \beta \omega_1 \frac{e_{f_r}(\|x - x_*\|, 0) + \sqrt{2} c \beta (f_r'(\|x - x_*\|) + 1)}{1 - \beta (f_r'(\|x - x_*\|) + 1)} \\ + \omega_1 \vartheta \|x - x_*\| + \omega_2 \|x - x_*\|.$$

Agora, usando (2.9) e simples cálculos, concluímos da última inequação que

$$\|x_+ - x_*\| \leq \\ (1 + \vartheta) \beta \omega_1 \frac{f_r'(\|x - x_*\|) \|x - x_*\| - f_r(\|x - x_*\|) + \sqrt{2} c \beta (f_r'(\|x - x_*\|) + 1)}{1 - \beta (f_r'(\|x - x_*\|) + 1)} \\ + \omega_1 \vartheta \|x - x_*\| + \omega_2 \|x - x_*\|,$$

a qual é a inequação (4.8). Para concluir a prova, note que o lado da desigualdade em (4.8) é equivalente a

$$\left[(1 + \vartheta) \omega_1 \beta \frac{f_r'(\|x - x_*\|) \|x - x_*\| - f_r(\|x - x_*\|) + \sqrt{2} c \beta (f_r'(\|x - x_*\|) + 1)}{\|x - x_*\| [1 - \beta (f_r'(\|x - x_*\|) + 1)]} \right. \\ \left. + \omega_1 \vartheta + \omega_2 \right] \|x - x_*\|.$$

Por outro lado, como $x \in B(x_*, r)/\{x_*\}$, i.e., $0 < \|x - x_*\| < r \leq \bar{\rho}$ segue da Proposição 2.5 com $t = \|x - x_*\|$ que o termo entre colchetes na equação acima é menor que um. Portanto, a última inequação do lema segue. ∎

4.1.1 Prova do Teorema 4.1

Façamos agora a demonstração do Teorema 4.1.

Demonstração: Desde que $x_0 \in B(x_*, r)/\{x_*\}$, i.e., $0 < \|x_0 - x_*\| < r$, usando um argumento de indução o Lema 2.1 e a última desigualdade do Lema 4.1, é fácil ver que $\{x_k\}$ está bem definida e contida em $B(x_*, r)$.

Agora, iremos provar que $\{x_k\}$ converge para x_*. Como $\{x_k\}$ está bem definida e contida em $B(x_*, r)$, aplicando o Lema 4.1 com $x_+ = x_{k+1}$, $x = x_k$, $r = r_k$, $B(x) = B(x_k)$, $P = P_k$, e $\theta = \theta_k$, obtemos que

$$\begin{aligned}\|x_{k+1} - x_*\| \leq {} & (1+\vartheta)\omega_1\beta \frac{[f_r'(\|x_k - x_*\|)\|x_k - x_*\| - f_r(\|x_k - x_*\|)]}{\|x_k - x_*\|^2[1 - \beta(f_r'(\|x_k - x_*\|) + 1)]}\|x_k - x_*\|^2 \\ & + \left(\frac{(1+\vartheta)\omega_1\sqrt{2}c\beta^2[f_r'(\|x_k - x_*\|) + 1]}{\|x_k - x_*\|[1 - \beta(f_r'(\|x_k - x_*\|) + 1)]} + \omega_1\vartheta + \omega_2\right)\|x_k - x_*\|,\end{aligned}$$

para todo $k = 0, 1, \ldots$. Por outro lado, a última inequação do Lema 4.1 implica que

$$\|x_k - x_*\| < \|x_0 - x_*\|, \qquad k = 1, 2 \ldots. \tag{4.9}$$

Daí, combinando as duas últimas inequações e a última parte da Proposição 2.3, obtemos que

$$\begin{aligned}\|x_{k+1} - x_*\| \leq {} & (1+\vartheta)\omega_1\beta \frac{[f_r'(\|x_0 - x_*\|)\|x_0 - x_*\| - f_r(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]}\|x_k - x_*\|^2 \\ & + \left(\frac{(1+\vartheta)\omega_1\sqrt{2}c\beta^2[f_r'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\|[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]} + \omega_1\vartheta + \omega_2\right)\|x_k - x_*\|,\end{aligned}$$

para todo $k = 0, 1, \ldots$, a qual é equivalente a inequação em (4.4). Agora, usando (4.9) e a última inequação temos que

$$\begin{aligned}&\|x_{k+1} - x_*\| \leq \\ &\Bigg[(1+\vartheta)\omega_1\beta \frac{f_r'(\|x_0 - x_*\|)\|x_0 - x_*\| - f_r(\|x_0 - x_*\|) + \sqrt{2}c\beta(f_r'(\|x_0 - x_*\|) + 1)}{\|x_0 - x_*\|[1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]} \\ &\qquad\qquad + \omega_1\vartheta + \omega_2\Bigg]\|x_k - x_*\|,\end{aligned}$$

para todo $k = 0, 1, \ldots$. Aplicando a Proposição 2.5 com $t = \|x_0 - x_*\|$ é imediato concluir da última inequação que $\{\|x_k - x_*\|\}$ converge para zero. Então, $\{x_k\}$ converge para x_*. ∎

Para o importante caso $\vartheta = 0$, ou seja, método quase Gauss-Newton sob condição majorante, o Teorema 4.1 torna-se:

Corolário 4.1 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continuamente diferenciável tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$, $R > 0$,*

$$c := \|F(x_*)\|, \qquad \beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\left\{t \in [0, R) : B(x_*, t) \subset \Omega\right\}.$$

Suponha que $F'(x_*)^* F(x_*) = 0$, $F'(x_*)$ é injetiva e que existe uma função continuamente diferenciável $f_r : [0,\ R) \to \mathbb{R}$ tal que

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq f_r'\left(\|x - x_*\|\right) - f_r'\left(\tau\|x - x_*\|\right),$$

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

h1) *$f_r(0) = 0$ and $f_r'(0) = -1$;*

h2) *f_r' é convexa e estritamente crescente;*

h3) *$\mu := \sqrt{2}\, c\, \beta^2 D^+ f_r'(0) < 1$.*

São dadas as constantes não-negativas $0 \leq \omega_2 < \omega_1$ tais que $\omega_1 \mu + \omega_2 < 1$. Sejam as constantes positivas $\nu := \sup\left\{t \in [0, R) : \beta[f_r'(t) + 1] < 1\right\}$,

$$\bar{\rho} := \sup\left\{t \in (0, \nu) : \omega_1 \beta \frac{t f_r'(t) - f_r(t) + \sqrt{2} c \beta [f_r'(t) + 1]}{t[1 - \beta(f_r'(t) + 1)]} + \omega_2 < 1\right\},$$

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

Considere $B(x_k)$ uma aproximação invertível de $F'(x_k)^* F'(x_k)$ satisfazendo

$$\|B(x_k)^{-1} F'(x_k)^* F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1} F'(x_k)^* F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \ldots.$$

Então, o método quase Gauss-Newton para resolver (4.1), com ponto inicial $x_0 \in B(x_*, r) \backslash \{x_*\}$

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

está bem definido, a sequência $\{x_k\}$ está contida em $B(x_*, r)$, converge para x_* e vale

$$\begin{aligned}\|x_{k+1} - x_*\| \leq{}& \omega_1 \beta \frac{[f_r'(\|x_0 - x_*\|)\|x_0 - x_*\| - f_r(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2 [1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]} \|x_k - x_*\|^2 \\ &+ \left(\frac{\omega_1 \sqrt{2} c \beta^2 [f_r'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\| [1 - \beta(f_r'(\|x_0 - x_*\|) + 1)]} + \omega_2\right) \|x_k - x_*\|, \quad k = 0, 1, \ldots.\end{aligned}$$

Observação 4.3 *Apesar do fato que o corolário acima é um caso particular do Teorema 4.1, os resultados obtidos nele estendem os resultados de Chen e Li em [11], pois os resultados obtidos em [11] são somente para o caso* $c = 0$.

4.2 Casos Especiais

Nesta seção, apresentaremos dois casos especiais do Teorema 4.1, a saber, resultados de convergência sob condição Lipschitz clássica e resultados de convergência sob condição de Smale para funções analíticas.

4.2.1 Resultados de Convergência Sob Condição Lipschitz Clássica

Nesta seção mostraremos um teorema correspondente ao Teorema 4.1, onde a condição geral (3.2) é substituida pela condição Lipschitz clássica (ver [6] e [7]).

Teorema 4.2 *Sejam* X, Y *espaços de Hilbert,* $\Omega \subseteq X$ *um conjunto aberto e* $F : \Omega \to Y$ *uma função continuamente diferenciável tal que* $F'(x)$ *tem imagem fechada em* Ω. *Tome* $x_* \in \Omega$,

$$c := \|F(x_*)\|, \qquad \beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\left\{t > 0 : B(x_*, t) \subset \Omega\right\}.$$

Suponha que $F'(x_*)^*F(x_*) = 0$, $F'(x_*)$ *é injetiva e existe um* $K > 0$ *tal que*

$$\mu := \sqrt{2}c\beta^2 K < 1, \qquad \|F'(x) - F'(y)\| \leq K\|x - y\|, \qquad \forall\, x, y \in B(x_*, \kappa).$$

São dadas as constantes não-negativas $0 \leq \vartheta < 1$, $0 \leq \omega_2 < \omega_1$ *tais que* $\omega_1(\mu + \mu\vartheta + \vartheta) + \omega_2 < 1$. *Seja*

$$r := \min\left\{\kappa, \frac{2(1 - \omega_1\vartheta - \omega_2) - 2\sqrt{2}cK\beta^2\omega_1(1 + \vartheta)}{\beta K\,(2 + \omega_1 - \vartheta\omega_1 - 2\omega_2)}\right\}.$$

Considere $B(x_k)$ *uma aproximação invertível de* $F'(x_k)^*F'(x_k)$ *satisfazendo*

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \ldots,$$

Então, o método quase Gauss-Newton inexato para resolver (4.1), *com ponto inicial* $x_0 \in B(x_*, r)\backslash\{x_*\}$

$$x_{k+1} = x_k + S_k, \qquad B(x_k)S_k = -F'(x_k)^*F(x_k) + r_k, \qquad k = 0, 1, \ldots, \tag{4.10}$$

onde o resíduo r_k, a matriz invertível P_k (precondicionando o sistema linear em (4.10)*) e o termo "forcing"θ_k satisfazem as seguintes condições*

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^* F(x_k)\|, \quad 0 \leq \theta_k cond(P_k F'(x_k)^* F'(x_k)) \leq \vartheta, \quad k = 0, 1, \ldots,$$

está bem definido, a sequência $\{x_k\}$ está contida em $B(x_*, r)$, converge para x_* e vale

$$\begin{aligned} \|x_{k+1} - x_*\| \leq & \frac{(1+\vartheta)\beta\omega_1 K}{2(1 - \beta K\|x_0 - x_*\|)}\|x_k - x_*\|^2 \\ & + \left(\frac{(1+\vartheta)\omega_1\sqrt{2}c\beta^2 K}{1 - \beta K\|x_0 - x_*\|} + \omega_1\vartheta + \omega_2 \right) \|x_k - x_*\|, \end{aligned}$$

