

**ALGORITMOS PARA  
COMPLEMENTARIDADE NÃO LINEAR E  
PROBLEMAS RELACIONADOS**

**Rosana Pérez Mera**

**Rosana Pérez Mera**

**ALGORITMOS PARA  
COMPLEMENTARIDADE NÃO LINEAR E  
PROBLEMAS RELACIONADOS**

**Tese apresentada ao Curso de Doutorado do Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas, como requisito parcial à obtenção do título de Doutor em Matemática Aplicada.**

**Area: Otimização**

**Orientador: Prof. Dr. José Mario Martínez**

**Co-Orientador: Profa. Dra. Vera L.R. Lopes**

**Campinas  
IMECC - UNICAMP  
1997**

UNIDADE	30
N.º CHAMADA :	
V.	Ex.
TOMBO BC	30132
PROC.	395/98
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREÇO	R\$ 11,00
DATA	28/04/98
N.º CPD	

CM-00105540-0

**FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DO IMECC DA UNICAMP**

Pérez Mera, Rosana

P415a      Algoritmos para complementaridade não linear e problemas relacionados / Rosana Pérez Mera -- Campinas, [S.P. :s.n.], 1997.

Orientador : José Mario Martínez

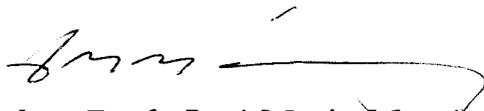
Tese (doutorado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Sistemas não diferenciáveis. 2. Complementaridade não linear. 3. Métodos quase-Newton. I. Martínez, José Mario. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

# ALGORITMOS PARA COMPLEMENTARIDADE NÃO LINEAR E PROBLEMAS RELACIONADOS

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida por Rosana Pérez Mera e aprovada pela Comissão Julgadora.

Campinas, 11 de Dezembro de 1997



Orientador: Prof. José Mario Martínez

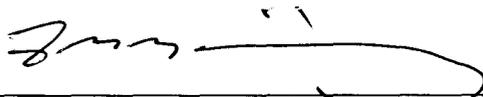


Co-Orientador: Profa. Vera L.R. Lopes

Tese apresentada ao Curso de Doutorado do Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas, como requisito parcial à obtenção do título de Doutor em Matemática Aplicada.

**Tese de Doutorado defendida e aprovada em 11 de dezembro de 1997**

**Pela Banca Examinadora composta pelos Profs. Drs.**



---

**Prof (a). Dr (a). JOSÉ MÁRIO MARTÍNEZ PÉREZ**



---

**Prof (a). Dr (a). IGOR LITVINCHEV**



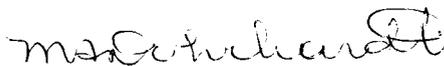
---

**Prof (a). Dr (a). CLOVIS PERIN FILHO**



---

**Prof (a). Dr (a). ROBERTO ANDREANI**



---

**Prof (a). Dr (a). MARIA APARECIDA DINIZ EHRHARDT**

# Agradecimentos

Ao Professor, **José Mario Martínez** pelas idéias, comentários oportunos e pelo estímulo, os quais contribuíram em forma significativa ao desenvolvimento desta tese.

À Professora, **Vera L. R. Lopes** cujas sugestões e contribuições foram fundamentais para conferir e avaliar vários resultados deste trabalho. Também agradeço por sua particular preocupação na redação desta tese e por seu apoio e estímulo constantes.

Ao professor, **Lúcio Tunes dos Santos**, pela sua colaboração em alguns dos aspectos computacionais desta tese.

Aos professores e colegas do grupo de Otimização.

À secretaria de Pós-graduação e do Departamento de Matemática Aplicada.

Agradeço à FAPESP pelo suporte financeiro.

*Aos meus pais*

## Resumo

É apresentada uma família de métodos **Least Change Secant Update** para resolver problemas de Complementaridade Não Linear (PCN), baseados em sistemas de equações não lineares não diferenciáveis. Resultados de convergência local e superlinear são provados. São comparadas duas reformulações distintas do PCN como sistema de equações não lineares, tanto do ponto de vista teórico como prático. Um algoritmo global para resolver o PCN, que usa os algoritmos introduzidos aqui é também apresentado. Alguns experimentos numéricos mostram um bom desempenho deste algoritmo.