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

Demonstração: É imediato que F, x_* e $f_r : [0, \kappa) \to \mathbb{R}$ definida por $f_r(t) = Kt^2/2 - t$, satisfazem as hipóteses (4.2), **h1**, **h2**, **h3** e **h4** no Teorema 4.1. Neste caso, é fácil ver que as constantes ν e $\bar{\rho}$ como definidas no teorema 4.1, são

$$0 < \bar{\rho} = \frac{2(1 - \omega_1\vartheta - \omega_2) - 2\sqrt{2}cK\beta^2\omega_1(1+\vartheta)}{\beta K\,(2 + \omega_1 - \vartheta\omega_1 - 2\omega_2)} \leq \nu = 1/\beta K,$$

Como consequência, $0 < r = \min\{\kappa, \bar{\rho}\}$. Assim F, r, f_r e x_* satisfazem todas as hipóteses do Teorema 4.1. Portanto, as afirmações deste teorema seguem do Teorema 4.1. ∎

Para o caso $\vartheta = 0$, o Teorema 4.2 torna-se:

Corolário 4.2 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função continuamente diferenciável tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$, $R > 0$,*

$$c := \|F(x_*)\|, \qquad \beta := \left\|F'(x_*)^\dagger\right\|, \qquad \kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

Suponha que $F'(x_*)^* F(x_*) = 0$, $F'(x_*)$ é injetiva e existe um $K > 0$ tal que

$$\mu := \sqrt{2}c\beta^2 K < 1, \qquad \|F'(x) - F'(y)\| \leq K\|x - y\|, \qquad \forall\, x, y \in B(x_*, \kappa).$$

São dadas as constantes não-negativas $0 \leq \omega_2 < \omega_1$ tais que $\omega_1\mu + \omega_2 < 1$. Seja

$$r := \min\left\{\kappa, \frac{2(1-\omega_2) - 2\sqrt{2}cK\beta^2\omega_1}{\beta K\,(2 + \omega_1 - 2\omega_2)}\right\}.$$

Considere $B(x_k)$ uma aproximação invertível de $F'(x_k)^*F'(x_k)$ satisfazendo

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \ldots.$$

Então, o método quase Gauss-Newton inexato para resolver (4.1), *com ponto inicial $x_0 \in B(x_*, r)\backslash\{x_*\}$*

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

está bem definido, a sequência $\{x_k\}$ está contida em $B(x_*, r)$, converge para x_* e vale

$$\|x_{k+1} - x_*\| \leq \frac{\beta\omega_1 K}{2(1 - \beta K\|x_0 - x_*\|)}\|x_k - x_*\|^2 + \left(\frac{\omega_1\sqrt{2}c\beta^2 K}{1 - \beta K\|x_0 - x_*\|} + \omega_2\right)\|x_k - x_*\|,$$

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

Se ainda $c = 0$ no corolário acima, obtemos o Corolário 6.1 de [11].

4.2.2 Resultados de Convergência sob Condição de Smale

Nesta seção, mostraremos um teorema correspondente ao Teorema 4.1, onde a condição geral (4.2) é substituida pela condição de Smale (ver [14, 27]) para funções analíticas.

Teorema 4.3 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função analítica tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$,*

$$c := \|F(x_*)\|, \qquad \beta := \|F'(x_*)^\dagger\|, \qquad \kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}.$$

Suponha que $F'(x_*)^*F(x_*) = 0$, $F'(x_*)$ é injetiva e

$$\gamma := \sup_{n>1}\left\|\frac{F^{(n)}(x_*)}{n!}\right\|^{1/(n-1)} < +\infty, \qquad\qquad \mu := 2\sqrt{2}c\beta^2\gamma < 1. \tag{4.11}$$

São dadas as constantes não-negativas $0 \leq \vartheta < 1$, $0 \leq \omega_2 < \omega_1$ tais que $\omega_1(\mu + \mu\vartheta + \vartheta) + \omega_2 < 1$. Sejam $a := (1 - \vartheta\omega_1 - \omega_2)$, $b := (1 + \vartheta)\omega_1\beta$, $\bar{a} := b + 2a(1 + \beta) - \sqrt{2}\gamma\beta bc$,

$$r := \min\left\{\kappa, \frac{\bar{a} - \sqrt{\bar{a}^2 - 4a(1 + \beta)(a - 2\sqrt{2}c\beta b\gamma)}}{2a\gamma(1 + \beta)}\right\}.$$

Considere $B(x_k)$ uma aproximação invertível de $F'(x_k)^*F'(x_k)$ satisfazendo

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \ldots.$$

Então, o método quase Gauss-Newton inexato para resolver (4.1), com ponto inicial $x_0 \in B(x_*, r)\backslash\{x_*\}$

$$x_{k+1} = x_k + S_k, \qquad B(x_k)S_k = -F'(x_k)^*F(x_k) + r_k, \qquad k = 0, 1, \ldots, \quad (4.12)$$

onde o resíduo r_k*, a matriz invertível* P_k *(precondicionando o sistema linear em* (4.12)*) e o termo "forcing"*θ_k *satisfazem as seguintes condições*

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^* F(x_k)\|, \quad 0 \leq \theta_k cond(P_k F'(x_k)^* F'(x_k)) \leq \vartheta, \quad k = 0, 1, \ldots,$$

está bem definido, a sequência $\{x_k\}$ *está contida em* $B(x_*, r)$*, converge para* x_* *e vale*

$$\|x_{k+1} - x_*\| \leq \frac{(1+\vartheta)\omega_1\beta\gamma}{(1-\gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)}\|x_k - x_*\|^2$$
$$+ \left(\frac{(1+\vartheta)\omega_1\sqrt{2}c\beta^2\gamma(2-\gamma\|x_0 - x_*\|)}{(1-\gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} + \omega_1\vartheta + \omega_2\right)\|x_k - x_*\|,$$

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

Para provar o teorema acima, utilizaremos os Lemas 3.4 e 3.5 da Seção 3.2.2 do Capítulo 3

[Prova do Teorema 4.3]. *Considere a função real* $f_r : [0, 1/\gamma) \to \mathbb{R}$ *defina por*

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

É simples mostrar que f_r *é analítica e*

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

para $n \geq 2$. *Segue das equações acima que* f_r *satisfaz as hipóteses* **h1***,* **h2***, e* **h3** *e* **h4** *no Teorema 4.1. Como* $f_r''(t) = (2\gamma)/(1-\gamma t)^3$*, combinando os Lemas 3.4 e 3.5 concluimos que* F *e* f_r *satisfaz* (4.2) *com* $R = 1/\gamma$*. Neste caso, é fácil ver que as constantes* ν *e* $\bar{\rho}$ *como definidas no Teorema 4.1, são*

$$0 < \bar{\rho} = \frac{\bar{a} - \sqrt{\bar{a}^2 - 4a(1+\beta)(a - 2\sqrt{2}c\beta b\gamma)}}{2a\gamma(1+\beta)} < \nu = ((1+\beta) - \sqrt{\beta(1+\beta)})/(\gamma(1+\beta)) < 1/\gamma.$$

Como consequência, $0 < r = \min\{\kappa, \bar{\rho}\}$. *Assim,* F*,* $\bar{\rho}$*,* f_r *e* x_* *satisfazem todas as hipóteses do Teorema 4.1. Portanto, as afirmações deste teorema seguem do Teorema 4.1.*

Para o caso $\vartheta = 0$, o Teorema 4.3 torna-se:

Corolário 4.3 *Sejam X, Y espaços de Hilbert, $\Omega \subseteq X$ um conjunto aberto e $F : \Omega \to Y$ uma função analítica tal que $F'(x)$ tem imagem fechada em Ω. Tome $x_* \in \Omega$,*

$$c := \|F(x_*)\|, \qquad \beta := \|F'(x_*)^\dagger\|, \qquad \kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}.$$

Suponha que $F'(x_*)^* F(x_*) = 0$, $F'(x_*)$ é injetiva e

$$\gamma := \sup_{n>1} \left\| \frac{F^{(n)}(x_*)}{n!} \right\|^{1/(n-1)} < +\infty, \qquad \mu := 2\sqrt{2}c\beta^2\gamma < 1.$$

São dadas as constantes não-negativas $0 \leq \omega_2 < \omega_1$ tais que $\omega_1\alpha + \omega_2 < 1$. Sejam $\bar{a} := \omega_1\beta + 2(1-\omega_2)(1+\beta) - \sqrt{2}\gamma\beta^2\omega_1 c$ e

$$r := \min\left\{\kappa, \frac{\bar{a} - \sqrt{\bar{a}^2 - 4(1-\omega_2)(1+\beta)(1-\omega_2 - 2\sqrt{2}c\beta^2\omega_1\gamma)}}{2(1-\omega_2)\gamma(1+\beta)}\right\}.$$

Considere $B(x_k)$ uma aproximação invertível de $F'(x_k)^*F'(x_k)$ satisfazendo

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \ldots,$$

Então, o método quase Gauss-Newton inexato para resolver (4.1), com ponto inicial $x_0 \in B(x_*, r)\backslash\{x_*\}$

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

está bem definido, a sequência $\{x_k\}$ está contida em $B(x_*, r)$, converge para x_* e vale

$$\begin{aligned}\|x_{k+1} - x_*\| \leq &\frac{\omega_1\beta\gamma}{(1-\gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)}\|x_k - x_*\|^2 \\ &+ \left(\frac{\omega_1\sqrt{2}c\beta^2\gamma(2-\gamma\|x_0 - x_*\|)}{(1-\gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} + \omega_2\right)\|x_k - x_*\|,\end{aligned}$$

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

Se ainda $c = \omega_2 = 0$ e $\omega_1 = 1$ no corolário acima, obtemos o Exemplo 1 de [11].

Capítulo 5

Análise Semi-Local do Método Gauss-Newton

Neste capítulo, apresentaremos uma análise de convergência semi-local para a sequência gerada pelo algoritmo de Gauss-Newton para resolver problemas de otimização de composição convexa. Comprovaremos que sob certas condições a sequência gerada pelo algoritmo de Gauss-Newton está bem definida e converge linearmente para uma solução do problema em consideração. Além disso, aplicaremos os resultados obtidos para o caso onde o ponto inicial é um ponto regular e o caso onde o ponto inicial satisfaz a condição de Robinson. Também daremos resultados de convergência sob condição Lipschitz clássica e condição de Smale para funções analíticas

Inicialmente, considere o seguinte *problema de otimização de composição convexa*

$$\min\ h(F(x)), \tag{5.1}$$

onde $h : \mathbb{R}^m \to \mathbb{R}$ é uma função convexa de valores reais e $F : \mathbb{R}^n \to \mathbb{R}^m$ é continuamente diferenciável.

É fácil ver que o estudo de (5.1) está relacionado com o problema de inclusão convexa

$$F(x) \in C := \{z \in \mathbb{R}^m : h(z) \leq h(y),\, \forall y \in \mathbb{R}^m\}. \tag{5.2}$$

Em ordem para afirmar o algoritmo de Gauss-Newton, para resolver (5.1), precisamos da seguinte definição: Dados $\Delta \in (0, +\infty)$ e $x \in \mathbb{R}^n$ defina

$$D_\Delta(x) := \operatorname{argmin}\left\{h(F(x) + F'(x)d) :\ d \in \mathbb{R}^n,\ \|d\| \leq \Delta\right\}, \tag{5.3}$$

isto é, $D_\Delta(x)$ é o conjunto solução do seguinte problema

$$\min\left\{h(F(x) + F'(x)d) : d \in \mathbb{R}^n,\ \|d\| \leq \Delta\right\}. \tag{5.4}$$

Dados $\Delta \in (0, +\infty]$, $\eta \in [1, +\infty)$ e um ponto $x_0 \in \mathbb{R}^n$, *o algoritmo tipo Gauss-Newton* associado a (Δ, η, x_0) como definido em [3] (ver também, [4, 5]) é como segue:

Algoritmo 5.1

INICIALIZAÇÃO. *Dados* $\Delta \in (0, +\infty]$, $\eta \in [1, +\infty)$ *e* $x_0 \in \mathbb{R}^n$. *Faça* $k = 0$.
CRITÉRIO DE PARADA. *Calcule* $D_\Delta(x_k)$. *Se* $0 \in D_\Delta(x_k)$, *Pare. Caso contrário.*
PASSO ITERATIVO . *Calcule* d_k *satisfazendo*

$$d_k \in D_\Delta(x_k), \qquad \|d_k\| \leq \eta d(0, D_\Delta(x_k)),$$

e faça

$$x_{k+1} = x_k + d_k,$$

$k = k + 1$ *e vá para* CRITÉRIO DE PARADA.

Observe que o conjunto $D_\Delta(x)$ é não-vazio para todo $x \in \mathbb{R}^n$, pois (5.4) é um problema de otimização convexo sobre um conjunto compacto. Portanto, a sequência gerada pelo Algoritmo 5.1 está bem definida.

A seguir, faremos uma análise de convergência semi-local do método definido acima.

5.1 Convergência Semi-Local do Método de Gauss-Newton

Nesta seção, provaremos a convergência semi-local da sequência gerada pelo **Algoritmo 5.1** para resolver (5.1). Sob as hipóteses que o ponto inicial é um ponto quase-regular para a inclusão (5.2) e a função não-linear F satisfaz nossa condição majorante esférica, provaremos que a sequência gerada pelo **Algoritmo 5.1** converge para um ponto x_* tal que $F(x_*) \in C$, em particular, x_* resolve (5.1). Estas afirmações se encontram no teorema:

Teorema 5.1 *Seja* $F : \mathbb{R}^n \to \mathbb{R}^m$ *uma função continuamente diferenciável. Suponha que existem* $R > 0$, $x_0 \in \mathbb{R}^n$ *e uma função majorante esférica* $f_e : [0,\ R) \to \mathbb{R}$ *para a função* F *em* $B(x_0, R)$. *Dadas as constantes* $\alpha > 0$ *e* $\xi > 0$, *considere a função auxiliar* $f_{\xi,\alpha} : [0, R) \to \mathbb{R}$,

$$f_{\xi,\alpha}(t) = \xi + (\alpha - 1)t + \alpha f_e(t).$$

Se $f_{\xi,\alpha}$ *satisfaz* **a3**, *i.e.,* t_* *é o menor zero de* $f_{\xi,\alpha}$, *então a sequência gerada pelo* *método de Newton para resolver $f_{\xi,\alpha}(t) = 0$, com ponto inicial $t_0 = 0$,*

$$t_{k+1} = t_k - f'_{\xi,\alpha}(t_k)^{-1} f_{\xi,\alpha}(t_k), \quad k = 0, 1, \ldots, \tag{5.5}$$

está bem definida, $\{t_k\}$ é estritamente crescente, está contida em $[0, t_*)$, e converge Q-linearmente para t_*. Sejam $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$ e $h : \mathbb{R}^m \to \mathbb{R}$ uma função convexa de valores-reais com conjunto de mínimos C não-vazio. Assuma que x_0 é um ponto quase-regular da inclusão

$$F(x) \in C,$$

com o raio quase-regular r_{x_0} e função limitante quase-regular β_{x_0} como definidos em (1.5) e (1.6), respectivamente. Se $d(F(x_0), C) > 0$, $t_* \le r_{x_0}$,

$$\Delta \ge \xi \ge \eta \beta_{x_0}(0) d(F(x_0), C), \tag{5.6}$$

$$\alpha \ge \sup \left\{ \frac{\eta \beta_{x_0}(t)}{\eta \beta_{x_0}(t)(f'_e(t) + 1) + 1} : \xi \le t < t_* \right\},$$

então a sequência gerada pelo **Algoritmo 5.1**, *denotada por $\{x_k\}$, está contida em $B(x_0, t_*)$,*

$$F(x_k) + F'(x_k)(x_{k+1} - x_k) \in C, \quad k = 0, 1, \ldots, \tag{5.7}$$

satisfaz as inequações

$$\|x_{k+1} - x_k\| \le |t_{k+1} - t_k|, \quad k = 0, 1 \ldots, \tag{5.8}$$

$$\|x_{k+1} - x_k\| \le \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2} \|x_k - x_{k-1}\|^2, \quad k = 1, 2, \ldots, \tag{5.9}$$

converge para um ponto $x_* \in B[x_0, t_*]$ tal que $F(x_*) \in C$,

$$\|x_* - x_k\| \le |t_* - t_k|, \qquad k = 0, 1, \ldots \tag{5.10}$$

e a convergência é R-linear. Se, adicionalmente, $f_{\xi,\alpha}$ satisfaz **a4** *então as sequências $\{t_k\}$ e $\{x_k\}$ convergem Q-quadraticamente e R-quadraticamente para t_* and x_*, respectivamente.*

Observação 5.1 *Observamos que todas as afirmações que faremos neste capítulo envolvendo a função majorante esférica f_e e a função auxiliar $f_{\xi,\alpha}$ foram provadas na Seção 2.2. Além disso, todas as afirmações feitas no Teorema 5.1 para a sequência t_k foram provadas no Corolário 2.1.*

Observação 5.2 *Se*

$$\alpha > \bar{\alpha} := \sup\left\{\frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)(f'_e(t)+1)+1} : \xi \leq t < t_*\right\},$$

então a sequência $\{x_k\}$ *converge R-quadraticamente para* x_*. *Para provar essa afirmação, note que do item* **iii** *da Proposição 2.13, temos que* $\bar{t}_* < t_*$. *Daí, usando que* $f'_{\xi,\bar{\alpha}}$ *é estritamente crescente e item* **ii** *da Proposição 2.13, obtemos que*

$$f'_{\xi,\bar{\alpha}}(\bar{t}_*) < f'_{\xi,\bar{\alpha}}(t_*) < f'_{\xi,\alpha}(t_*),$$

o qual, combinado com Proposição 2.8 implica que $f'_{\xi,\bar{\alpha}}(t_*) < 0$. *Portanto, a afirmação segue se* $f_{\xi,\alpha}$ *é trocado por* $f_{\xi,\bar{\alpha}}$ *no Teorema 5.1.*

Para facilitar a demonstração do Teorema 5.1, apresentaremos a seguir uma seção contendo três lemas auxiliares.

5.1.1 Resultados auxiliares

Nesta subseção nosso objetivo é demonstrar três lemas que auxiliarão na prova de convergência do Teorema 5.1.

Como vimos na introdução deste capítulo, $D_\Delta(x) \neq \emptyset$ para todo $x \in \mathbb{R}^n$, portanto a sequência $\{x_k\}$ está bem definida. Mas isso não nos capacita a provar a convergência da sequência $\{x_k\}$ para algum ponto $x_* \in \mathbb{R}^n$ tal que $F(x_*) \in C$, pois não temos nenhuma relação entre o conjunto de direções $D_\Delta(x)$ e o conjunto de soluções da inclusão

$$F(x) + F'(x)d \in C, \qquad \|d\| \leq \Delta.$$

Agora, se provarmos que $D_\Delta(x) \subset D_C(x)$ para alguns pontos, podemos usar os resultados de regularidade para relacionar os conjuntos mencionados acima.

Primeiro, definimos alguns subconjuntos em $B(x_0, t_*)$ nos quais, como iremos provar, a inclusão desejada é válida para todos os pontos nestes conjuntos.

$$K(t) := \left\{x \in \mathbb{R}^n \,:\, \|x - x_0\| \leq t,\ \eta d(0, D_C(x)) \leqslant -\frac{f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)}\right\}, \quad t \in [0, t_*)\,, \tag{5.11}$$

$$K := \bigcup_{t\in[0,t_*)} K(t). \tag{5.12}$$

Em (5.11), $0 \leqslant t < t_*$, consequentemente, $f'_{\xi,\alpha}(t) \neq 0$ (Proposição 2.8). Portanto, as definições são consistentes.

Proposição 5.1 *Se* $x \in K$, *então*

$$D_\Delta(x) = \{d \in \mathbb{R}^n : F(x) + F'(x)d \in C, \|d\| \leq \Delta\} \subset D_C(x)$$

e

$$d(0, D_\Delta(x)) = d(0, D_C(x)).$$

Demonstração: Pela Proposição 1.3 é suficiente mostrar que $D_C(x) \neq \emptyset$ e $d(0, D_C(x)) \leq \Delta$. Seja $x \in K$, então $x \in K(t)$ para algum $t \in [0, t_*)$, o que implica que $x \in B(x_0, t_*)$. Como $t_* \leq r_{x_0}$ e x_0 é um ponto quase-regular, segue da Definição 1.3 e definição de raio quase-regular em (1.5) que $D_C(x) \neq \emptyset$. Por hipóteses $\eta \geq 1$ e $\xi \leq \Delta$. Daí, como $x \in K(t)$, usando a Definição em (5.11) Proposições 2.11 e 2.7, obtemos

$$d(0, D_C(x)) \leqslant \eta d(0, D_C(x)) \leqslant -\frac{f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)} \leqslant -\frac{f_{\xi,\alpha}(0)}{f'_{\xi,\alpha}(0)} = \xi \leqslant \Delta,$$

o que prova o resultado desejado ∎

Para cada $x \in \mathbb{R}^n$, definimos o conjunto $\bar{D}_\Delta(x)$ como

$$\bar{D}_\Delta(x) := \{d \in D_\Delta(x);\ \|d\| \leq \eta d(0, D_\Delta(x))\}. \tag{5.13}$$

Como $D_\Delta(x) \neq \emptyset$ para todo $x \in \mathbb{R}^n$, segue que $\bar{D}_\Delta(x) \neq \emptyset$ para todo $x \in \mathbb{R}^n$, assim temos a boa definição da multifunção iteração de Gauss-Newton $\bar{G}_F$

$$\begin{array}{rcl} \bar{G}_F : B(x_0, t_*) & \to & P(\mathbb{R}^n) \\ x & \mapsto & x + \bar{D}_\Delta(x). \end{array} \tag{5.14}$$

Iremos provar que a multifunção iteração de Gauss-Newton está bem definida nos subconjuntos definidos em (5.11), mas antes precisamos do seguinte resultado técnico.