## Abstract:

A family of **Least Change Secant Update** methods for solving **Nonlinear Complementarity Problems** based on **Nonsmooth Systems of Equations** is introduced. **Local and superlinear convergence results** for the algorithms are proved. **Two different reformulations** of the **Nonlinear Complementarity Problem** as a **nonsmooth system** are compared, both from the **theoretical and the practical point of view**. A **global algorithm** for solving the **Nonlinear Complementarity Problem** which uses the algorithms introduced here is also presented. **Some numerical experiments** show a **good performance** of this algorithm.

# Notação

A notação que será usada ao longo da tese é a seguinte:

$\| \cdot \|$  : Uma norma arbitrária associada a um produto escalar  $\langle , \rangle$ .

$F'(x)$  : Matriz Jacobiana de  $F$  em  $x$ .

$e_i$  :  $i$ -ésimo vetor coordenado unitário de  $\mathbf{R}^n$ .

$f'_i(x)$  :  $(\frac{\partial f_i}{\partial x_1}(x), \dots, \frac{\partial f_i}{\partial x_n}(x)) = \nabla^t f_i(x)$ .

$B(x, \epsilon)$  : Bola de centro  $x$  e raio  $\epsilon$ .

$V(x, y)$  : Subespaço linear gerado por  $x$  e  $y$ , contido num espaço de dimensão finita  $X$ .

$\| \cdot \|_{xy}$  : Norma sobre  $X$  relativa ao produto escalar  $\langle , \rangle_{xy}$ .

$P_{xy}$  : Projeção ortogonal sobre  $V(x, y)$  com respeito à norma  $\| \cdot \|_{xy}$ .

$D_G$  : Conjunto de pontos onde a função  $G$  é diferenciável.

$\partial_B G(x^*)$  : B-Jacobiano generalizado da função  $G$  no vetor  $x^*$ , o qual é definido por:  $\{ \lim_{x^k \rightarrow x} G'(x^k) : x^k \in D_G \}$ .

# Índice

<b>1</b>	<b>Introdução</b>	<b>2</b>
<b>2</b>	<b>Reformulação do PCN usando a função Mínimo</b>	<b>5</b>
2.1	Hipóteses Locais . . . . .	9
2.2	Resultados de Convergência . . . . .	9
2.3	Família LCSU para $G(x) = 0$ . . . . .	18
<b>3</b>	<b>Reformulação do PCN usando a função de Fischer</b>	<b>28</b>
3.1	Preliminares . . . . .	30
3.2	Hipóteses Locais . . . . .	34
3.3	Resultados de Convergência . . . . .	34
3.4	Família LCSU para $\Phi(x) = 0$ . . . . .	48
<b>4</b>	<b>A hipótese de não singularidade</b>	<b>54</b>

*Introdução* 2

**5 Testes numéricos 66**

5.1 Sensibilidade em soluções degeneradas . . . . . 67

5.2 Comportamento local dos algoritmos 2.1 e 3.1 . . . . . 70

**6 Estratégia de globalização 73**

6.1 Iterações ordinárias e especiais . . . . . 74

6.2 Algoritmo híbrido . . . . . 76

6.3 Testes numéricos . . . . . 77

6.4 Uma aplicação . . . . . 79

**7 Comentários finais 84**

# Capítulo 1

## Introdução

Dada uma função  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $F(x) = (f_1(x), \dots, f_n(x))^T$ ,  $F \in C^1(\mathbb{R}^n)$ , o Problema de Complementaridade Não Linear (PCN) consiste em encontrar um vetor  $x \in \mathbb{R}^n$  tal que

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0.$$

Relacionados com este problema estão os de inequações variacionais, complementaridade linear, complementaridade mista, complementaridade horizontal e outros. Este tipo de problema tem muitas aplicações em Física, Engenharia e Economia ([5], [7], [18]). A importância do PCN nestas áreas deve-se ao fato do conceito de complementaridade ser sinônimo da noção de sistema em equilíbrio.

Nos últimos tempos tem havido uma intensa atividade no sentido de encontrar métodos Newtonianos eficientes para o PCN. Por um lado, uma investigação orientou-se no sentido de encontrar funções de mérito eficientes cujos minimizadores coincidam com as soluções do PCN. Por outro lado, na procura de métodos com convergência local de ordem alta ([11]).