Lema 5.1 *Para cada $t \in [0, t_*)$, $x \in K(t)$ e $y \in G_F(x)$ as seguintes desigualdades são válidas:*

i) $\|y - x\| \leq n_{f_{\xi,\alpha}}(t) - t$;

ii) $\|y - x_0\| \leq n_{f_{\xi,\alpha}}(t) < t_*$;

iii) $\eta d(0, D_C(y)) \leq -\dfrac{f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t))}{f'_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t))} \left(\dfrac{\|y - x\|}{n_{f_{\xi,\alpha}}(t) - t}\right)^2.$

Demonstração: Como $t \in [0, t_*)$ and $x \in K(t)$ usando a definição em (5.11), Proposição 5.1 e as duas primeiras afirmações na Proposição 2.10, obtemos

$$\|x - x_0\| \leq t, \quad \eta d(0, D_\Delta(x)) = \eta d(0, D_C(x)) \leqslant -\frac{f_{\xi,\alpha}(t)}{f'_{\xi,\alpha}(t)}, \quad t < n_{f_{\xi,\alpha}}(t) < t_*. \tag{5.15}$$

Agora, como $y \in G_F(x)$ segue que existe $d \in \bar{D}_\Delta(x)$ tal que $y = x + d$. Usando a definição do conjunto $\bar{D}_\Delta(x)$ em (5.13) e a segunda inequação em (5.15) temos que

$$\|d\| \leq \eta d(0, D_\Delta(x)) = \eta d(0, D_C(x)) \leq -f_{\xi,\alpha}(t)/f'_{\xi,\alpha}(t).$$

Como $d = y - x$, a última inequação juntamente com a definição em (2.15) implica o item **i**.

Combinado desigualdade triangular, primeira inequação em (5.15), item **i** e a última desigualdade em (5.15) obtemos que

$$\|y - x_0\| \leq \|y - x\| + \|x - x_0\| \leq n_{f_{\xi,\alpha}}(t) < t_*,$$

o que prova o item **ii**.

Prova do item **iii**. Como $\|y - x_0\| < t_* \leq r_{x_0}$, segue da quase regularidade que

$$D_C(y) \neq \emptyset, \qquad d(0, D_C(y)) \leq \beta_{x_0}(\|y - x_0\|) d(F(y), C).$$

Usando que $x \in K(t) \subset K$, $y - x = d \in D_\Delta(x)$ e Proposição 5.1 temos que

$$F(x) + F'(x)(y - x) \in C.$$

Portanto, combinando $\eta \geq 1$, a inequação acima e a última inclusão é fácil ver que

$$\eta d(0, D_C(y)) \leq \eta \beta_{x_0}(\|y - x_0\|) \|F(y) - F(x) - F'(x)(y - x)\|.$$

Agora, do item **i** temos que $\|y - x\| \leq n_{f_{\xi,\alpha}}(t) - t$. Daí, como $\|x - x_0\| \leq t$ segue da última inequação e do Lemma 2.2 que

$$\eta d(0, D_C(y)) \leq \eta \beta_{x_0}(\|y - x_0\|) e_f(t, n_{f_{\xi,\alpha}}(t)) \left(\frac{\|y - x\|}{n_{f_{\xi,\alpha}}(t) - t} \right)^2. \tag{5.16}$$

Por outro lado, usando definições em (2.12), (2.15), (2.9) e simples cálculos, temos

$$\begin{aligned} f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t)) &= f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t)) - f_{\xi,\alpha}(t) - f'_{\xi,\alpha}(t)(n_{f_{\xi,\alpha}}(t) - t) \\ &= \xi_1 + (\alpha - 1) n_{f_{\xi,\alpha}}(t) + \alpha f_e(n_{f_{\xi,\alpha}}(t)) - [\xi_1 + (\alpha - 1)t + \alpha f_e(t)] \\ &\quad - [(\alpha - 1) + \alpha f'_e(t)](n_{f_{\xi,\alpha}}(t) - t) \\ &= \alpha \left(f_e(n_{f_{\xi,\alpha}}(t)) - f_e(t) - f'_e(t)(n_{f_{\xi,\alpha}}(t) - t) \right) \\ &= \alpha e_{f_e}(t, n_{f_{\xi,\alpha}}(t)). \end{aligned}$$

Segue das duas últimas inequações, $\beta_{x_0}(t)$ uma função crescente e item **ii** que

$$\eta d(0, D(y)) \leq \frac{\eta \beta_{x_0}(n_{f_{\xi,\alpha}}(t))}{\alpha} f_\alpha(n_{f_{\xi,\alpha}}(t)) \left(\frac{\|y - x\|}{n_{f_{\xi,\alpha}}(t) - t} \right)^2.$$

Da Proposição 2.9 e primeira afirmação na Proposição 2.10 temos que $\xi \leq n_{f_{\xi,\alpha}}(t) < t_*$. Portanto, para concluir a prova combine a última desigualdade com Proposição 2.12. ∎

No próximo resultado, iremos mostrar que a multifunção iteração de Gauss-Newton está bem definida nos subconjuntos definidos em (5.11).

Lema 5.2 *Para cada $t \in [0, t_*)$, $K(t) \subset B(x_0, t_*)$ e*

$$\bar{G}_F\left(K(t)\right) \subset K\left(n_{f_{\xi,\alpha}}(t)\right).$$

Como consequência, $K \subset B(x_0, t_*)$ e $\bar{G}_F(K) \subset K$.

Demonstração: A primeira inclusão segue trivialmente da definição de $K(t)$.

Tome $x \in K(t)$ e $y \in \bar{G}_F(x)$. Combinando os itens **i** e **iii** do Lema 5.1 temos que

$$\eta d(0, D_C(y)) \leq -\frac{f_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t))}{f'_{\xi,\alpha}(n_{f_{\xi,\alpha}}(t))}.$$

Este resultado, juntamente com o item **ii** do Lema 5.1 e definição em (5.11) mostra que $y \in K(n_{f_{\xi,\alpha}}(t))$, o que prova a segunda inclusão.

A próxima inclusão, segue trivialmente da definições em (5.11) e (5.12). Para verificar a última inclusão, tome $x \in K$. Então $x \in K(t)$ para algum $t \in [0, t_*)$. Usando a primeira parte do lema, concluímos que $\bar{G}_F(x) \subset K(n_{f_{\xi,\alpha}}(t))$. Para concluir a prova, use $n_{f_{\xi,\alpha}}(t) \in [0, t_*)$ e definição de K. ∎

5.1.2 Prova do Teorema 5.1

Primeiro, note que a sequência gerada pelo **Algoritmo 5.1** satisfaz

$$x_{k+1} \in \bar{G}_F(x_k), \qquad k = 0, 1, \ldots, \tag{5.17}$$

o que necessariamente é uma definição equivalente para esta sequência.

Demonstração: Primeiro lembramos que todas as afirmações feitas no Teorema 5.1 para a sequência t_k foram provadas no Corolário 2.1.

Agora, como $x_0 \in B(x_0, t_*) \subseteq B(x_0, r_{x_0})$, usando a quase regularidade, $\eta \geq 1$, inequação em (5.6) e Proposição 2.7, obtemos

$$D_C(x_0) \neq \emptyset, \qquad \eta d(0, D_C(x_0)) \leq \eta \beta_{x_0}(0) d(F(x_0), C) \leq \xi = -\frac{f_{\xi,\alpha}(0)}{f'_{\xi,\alpha}(0)}.$$

Portanto,
$$x_0 \in K(0) \subset K,$$
onde a segunda inclusão segue trivialmente de (5.12). Usando a inclusão acima, a inclusão $\bar{G}_F(K) \subset K$ (Lemma 5.2) e (5.17) concluímos que a sequência $\{x_k\}$ permanece em K, em particular, temos que $\{x_k\}$ está contida em $B(x_0, t_*)$. Como $\{x_k\} \subset K$, combinando a Proposição 5.1 e **Algoritmo 5.1** segue que a inclusão em (5.7) vale. Iremos provar por indução que

$$x_k \in K(t_k), \qquad k = 0, 1, \dots . \tag{5.18}$$

A inclusão acima, para $k = 0$ é o primeiro resultado dentro desta prova. Assuma que $x_k \in K(t_k)$. De (2.19) temos que $t_{k+1} = \eta_{f_{\xi,\alpha}}(t_k)$ e como $x_k \in K(t_k)$ o Lema 5.2 implica que $\bar{G}_F(x_k) \subset K(t_{k+1})$, o que combinado com (5.17) completa a prova de indução de (5.18).

Agora, usando **Algoritmo 5.1**, (5.18), Proposição 5.1 e (5.11), temos

$$\|x_{k+1} - x_k\| \le \eta d(0, D_\Delta(x_k)) = \eta d(0, D(x_k)) \leqslant -\frac{f_{\xi,\alpha}(t_k)}{f'_{\xi,\alpha}(t_k)}, \qquad k = 0, 1, \dots \tag{5.19}$$

o qual, usando (5.5) é equivalente à

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

Então, a inequação em (5.8) vale. Por outro lado, como $\{t_k\}$ converge para t_*, a inequação acima implica que

$$\sum_{k=k_0}^{\infty} \|x_{k+1} - x_k\| \le \sum_{k=k_0}^{\infty} t_{k+1} - t_k = t_* - t_{k_0} < +\infty,$$

para qualquer $k_0 \in \mathbb{N}$. Daí, $\{x_k\}$ é uma sequência de Cauchy em $B(x_0, t_*)$ e assim, converge para algum $x_* \in B[x_0, t_*]$. Além disso, a última inequação também implica (5.10), i.e., $\|x_* - x_k\| \le t_* - t_k$, para qualquer k. Como C é fechado, $\{x_k\}$ converge para x_*,
$$F(x_k) + F'(x_k)(x_{k+1} - x_k) \in C$$
e F é uma função continuamente diferenciável, temos que $F(x_*) \in C$.

Agora para provar a inequação em (5.9), primeiro note que $x_k \in K(t_k)$ e $t_{k+1} = n_{f_{\xi,\alpha}}(t_k)$, para todo $k = 0, 1, \dots$. Deste modo, tome um k arbitrário e aplique o item **iii** do Lema 5.1 com $y = x_k$, $x = x_{k-1}$ e $t = t_{k-1}$, para obter

$$\eta d(0, D_C(x_k)) \le -\frac{f_{\xi,\alpha}(t_k)}{f'_{\xi,\alpha}(t_k)} \left(\frac{\|x_k - x_{k-1}\|}{t_k - t_{k-1}} \right)^2,$$

o qual usando (5.5) e a primeira inequação em (5.19) prova a inequação desejada.

Para concluir a prova, combine (5.10) com a última inequação do Corolário 2.1. ∎

5.2 Casos Especiais

Nesta seção, apresentaremos casos especiais do Teorema 5.1, i.e., o caso onde x_0 é um ponto regular da inclusão (5.2) e o caso onde x_0 satisfaz a condição de Robinson. Além disso, apresentaremos resultados sob condição Lipschitz clássica e condição de Smale para funções analíticas.