Uma outra forma bem conhecida de tratar com o PCN é a reformulação do mesmo como um sistema de equações não lineares não diferenciáveis ([19], [23]). Neste trabalho usamos duas reformulações. A primeira é baseada na clássica

função Mínimo ([19]), definida por

$$G : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$x \longmapsto G(x) = \begin{pmatrix} \min \{x_1, f_1(x)\} \\ \vdots \\ \min \{x_n, f_n(x)\} \end{pmatrix}.$$

Observe que, se existe  $i$  tal que  $x_i = f_i(x)$ , a função  $G$  pode não ser diferenciável. É fácil verificar que resolver o PCN é equivalente a resolver o sistema de equações não lineares:

$$G(x) = 0. \quad (1.1)$$

A segunda reformulação é baseada na função Fischer-Burmeister (Ver [6], [2], [4]), definida por :

$$\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$x \longmapsto \Phi(x) = \begin{pmatrix} \phi(x_1, f_1(x)) \\ \vdots \\ \phi(x_n, f_n(x)) \end{pmatrix}$$

onde,  $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}$  é dada por

$$\phi(x, y) = \|(x, y)\|_2 - x - y.$$

Esta função é tal que  $\phi(x, y) = 0$  se e somente se  $x \geq 0, y \geq 0, xy = 0$ . Assim, é óbvio que o PCN é equivalente ao sistema não linear

$$\Phi(x) = 0. \quad (1.2)$$

Se existe  $i$  tal que  $x_i = f_i(x) = 0$ , a função  $\Phi$  não é diferenciável em  $x$ .

Chamaremos a função  $G$  “função Mínimo”, enquanto  $\Phi$  será chamada “função Fischer” e assim (1.1) e (1.2) serão chamadas reformulações do PCN usando as funções Mínimo e Fischer, respectivamente.

Neste trabalho, desenvolvemos e analisamos (tanto na forma teórica como prática) métodos para resolver o PCN usando os sistemas (1.1) e (1.2). Em ambos os casos desenvolvemos uma família de métodos **Least Change Secant Update LCSU**, seguindo a teoria desenvolvida em ([15]). Para tais famílias provamos convergência superlinear sob hipóteses apropriadas. Apresentamos ainda um algoritmo global para resolver o PCN que aproveita o comportamento local da função Mínimo e o comportamento global da função Fischer. Alguns experimentos numéricos mostram o bom desempenho deste algoritmo.

Este trabalho é organizado como segue.

No **Capítulo 2**, sob certas hipóteses desenvolvemos a teoria **LCSU** para a reformulação do PCN baseada na função Mínimo (1.1).

No **Capítulo 3** estabelecemos hipóteses semelhantes às dadas no **Capítulo 2**, sob as quais desenvolvemos a teoria **LCSU** para a reformulação do PCN baseada na função Fischer (1.2).

O **Capítulo 4** é dedicado a mostrar que a hipótese de não singularidade dos elementos em  $\partial_B G(x^*)$  feita no **Capítulo 2**, não é equivalente à não singularidade dos elementos de um subconjunto  $Z_*$  de  $\partial_B \Phi(x^*)$  feita no **Capítulo 3**.

No **Capítulo 5** apresentamos alguns resultados numéricos que mostram a sensibilidade da função Fischer a degenerações na solução do PCN. Também apresentamos o desempenho numérico dos algoritmos propostos nos **Capítulos 2 e 3** respectivamente, quando são aplicados a 16 problemas teste dados em ([14]).

No **Capítulo 6** apresentamos uma estratégia de globalização para resolver o PCN a qual usa idéias do algoritmo híbrido dado em ([8]), usando como algoritmo global o dado em ([2]). Apresentamos alguns resultados numéricos os quais mostram o bom desempenho de nosso algoritmo. Estes resultados incluem a aplicação de nosso algoritmo híbrido a um problema de valor de capital ótimo invariante.

No **Capítulo 7** fazemos alguns comentários finais e apresentamos propostas de trabalhos futuros.