5.2.1 Resultados de convergência para ponto incial regular

Nesta seção mostraremos um teorema correspondente ao Teorema 5.1, onde assumimos que o ponto inicial é um ponto regular da inclusão (5.2), ver [3] e as referências dele. Além disso, daremos resultados sob condição Lipschitz clássica e condição de Smale para funções analíticas. Começamos definindo regularidade.

Definição 5.1 *Sejam $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função continuamente diferenciável e $h : \mathbb{R}^m \to \mathbb{R}$ uma função convexa de valores-reais com conjunto de mínimos C não-vazio. Um ponto $x_0 \in \mathbb{R}^n$ é um ponto regular da inclusão $F(x) \in C$ se*

$$Ker(F'(x_0)^T) \cap (C - F(x_0))^o = \{0\},$$

Como sabemos (ver [4]) a definição de ponto quase-regular estende a definição de ponto regular. A seguinte proposição relaciona os dois conceitos, onde a existência das constantes r e β é devido a Burke e Ferris em [3], e a segunda afirmação segue da Observação 1.2.

Proposição 5.2 *Seja $x_0 \in \mathbb{R}^n$ um ponto regular da inclusão $F(x) \in C$. Então existem constantes $r > 0$ e $\beta > 0$ tais que*

$$D_C(x) \neq \emptyset \quad e \quad d(0, D_C(x)) \leq \beta d(F(x), C), \qquad \forall x \in B(x_0, r).$$

Consequentemente, x_0 é um ponto quase-regular da inclusão $F(x) \in C$ com raio quase-regular $r_{x_0} \geq r$ e função limitante quase-regular $\beta_{x_0}(\cdot) \leq \beta$ em $[0, r)$, como definidos em (1.5) *e* (1.6)*, respectivamente.*

Daqui em diante, para cada ponto regular $x_0 \in \mathbb{R}^n$ da inclusão $F(x) \in C$, denotaremos por $r > 0$ e $\beta > 0$ as constantes associadas dadas pela última proposição.

Teorema 5.2 *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função continuamente diferenciável. Suponha que existem $R > 0$, $x_0 \in \mathbb{R}^n$ e uma função majorante esférica $f_e : [0,\ R) \to$*

$\mathbb{R}$ para a função F em $B(x_0, R)$. Dadas as constantes $\alpha > 0$ e $\xi > 0$, considere a função auxiliar $f_{\xi,\alpha} : [0, R) \to \mathbb{R}$,

$$f_{\xi,\alpha}(t) = \xi + (\alpha - 1)t + \alpha f_e(t).$$

*Se $f_{\xi,\alpha}$ satisfaz **a3**, i.e., t_* é o menor zero de $f_{\xi,\alpha}$, então a sequência gerada pelo método de Newton para resolver $f_{\xi,\alpha}(t) = 0$, com ponto inicial $t_0 = 0$,*

$$t_{k+1} = t_k - f'_{\xi,\alpha}(t_k)^{-1} f_{\xi,\alpha}(t_k), \quad k = 0, 1, \dots ,$$

está bem definida, $\{t_k\}$ é estritamente crescente, está contida em $[0, t_*)$, e converge Q-linearmente para t_*. Sejam $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$ e $h : \mathbb{R}^m \to \mathbb{R}$ uma função convexa de valores-reais com conjunto de mínimos C não-vazio. Assuma que x_0 é um ponto regular da inclusão $F(x) \in C$ com constantes associadas $r > 0$ e $\beta > 0$. Se $d(F(x_0), C) > 0$, $t_* \le r$,

$$\Delta \ge \xi \ge \eta\beta d(F(x_0), C), \qquad \alpha \ge \eta\beta/(\eta\beta[f'_e(\xi) + 1] + 1),$$

*então a sequência gerada pelo **Algoritmo 5.1**, denotada por $\{x_k\}$, está contida em $B(x_0, t_*)$,*

$$F(x_k) + F'(x_k)(x_{k+1} - x_k) \in C, \quad k = 0, 1, \dots ,$$

satisfaz as inequações

$$\|x_{k+1} - x_k\| \le |t_{k+1} - t_k|, \quad k = 0, 1 \dots ,$$

$$\|x_{k+1} - x_k\| \le \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2} \|x_k - x_{k-1}\|^2, \quad k = 1, 2, \dots ,$$

converge para um ponto $x_* \in B[x_0, t_*]$ tal que $F(x_*) \in C$,

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

*e a convergência é R-linear. Se, adicionalmente, $f_{\xi,\alpha}$ satisfaz **a4** então as sequências $\{t_k\}$ e $\{x_k\}$ convergem Q-quadraticamente e R-quadraticamente para t_* and x_*, respectivamente.*

Demonstração: Como x_0 é um ponto regular da inclusão $F(x) \in C$, temos da Proposição 5.2 que x_0 é um ponto quase-regular da inclusão $F(x) \in C$ com raio quase-regular $r_{x_0} \ge r$. Daí, usando que $t_* \le r$ obtemos

$$t_* < r_{x_0}.$$

Além disso, a Proposição 5.2 implica que a função limitante quase-regular satisfaz

$$\beta_{x_0}(t) \leq \beta, \qquad \forall\, t \in [0, r). \tag{5.20}$$

Como $\Delta \geq \xi \geq \eta\beta d(F(x_0)C)$ e a última inequação implica que $\beta_{x_0}(0) \leq \beta$, segue que

$$\Delta \geq \xi \geq \eta\beta_{x_0}(0)d(F(x_0), C).$$

Agora, combinando as hipóteses que $0 < \xi$ e $t_* \leq r$ com a primeira afirmação na Proposição 2.12 concluímos que $0 < \xi < t_* \leq r$. Portanto, usando (5.20), $f_e'(0) = -1$, f_e' é estritamente crescente e $\eta \geq 1$ obtemos, após simples cálculo que

$$\frac{\eta\beta}{\eta\beta(f_e'(\xi)+1)+1} \geq \frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f_e'(t)+1]+1}, \qquad \forall\, t \in [\xi, t_*).$$

Daí, a hipótese $\alpha \geq \eta\beta/[\eta\beta(f_e'(\xi)+1)+1]$ e última inequação implica que

$$\alpha \geq \sup\left\{\frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f_e'(t)+1]+1} : \xi \leq t < t_*\right\}.$$

Assim, F e x_0 satisfazem todas as hipóteses do Teorema 5.1 e as afirmações deste teorema seguem do Teorema 5.1. ∎

Quando F satisfaz a condição Lipschitz, o Teorema 5.2 torna-se:

Teorema 5.3 *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função continuamente diferenciável. Suponha que existem $R > 0$, $x_0 \in \mathbb{R}^n$ e $K > 0$, tais que*

$$\|F'(y) - F'(x)\| \leq K\|y - x\|, \qquad x, y \in B(x_0, R).$$

Dadas as constantes $\alpha > 0$ e $\xi > 0$, considere a função auxiliar $f_{\xi,\alpha} : [0, R) \to \mathbb{R}$,

$$f_{\xi,\alpha}(t) = \xi - t + (\alpha K t^2)/2.$$

Se $2\alpha K\xi \leq 1$, então $t_* = (1 - \sqrt{1 - 2\alpha K\xi}\,)/(\alpha K)$ é o menor zero de $f_{\xi,\alpha}$, a sequência gerada pelo método de Newton para resolver $f_{\xi,\alpha}(t) = 0$, com ponto inicial $t_0 = 0$,

$$t_{k+1} = t_k - f_{\xi,\alpha}'(t_k)^{-1} f_{\xi,\alpha}(t_k), \quad k = 0, 1, \ldots,$$

está bem definida, $\{t_k\}$ é estritamente crescente, está contida em $[0, t_*)$, e converge Q-linearmente para t_*. Sejam $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$ e $h : \mathbb{R}^m \to \mathbb{R}$ uma função convexa de valores-reais com conjunto de mínimos C não-vazio. Assuma que x_0 é um ponto regular da inclusão $F(x) \in C$ com constantes associadas $r > 0$ e $\beta > 0$.

Se $d(F(x_0), C) > 0$, $t_* \leq r$,

$$\Delta \geq \xi \geq \eta\beta d(F(x_0), C), \qquad \alpha \geq \eta\beta/(K\eta\beta\xi + 1),$$

então a sequência gerada pelo **Algoritmo 5.1**, *denotada por* $\{x_k\}$, *está contida em* $B(x_0, t_*)$,

$$F(x_k) + F'(x_k)(x_{k+1} - x_k) \in C, \quad k = 0, 1, \dots ,$$

satisfaz as inequações

$$\|x_{k+1} - x_k\| \leq |t_{k+1} - t_k|, \quad k = 0, 1 \dots ,$$

$$\|x_{k+1} - x_k\| \leq \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2}\|x_k - x_{k-1}\|^2, \quad k = 1, 2, \dots ,$$

converge para um ponto $x_* \in B[x_0, t_*]$ *tal que* $F(x_*) \in C$,

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

e a convergência é R-linear. Se, adicionalmente, $2\alpha K\xi < 1$ *então as sequências* $\{t_k\}$ *e* $\{x_k\}$ *convergem Q-quadraticamente e R-quadraticamente para* t_* *and* x_*, *respectivamente.*

Demonstração: É imediato provar que $f_e : [0, R) \to \mathbb{R}$ definida por $f_e(t) = Kt^2/2 - t$ é uma função majorante esférica para F em $B(x_0, R)$. Daí,

$$f_{\xi,\alpha}(t) = \xi - t + (\alpha K t^2)/2 = \xi + (\alpha - 1)t + \alpha f_e(t),$$

e como $2\alpha K\xi \leq 1$, concluímos que $f_{\xi,\alpha}$ satisfaz **a3** e $t_* = (1 - \sqrt{1 - 2\alpha K\xi}\,)/(\alpha K)$ é sua menor raiz. Neste caso, a constante α satisfaz

$$\alpha \geq \frac{\eta\beta}{1 + K\eta\beta\xi} = \frac{\eta\beta}{\eta\beta[f_e'(\xi) + 1] + 1}.$$

Portanto, tomando α, $f_{\xi,\alpha}$ e t_* como definido acima, a primeira parte do teorema segue do Teorema 5.2. Para provar a segunda parte, é suficiente notar que a hipótese $2\alpha K\xi < 1$ implica que $f_{\xi,\alpha}$ satisfaz **a4**. ∎

Quando F é analítica e satisfaz a condição de Smale, o Teorema 5.2 torna-se:

Teorema 5.4 *Seja* $F : \mathbb{R}^n \to \mathbb{R}^m$ *uma função analítica. Suponha que* $x_0 \in \mathbb{R}^n$ *e*

$$\gamma := \sup_{n>1} \left\| \frac{F^{(n)}(x_0)}{n!} \right\|^{1/(n-1)} < +\infty. \tag{5.21}$$

Dadas as constantes $\alpha > 0$ e $\xi > 0$, considere a função auxiliar $f_{\xi,\alpha} : [0, 1/\gamma) \to \mathbb{R}$,

$$f_{\xi,\alpha}(t) = \frac{\alpha\gamma}{1 - \gamma t} t^2 - t + \xi.$$

Se $\xi\gamma \leq 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)}$, então

$$t_* = \frac{1 + \gamma\xi - \sqrt{(1+\gamma\xi)^2 - 4(1+\alpha)\gamma\xi}}{2(1+\alpha)\gamma}$$

é o menor zero de $f_{\xi,\alpha}$, a sequência gerada pelo método de Newton para resolver $f_{\xi,\alpha}(t) = 0$, com ponto inicial $t_0 = 0$,

$$t_{k+1} = t_k - f'_{\xi,\alpha}(t_k)^{-1} f_{\xi,\alpha}(t_k), \quad k = 0, 1, \dots,$$

está bem definida, $\{t_k\}$ é estritamente crescente, está contida em $[0, t_*)$, e converge Q-linearmente para t_*. Sejam $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$ e $h : \mathbb{R}^m \to \mathbb{R}$ uma função convexa de valores-reais com conjunto de mínimos C não-vazio. Assuma que x_0 é um ponto regular da inclusão $F(x) \in C$ com constantes associadas $r > 0$ e $\beta > 0$. Se $d(F(x_0), C) > 0$, $t_* \leq r$,

$$\Delta \geq \xi \geq \eta\beta d(F(x_0), C), \qquad \alpha \geq \frac{\eta\beta(1-\gamma\xi)^2}{\eta\beta + (1-\eta\beta)(1-\gamma\xi)^2},$$

então a sequência gerada pelo **Algoritmo 5.1**, *denotada por $\{x_k\}$, está contida em $B(x_0, t_*)$,*

$$F(x_k) + F'(x_k)(x_{k+1} - x_k) \in C, \quad k = 0, 1, \dots,$$

satisfaz as inequações

$$\|x_{k+1} - x_k\| \leq |t_{k+1} - t_k|, \quad k = 0, 1 \dots,$$

$$\|x_{k+1} - x_k\| \leq \frac{t_{k+1} - t_k}{(t_k - t_{k-1})^2} \|x_k - x_{k-1}\|^2, \quad k = 1, 2, \dots,$$

converge para um ponto $x_* \in B[x_0, t_*]$ tal que $F(x_*) \in C$,

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

e a convergência é R-linear. Se, adicionalmente, $\xi\gamma < 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)}$ então as sequências $\{t_k\}$ e $\{x_k\}$ convergem Q-quadraticamente e R-quadraticamente para t_* and x_*, respectivamente.

Os seguintes resultados serão necessários na prova do teorema acima.

Lema 5.3 *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função analítica. Assuma que $x_0 \in \mathbb{R}^n$ e γ como* *definido* (5.21). *Então, para todo* $x \in B(x_0, 1/\gamma)$ *vale*

$$\|F''(x)\| \leqslant (2\gamma)/(1 - \gamma\|x - x_0\|)^3.$$

Demonstração: A prova segue o mesmo padrão do Lema 3.4 do capítulo 3. ∎

Lema 5.4 *Seja* $F : \mathbb{R}^n \to \mathbb{R}^m$ *uma função duas vezes diferenciável. Se existe uma função* $f : [0, R) \to \mathbb{R}$ *duas vezes continuamente diferenciável tal que*

$$\|F''(x)\| \leqslant f''(\|x - x_0\|),$$

para todo $x \in \mathbb{R}^n$ *e* $\|x - x_0\| < R$. *Então* F *e* f *satisfaz* (2.11).

Demonstração: A prova segue o mesmo padrão do Lema 3.5 do capítulo 3. ∎

Prova do Teorema 5.4. *Considere a função real* $f_e : [0, 1/\gamma) \to \mathbb{R}$ *definida por*

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

É simples mostrar que f *é analítica e*

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

para $n \geq 2$. *Segue das equações acima que* f_e *satisfaz* **a1** *e* **a2** *na definição 2.2. Agora, como* $f_e''(t) = (2\gamma)/(1 - \gamma t)^3$, *combinando os Lemas 5.3 e 5.4, temos que* F *e* f_e *satisfaz* (2.11) *com* $R = 1/\gamma$. *Portanto,* f_e *é uma função majorante esférica para* F *em* $B(x_0, 1/\gamma)$. *Daí,*

$$f_{\xi,\alpha}(t) = \frac{\alpha\gamma}{1 - \gamma t} t^2 - t + \xi = \xi + (\alpha - 1)t + \alpha f_e(t),$$

e como $\xi\gamma \leq 1 + 2\alpha - 2\sqrt{\alpha(1 + \alpha)}$, *concluímos que* $f_{\xi,\alpha}$ *satisfaz* **a3** *e*

$$t_* = \frac{1 + \gamma\xi - \sqrt{(1 + \gamma\xi)^2 - 4(1 + \alpha)\gamma\xi}}{2(1 + \alpha)\gamma}$$

é sua menor raiz. Neste caso, a constante α *satisfaz*

$$\alpha \geq \frac{\eta\beta(1 - \gamma\xi)^2}{\eta\beta + (1 - \eta\beta)(1 - \gamma\xi)^2} = \frac{\eta\beta}{\eta\beta[f'(\xi) + 1] + 1}.$$

Portanto, tomando α, $f_{\xi,\alpha}$ e t_* como definido acima, a primeira parte do teorema segue do Teorema 5.2. Para provar a segunda parte, é suficiente notar que a hipótese $\xi\gamma < 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)}$ implica que $f_{\xi,\alpha}$ satisfaz **a4**.

5.2.2 Resultado de convergence sob condição de Robinson

Nesta seção mostraremos um teorema correspondente ao Teorema 5.1, onde assumimos que o ponto inicial satisfaz a condição de Robinson, ver [4] e [32]. Além disso, daremos resultados sob condição Lipschitz clássica e condição de Smale para funções analíticas.

Inicialmente, sejam $C \subset \mathbb{R}^m$ um cone convexo fechado não-vazio, $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função continuamente diferenciável e $x \in \mathbb{R}^n$. Defina a multifunção $T_x : \mathbb{R}^n \to P(\mathbb{R}^m)$ por

$$T_x d = F'(x)d - C. \tag{5.22}$$

A multifunção T_x é um processo convexo de $\mathbb{R}^n$ em $\mathbb{R}^m$; processos convexos tem sido extensivamente estudados em [33, 34]. Como usual, o domínio, norma e inversa de T_x são definidos, respectivamente, por

$$\mathcal{D}(T_x) := \{d \in \mathbb{R}^n : T_x d \neq \emptyset\}, \qquad \|T_x\| := sup\,\{\|T_x d\| : x \in \mathcal{D}(T_x),\ \|d\| \leq 1\},$$

$$T_x^{-1} y := \{d \in \mathbb{R}^n : F'(x)d \in y + C\}, \qquad y \in \mathbb{R}^m.$$

onde $\|T_x d\| := \inf\{\|v\| \ : \ v \in T_x d\}$.

O ponto $x_0 \in \mathbb{R}^n$, satisfaz a *condição de Robinson* se a multifunção T_{x_0} leva $\mathbb{R}^n$ em $\mathbb{R}^m$, i.e.,

$$\forall\, y \in \mathbb{R}^m \quad \exists\, d \in \mathbb{R}^n,\ \exists\, c \in C; \quad y = F'(x_0)d - c. \tag{5.23}$$

Teorema 5.5 *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função continuamente diferenciável. Suponha que existem $R > 0$, $x_0 \in \mathbb{R}^n$ e uma função majorante esférica $f_e : [0,\ R) \to \mathbb{R}$ para a função F em $B(x_0, R)$. Dadas as constantes $\alpha > 0$ e $\xi > 0$, considere a função auxiliar $f_{\xi,\alpha} : [0, R) \to \mathbb{R}$,*

$$f_{\xi,\alpha}(t) = \xi + (\alpha - 1)t + \alpha f_e(t).$$

Se $f_{\xi,\alpha}$ satisfaz **a3**, *i.e., t_* é o menor zero de $f_{\xi,\alpha}$, então a sequência gerada pelo método de Newton para resolver $f_{\xi,\alpha}(t) = 0$, com ponto inicial $t_0 = 0$,*

$$t_{k+1} = t_k - f'_{\xi,\alpha}(t_k)^{-1} f_{\xi,\alpha}(t_k), \quad k = 0, 1, \ldots\,,$$

está bem definida, $\{t_k\}$ é estritamente crescente, está contida em $[0, t_*)$, e converge Q-linearmente para t_*. Sejam $\eta \in [1, \infty)$, $\Delta \in (0, \infty]$ e $h : \mathbb{R}^m \to \mathbb{R}$ uma função *convexa de valores-reais com conjunto de mínimos* C *não-vazio. Assuma que* C *é um cone e* x_0 *satisfaz à condição de Robinson. Seja* $\beta_0 = \|T_{x_0}^{-1}\|$. *Se* $d(F(x_0), C) > 0$, $t_* \leq r_{\beta_0} := \sup\{t \in [0, R) \ : \ \beta_0 - 1 + \beta_0 f_e'(t) < 0\}$,

$$\Delta \geq \xi \geq \eta\beta_0 d(F(x_0), C), \qquad \alpha \geq \frac{\eta\beta_0}{1 + (\eta - 1)\beta_0[f_e'(\xi) + 1]},$$

então a sequência gerada pelo **Algoritmo 5.1**, *denotada por* $\{x_k\}$, *está contida em* $B(x_0, t_*)$,

$$F(x_k) + F'(x_k)(x_{k+1} - x_k) \in C, \quad k = 0, 1, \ldots,$$

satisfaz as inequações

$$\|x_{k+1} - x_k\| \leq |t_{k+1} - t_k|, \quad k = 0, 1 \ldots,$$

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

converge para um ponto $x_* \in B[x_0, t_*]$ *tal que* $F(x_*) \in C$,

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

e a convergência é R-linear. Se, adicionalmente, $f_{\xi,\alpha}$ *satisfaz* **a4** *então as sequências* $\{t_k\}$ *e* $\{x_k\}$ *convergem Q-quadraticamente e R-quadraticamente para* t_* *and* x_*, *respectivamente.*

Os seguinte resultados serão necessários na prova do teorema acima.

Lema 5.5 *Seja* $F : \mathbb{R}^n \to \mathbb{R}^m$ *uma função continuamente diferenciável e* C *um cone convexo fechado não-vazio. Suponha que* $x_0 \in \mathbb{R}^n$ *satisfaz a condição de Robinson. Então*

$$\|T_{x_0}^{-1}\| < +\infty.$$

Além disso, se S *é uma transformação linear de* $\mathbb{R}^n$ *em* $\mathbb{R}^m$ *tal que* $\|T_{x_0}^{-1}\|\|S\| < 1$, *então o processo convexo* $\bar{T}$, *definido por* $\bar{T} := T_{x_0} + S$, *leva* $\mathbb{R}^n$ *em* $\mathbb{R}^m$, $\|\bar{T}^{-1}\| < +\infty$,

$$\|\bar{T}^{-1}\| \leq \frac{\|T_{x_0}^{-1}\|}{1 - \|T_{x_0}^{-1}\|\|S\|}.$$

Demonstração: Ver Teorema 1 na p.342 de [32]. ∎

Lema 5.6 *Seja* $F : \mathbb{R}^n \to \mathbb{R}^m$ *uma função continualmente diferenciável e* $h : \mathbb{R}^m \to \mathbb{R}$ *uma função convexa de valores-reais com conjunto de mínimos* C *não-vazio. Suponha que* $x_0 \in \mathbb{R}^n$ *satisfaz a condição Robinson. Então* x_0 *é um ponto* *regular da inclusão $F(x) \in C$, em particular, x_0 é um ponto quase-regular da inclusão $F(x) \in C$. Além diso, assuma que C é um cone, $R > 0$ e $f_e : [0,\ R) \to \mathbb{R}$ é uma função majorante esférica de F em $B(x_0, R)$. Seja $\xi > 0$, $\beta_0 = \|T_{x_0}^{-1}\|$, a função auxiliar $f_{\xi,\beta_0} : [0, R) \to \mathbb{R}$,*

$$f_{\xi,\beta_0}(t) := \xi + (\beta_0 - 1)t + \beta_0 f_e(t),$$

e $r_{\beta_0} := \sup\{t \in [0, R) : f'_{\xi,\beta_0}(t) < 0\}$. Se r_{x_0} é o raio quase-regular e $\beta_{x_0}(\cdot)$ é a função limitante quase-regular do ponto quase-regular x_0, então

$$r_{x_0} \geq r_{\beta_0}, \qquad \beta_{x_0}(t) \leq \frac{\beta_0}{1 - \beta_0[f'_e(t) + 1]}, \qquad \forall\, t \in [0,\ r_{\beta_0}).$$

Demonstração: Tome $y \in Ker(F'(x_0)^T) \cap (C - F(x_0))^o$. Daí,

$$0 = \langle F'(x_0)^T y, d\rangle = \langle y, F'(x_0)d\rangle, \quad \forall\, d \in \mathbb{R}^n, \qquad \langle y, c - F(x_0)\rangle \leq 0, \quad \forall\, c \in C.$$

Como x_0 satisfaz a condição de Robinson segue que existem $d \in \mathbb{R}^n$ e $c \in C$ tais que $-y - F(x_0) = F'(x_0)d - c$, o que combinado com as inequações acima implica

$$\langle y, y\rangle = \langle y, c - F(x_0) - F'(x_0)d\rangle = \langle y, c - F(x_0)\rangle \leq 0.$$

Portanto, $y = 0$ e obtemos da Definição 5.1 que x_0 e um ponto regular da inclução $F(x) \in C$.

Para estabelecer a segunda parte, primeiro tome $x \in \mathbb{R}^n$ tal que $\|x - x_0\| \leq r_{\beta_0}$. Usando que f_e é uma função majorante esférica de F em $B(x_0, R)$, definições de β_0, f_{ξ,β_0} e r_{β_0} , temos que

$$\|T_{x_0}^{-1}\|\|F'(x) - F'(x_0)\| \leq \beta_0[f'_e(\|x - x_0\|) - f'_e(0)] = f'_{\xi,\beta_0}(\|x - x_0\|) + 1 < 1. \quad (5.24)$$

Combinando a hipótese que x_0 satisfaz a condição de Robinson com a última inequação, obtemos do Lema 5.5 que o processo convexo

$$T_x d = F'(x)d - C = T_{x_0}d + [F'(x) - F'(x_0)]d, \qquad \forall\, d \in \mathbb{R}^n,$$

leva $\mathbb{R}^n$ em $\mathbb{R}^m$ e

$$\|{T_x}^{-1}\| \leq \frac{\|T_{x_0}^{-1}\|}{1 - \|T_{x_0}^{-1}\|\|F'(x) - F'(x_0)\|} \leq \frac{\beta_0}{1 - \beta_0[f'_e(\|x - x_0\|) - f'_e(0)]}, \qquad (5.25)$$

onde a última inequação segue da definição de β_0 e (5.24). Além disso, como T_x leva $\mathbb{R}^n$ em $\mathbb{R}^m$ é fácil ver que

$$D_C(x) = \{d \in \mathbb{R}^n : F(x) + F'(x)d \in C\} \neq \emptyset, \qquad \forall\, x \in B(x_0, r_{\beta_0}). \qquad (5.26)$$

Agora, seja $d \in T_x^{-1}(c - F(x))$. Usando a definição de T_x^{-1} segue que

$$F'(x)d \in c - F(x) + C = C - F(x),$$

ou seja, $F(x) + F'(x)d \in C$, o que combinado com a definição de $D_C(x)$ implicam

$$T_x^{-1}(c - F(x)) \subset D_C(x).$$

Portanto,

$$d(0, D_C(x)) \le \|T_x^{-1}(c - F(x))\| \le \|T_x^{-1}\|\|c - F(x)\|, \qquad \forall c \in C.$$

A última inequação juntamente com (5.25) implicam

$$d(0, D_C(x)) \le \|T_x^{-1}\| d(F(x), C) \le \frac{\beta_0}{1 - \beta_0[f_e'(\|x - x_0\|) - f_e'(0)]} d(F(x), C),$$

o que combinado com (5.26), definições de r_{x_0} e $\beta_{x_0}(\cdot)$ em (1.5) e (1.6), respectivamente, implicam a inequação desejada. ∎

[**Prova do Teorema 5.5**] *Como $x_0 \in \mathbb{R}^n$ satisfaz a condição de Robinson, temos do Lema 5.6 que x_0 é um ponto quase-regular da inclusão $F(x) \in C$ com raio quase-regular $r_{x_0} \ge r_{\beta_0}$. Daí, usando que $t_* \le r_{\beta_0}$ obtemos*

$$t_* < r_{x_0}.$$

Além disso, o Lema 5.6 implica que a função limitante quase-regular $\beta_{x_0}(\cdot)$ satisfaz

$$\beta_{x_0}(t) \le \frac{\beta_0}{1 - \beta_0[f_e'(t) + 1]}, \qquad \forall\, t \in [0,\ r_{\beta_0}). \tag{5.27}$$

Como $\Delta \ge \xi \ge \eta\beta_0 d(F(x_0), C)$ e a última inequação implica que $\beta_{x_0}(0) \le \beta_0$, segue que

$$\Delta \ge \xi \ge \eta\beta_{x_0}(0) d(F(x_0), C).$$

Agora, combinando as hipóteses que $0 < \xi$ e $t_* \le r_{\beta_0}$ com a primeira afirmação na Proposição 2.12 concluímos que $0 < \xi < t_* \le r_{\beta_0}$. Portanto, usando (5.27), $f'(0) = -1$, f_e' estritamente crescente e $\eta \ge 1$ obtemos, após simples cálculo que

$$\eta[f_e'(t)+1] + \frac{1}{\beta_{x_0}(t)} \ge \frac{1}{\beta_0} + (\eta-1)[f_e'(t)+1] \ge \frac{1}{\beta_0} + (\eta-1)[f_e'(\xi)+1], \qquad \forall\, t \in [\xi,\ t_*),$$

ou equivalentemente,

$$\frac{\eta\beta_0}{1 + (\eta - 1)\beta_0[f_e'(\xi) + 1]} \ge \frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f_e'(t) + 1] + 1}, \qquad \forall\, t \in [\xi,\ t_*). \tag{5.28}$$

Daí, a hipótese $\alpha \geq \eta\beta_0/[1+(\eta-1)\beta_0(f_e'(\xi)+1)]$ e última inequação implicam que

$$\alpha \geq \sup\left\{\frac{\eta\beta_{x_0}(t)}{\eta\beta_{x_0}(t)[f_e'(t)+1]+1} : \xi \leq t < t_*\right\}.$$

Assim, F e x_0 satisfazem todas as hipóteses do Teorema 5.1 e as afirmações deste teorema seguem do Teorema 5.1.

Observação 5.3 *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função continuamente diferenciável. Suponha que existem $R > 0$, $x_0 \in \mathbb{R}^n$ e $K > 0$, tais que*

$$\|F'(y) - F'(x)\| \leq K\|y - x\|, \qquad x, y \in B(x_0, R).$$

Note que, $f_e : [0, R) \to \mathbb{R}$ definida por $f_e(t) = Kt^2/2 - t$ é uma função majorante esférica para F em $B(x_0, R)$. Neste caso, é fácil ver que **a3**, **a4** *e t_* no Teorema 5.5 são*

$$2\alpha K\xi \leq 1, \qquad 2\alpha K\xi < 1, \qquad t_* = (1 - \sqrt{1 - 2\alpha K\xi}\,)/(\alpha K),$$

e α satisfiaz

$$\alpha \geq \frac{\eta\beta_0}{1 + (\eta - 1)K\beta_0\xi}.$$

Em particular, se $C = \{0\}$, $n = m$ a condição de Robinson é equivalente à condição de que $F'(x_0)^{-1}$ é não-singular. Daí, para $\eta = 1$ obtemos a convergência semi-local do método de Newton sob condição Lipschitz, ver [6].

Observação 5.4 *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma função analítica. Suponha que $x_0 \in \mathbb{R}^n$ e*

$$\gamma := \sup_{n>1} \left\| \frac{F^{(n)}(x_0)}{n!} \right\|^{1/(n-1)} < +\infty.$$

Note que a função real $f_e : [0, 1/\gamma) \to \mathbb{R}$ definida por $f_e(t) = t/(1 - \gamma t) - 2t$ é uma função majorante esférica para F em $B(x_0, 1/\gamma)$. Neste caso, é fácil ver que **a3**, **a4** *e t_* no Teorema 5.5 são*

$$\xi\gamma \leq 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)}, \qquad \xi\gamma < 1 + 2\alpha - 2\sqrt{\alpha(1+\alpha)},$$

$$t_* = \frac{1 + \gamma\xi - \sqrt{(1+\gamma\xi)^2 - 4(1+\alpha)\gamma\xi}}{2(1+\alpha)\gamma},$$

e α satisfaz

$$\alpha \geq \frac{\eta\beta_0(1-\gamma\xi)^2}{(\eta-1)\beta_0 + [1 - \beta_0(\eta-1)](1-\gamma\xi)^2}.$$

Em particular, se $C = \{0\}$, $n = m$ a condição de Robinson é equivalente à condição que $F'(x_0)^{-1}$ é não-singular. Daí, para $\eta = 1$ obtemos a convergência semi-local do método de Newton sob condição de Smale, ver [27].

Considerações Finais

Neste trabalho, estudamos e demonstramos resultados acerca da convergência local e semi-local dos métodos de Gauss-Newton.

A convergência dos métodos de Gauss-Newton também foi estudada nas recentes referências [4, 10, 11, 13, 18, 19, 24]). A principal diferença destas análises com as nossas análises é que em lugar de nossa condição majorante radial é utilizada a condição radial de Wang

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq \int_{\tau\|x-x_*\|}^{\|x-x_*\|} L(u)du, \qquad o \leq \tau \leq 1, \tag{5.29}$$

onde L é uma função escalar não-decrescente, e no lugar da nossa condição majorante esférica é utilizada a condição esférica de Wang

$$\|F'(y) - F'(x)\| \leq \int_{\|x-x_0\|}^{\|y-x\|+\|x-x_0\|} L(u)du, \qquad o \leq \tau \leq 1, \tag{5.30}$$

onde L é uma função escalar não-decrescente.

Entretanto, note que as condições (5.29) e (5.30) podem ser vistas como um caso particular da função majorante radial e esférica, respectivamente. Para verificarmos esta afirmação, basta escolhermos as funções majorantes radial e esférica da seguinte forma:

$$f_r(t) = f_e(t) = \int_0^t L(u)(t-u)du - t.$$

Temos que as funções $f_r(t)$ e $f_e(t)$ definidas acima satisfazem as hipóteses **h1**, **h2**, **a1** e **a2** . Além disso, também temos que

$$f_r'(t) = f_e'(t) = \int_0^t L(u)du - 1.$$

Assim, obtemos após alguns cálculos que

$$f_r'(\|x - x_*\|) - f_r'(\tau\|x - x_*\|) = \int_{\tau\|x-x_*\|}^{\|x-x_*\|} L(u)du,$$

$$f_e'(\|y - x\| + \|x - x_0\|) - f_e'(\|x - x_0\|) = \int_{\|x-x_0\|}^{\|y-x\|+\|x-x_0\|} L(u)du,$$

o que prova nossa afirmação. De fato, podemos mostrar que as condições da função majorante radial e esférica são equivalentes as condições (5.29) e (5.30) respectivamente. Mas, como adotadas aqui são mais naturais e deixam claro as suas relações com a função não-linear, refletindo de maneira simétrica em todos os resultados.

Nossa contribuição foi reformular os teoremas de convergência para os métodos de Gauss-Newton usando o princípio majorante de Kantorovich. Esta nova abordagem deixou clara a relação entre as funções majorantes com a função não-linear em consideração, o que não estava explícito nas outras demonstrações. Com isso, os resultados apresentados aqui tornaram as condições e demonstrações de convergência mais simples e mais didática. Além disso, no caso de convergência local dos métodos de Gauss-Newton o contexto foi estendido para espaços de Hilbert e no caso quase Gauss-Newton inexato adicionamos resultados de convergência para a condição de Smale. Ademais, na prova de convergência semi-local da sequência gerada pelo algoritmo de Gauss-Newton utilizamos uma técnica que no lugar de olhar apenas a sequência gerada, identificamos regiões onde a sequência de Gauss-Newton para o problema de otimização convexa é bem definido quando comparado com método de Newton aplicado a uma função auxiliar associada a função majorante.

Uma outra contribuição foi fazer um estudo completo das funções majorante, onde quase todos os resultados necessários para a convergência dos métodos de Gauss-Newton foram demonstrados, deixando o texto “auto-contido”.

Por fim, como proposta de trabalho futuro sugerimos o estudo de convergência de outras variações dos métodos de Newton e de Gauss-Newton usando nossa condição majorante.

Referências Bibliográficas

[1] Floudas, C. A., Pardalos, P. M. (Eds.). *Encyclopedia of Optimization, Second Edition*. Springer, 2009. ISBN: 978-0-387-74758-3.

[2] MORINI, B. "Convergence Behaviour of Inexact Newton Methods", *Math. Comp.*, v. 68, pp. 1605–1613, 1999.

[3] BURKE, J. V., FERRIS, M. C. "A Gauss-Newton method for convex composite optimization", *Math. Programming*, v. 71, n. 2, Ser. A, pp. 179–194, 1995. ISSN: 0025-5610. doi: 10.1007/BF01585997.

[4] LI, C., NG, K. F. "Majorizing functions and convergence of the Gauss-Newton method for convex composite optimization", *SIAM J. Optim*, pp. 613–642, 2007.

[5] LI, C., WANG, X. "On convergence of the Gauss-Newton method for convex composite optimization", *Math. Program.*, v. 91, pp. 349–356, 2002. ISSN: 0025-5610. 10.1007/s101070100249.

[6] DENNIS, JR., J. E., SCHNABEL, R. B. *Numerical methods for unconstrained optimization and nonlinear equations*, v. 16, *Classics in Applied Mathematics*. Philadelphia, PA, Society for Industrial and Applied Mathematics (SIAM), 1996. ISBN: 0-89871-364-1. Corrected reprint of the 1983 original.

[7] NOCEDAL, J., WRIGHT, S. J. *Numerical optimization*. Springer Series in Operations Research. New York, Springer-Verlag, 1999. ISBN: 0-387-98793-2.

[8] KANTOROVIČ, L. V. "The principle of the majorant and Newton's method", *Doklady Akad. Nauk SSSR (N.S.)*, v. 76, pp. 17–20, 1951.

[9] ARGYROS, I. K., HILOUT, S. "Improved generalized differentiability conditions for Newton-like methods", *J. Complexity*, v. In Press, Corrected Proof, pp. –, 2010. ISSN: 0885-064X. doi: DOI: 10.1016/j.jco.2009.12.001.

[10] CHEN, J., LI, W. "Convergence of Gauss-Newton's method and uniqueness of the solution", *Appl. Math. Comput.*, v. 170, n. 1, pp. 686–705, 2005. ISSN: 0096-3003.

[11] CHEN, J., LI, W. "Local convergence results of Gauss-Newton's like method in weak conditions", *J. Math. Anal. Appl.*, v. 324, n. 2, pp. 1381 – 1394, 2006. ISSN: 0022-247X. doi: DOI: 10.1016/j.jmaa.2006.01.032.

[12] DEDIEU, J. P., KIM, M. H. "Newton's method for analytic systems of equations with constant rank derivatives", *J. Complexity*, v. 18, n. 1, pp. 187–209, 2002. ISSN: 0885-064X. doi: 10.1006/jcom.2001.0612.

[13] CHEN, J., LI, W. "Convergence behaviour of inexact Newton methods under weak Lipschitz condition", *J. Comput. Appl. Math.*, v. 191, n. 1, pp. 143–164, 2006. ISSN: 0377-0427.

[14] DEDIEU, J. P., SHUB, M. "Newton's method for overdetermined systems of equations", *Math. Comp.*, v. 69, n. 231, pp. 1099–1115, 2000. ISSN: 0025-5718. doi: 10.1090/S0025-5718-99-01115-1.

[15] FERREIRA, O. P. "Local convergence of Newton's method in Banach space from the viewpoint of the majorant principle", *IMA J. Numer. Anal.*, v. 29, n. 3, pp. 746–759, 2009. ISSN: 0272-4979.

[16] FERREIRA, O. P., GONÇALVES, M. L. N. "Local convergence analysis of inexact Newton-like methods under majorant condition", *Comput. Optim. Appl.*, v. Article in Press, pp. 1–21, 2009. ISSN: 0926-6003.

[17] FERREIRA, O. P., SVAITER, B. F. "Kantorovich's majorants principle for Newton's method", *Comput. Optim. Appl.*, v. 42, n. 2, pp. 213–229, 2009. ISSN: 0926-6003.

[18] LI, C., HU, N., WANG, J. "Convergence behavior of Gauss-Newton's method and extensions of the Smale point estimate theory", *J. Complexity*, v. In Press, Corrected Proof, pp. –, 2010. ISSN: 0885-064X. doi: DOI: 10.1016/j.jco.2010.02.001.

[19] LI, C., ZHANG, W.-H., JIN, X.-Q. "Convergence and uniqueness properties of Gauss-Newton's method", *Comput. Math. Appl.*, v. 47, n. 6-7, pp. 1057–1067, 2004. ISSN: 0898-1221.

[20] PROINOV, P. D. "General local convergence theory for a class of iterative processes and its applications to Newton's method", *J. Complexity*, v. 25, n. 1, pp. 38 – 62, 2009. ISSN: 0885-064X. doi: DOI: 10.1016/j.jco.2008.05.006.

[21] PROINOV, P. D. "New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems", *J. Complexity*, v. 26, n. 1, pp. 3 – 42, 2010. ISSN: 0885-064X. doi: DOI: 10.1016/j.jco.2009.05.001.

[22] WANG, X. "Convergence of Newton's method and uniqueness of the solution of equations in Banach space", *IMA J. Numer. Anal.*, v. 20, n. 1, pp. 123–134, 2000. ISSN: 0272-4979.

[23] FERREIRA, O. P. "Local convergence of Newton's method under majorant condition", *J.Comput.Appl.Math.*, v. 235, n. 5, pp. 1515 – 1522, 2011. ISSN: 0377-0427. doi: DOI: 10.1016/j.cam.2010.08.038.

[24] CHEN, J. "The convergence analysis of inexact Gauss-Newton methods for nonlinear problems", *Comput. Optim. Appl.*, v. 40, n. 1, pp. 97–118, 2008. ISSN: 0926-6003. doi: 10.1007/s10589-007-9071-7.

[25] FERREIRA, O. P., SVAITER, B. F. "Kantorovich's Theorem on Newton's Method in Riemannian Manifolds", *J. Complexity*, v. 18, n. 1, pp. 304 – 329, 2002. ISSN: 0885-064X. doi: DOI: 10.1006/jcom.2001.0582.

[26] FERREIRA, O. P., GONÇALVES, M. L. N., OLIVEIRA, P. R. "Local convergence analysis of the Gauss-Newton method under a majorant condition", *J. Complexity*, v. 27, n. 1, pp. 111 – 125, 2011. ISSN: 0885-064X. doi: DOI: 10.1016/j.jco.2010.09.001.

[27] SMALE, S. "Newton's method estimates from data at one point". In: *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, Springer, pp. 185–196, New York, 1986.

[28] STEWART, G. W. "On the continuity of the generalized inverse", *SIAM J. Appl. Math.*, v. 17, pp. 33–45, 1969. ISSN: 0036-1399.

[29] WEDIN, P.-Å. "Perturbation theory for pseudo-inverses", *Nordisk Tidskr. Informationsbehandling (BIT)*, v. 13, pp. 217–232, 1973.

[30] HIRIART-URRUTY, J.-B., LEMARÉCHAL, C. *Convex analysis and minimization algorithms. I*, v. 305, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Berlin, Springer-Verlag, 1993. ISBN: 3-540-56850-6. Fundamentals.

[31] FERREIRA, O. P., GONÇALVES, M. L. N., OLIVEIRA, P. R. "Local convergence analysis of inexact Gauss-Newton like method under a majorant condition", *preprint, http://arxiv.org/abs/1008.1916v1.*

[32] ROBINSON, S. M. "Extension of Newton's method to nonlinear functions with values in a cone", *Numer. Math.*, v. 19, pp. 341–347, 1972. ISSN: 0029-599X. 10.1007/BF01404880.

[33] ROCKAFELLAR, R. T. *Monotone processes of convex and concave type.* Memoirs of the American Mathematical Society, No. 77. Providence, R.I., American Mathematical Society, 1967.

[34] ROCKAFELLAR, R. T. *Convex analysis.* Princeton Mathematical Series, No. 28. Princeton, N.J., Princeton University Press, 1970.

[35] BLUM, L., CUCKER, F., SHUB, M., et al. *Complexity and real computation.* New York, Springer-Verlag, 1998. ISBN: 0-387-98281-7. With a foreword by Richard M. Karp.