## Capítulo 2

# Reformulação do PCN usando a função Mínimo

É fácil ver que o PCN é equivalente ao problema de encontrar uma solução do sistema de equações não lineares  $G(x) = 0$ , onde a função  $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$  é dada por:

$$G(x) = \begin{pmatrix} \min \{x_1, f_1(x)\} \\ \vdots \\ \min \{x_n, f_n(x)\} \end{pmatrix}. \quad (2.1)$$

Nos pontos  $x$  onde existe  $i$  tal que  $x_i = f_i(x)$ , a função  $G$  pode ser não diferenciável.

O objetivo deste capítulo é definir e analisar métodos quase-Newton para resolver o PCN usando o sistema  $G(x) = 0$ , com  $G$  como em (2.1).

Dado  $x^0 \in \mathbb{R}^n$  uma aproximação inicial, o algoritmo genérico estudado neste capítulo é dado por

$$x^{k+1} = x^k - B_k^{-1} G(x^k), \quad (2.2)$$

onde,

$$B_k = \begin{pmatrix} [B^k]_1 \\ \vdots \\ [B^k]_n \end{pmatrix}, \quad (2.3)$$

com  $[B^k]_i \in \{[A^k]_i, \mathbf{e}_i\}$ , sendo que

$$[B^k]_i = \begin{cases} \mathbf{e}_i^T & \text{se } x_i^k < f_i(x^k), \\ [A^k]_i & \text{se } x_i^k > f_i(x^k), \\ \mathbf{e}_i^T \text{ ou } [A^k]_i & \text{se } x_i^k = f_i(x^k) = 0, \end{cases}$$

$\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  é a base canônica de  $\mathbb{R}^n$  e  $A_k = \begin{pmatrix} [A^k]_1 \\ \vdots \\ [A^k]_n \end{pmatrix}$ , é uma aproximação qualquer da matriz Jacobiana de  $F$  em  $x^k$ .

Para o sistema  $G(x) = 0$  desenvolvemos uma família do tipo **Least Change Secant Update LCSU**, a qual é gerada de acordo com as regras introduzidas por Martínez ([15]). Além disso, provamos que a família LCSU fornece algoritmos localmente convergentes e superlineares.

### Observações:

A escolha de  $[B^k]_i$ , sendo  $\mathbf{e}_i^T$  ou  $[A^k]_i$ , depende de

$$\min\{x_i^k, f_i(x^k)\} = x_i \text{ ou } \min\{x_i^k, f_i(x^k)\} = f_i(x^k),$$

respectivamente. Agora, se  $f_i(x^k) = x_i^k$  então

$$\min\{x_i^k, f_i(x^k)\} = x_i^k = f_i(x^k)$$

e portanto qualquer das duas escolhas serve para  $[B^k]_i$ .

Uma solução  $x^*$  do PCN é chamada de degenerada, se para algum  $i$ ,  $1 \leq i \leq n$

$$x_i^* = f_i(x^*) = 0.$$

em outro caso, é chamada de não degenerada.

Associada a cada  $B_k$  definimos uma matriz  $B_*$  da seguinte forma:

$$B_* = \begin{pmatrix} [B^*]_1 \\ \vdots \\ [B^*]_n \end{pmatrix}$$

onde, para  $i = 1, \dots, n$ ,

$$[B^*]_i = \begin{cases} e_i^T & \text{se } x_i^* < f_i(x^*), \\ f_i'(x^*) & \text{se } x_i^* > f_i(x^*), \\ e_i^T \text{ ou } f_i'(x^*) & \text{se } x_i^* = f_i(x^*) = 0, \end{cases}$$

Como existem no máximo  $2^m$  matrizes nessas condições, onde  $m$  é o número de componentes degeneradas em  $x^*$ , podemos definir  $\theta > 0$  que majore  $\|B_*^{-1}\|$  para todas elas.

Qualquer matriz  $B_*$ , na hipótese **A3**, pode ser vista como a matriz Jacobiana de um sistema diferenciável de equações não lineares. Para ver isto, suponha que  $x^*$  seja uma solução do PCN<sup>1</sup>:

Se a solução  $x^*$  é não degenerada, i.e., se para todo  $1 \leq i \leq n$

$$x_i^* \neq 0 \text{ ou } f_i(x^*) \neq 0,$$

a matriz  $B_*$  considerada na hipótese **A3** pode ser vista como a matriz Jacobiana  $G'(x^*)$ .

Se a solução  $x^*$  é degenerada, vamos supor que existe só uma degeneração sem perda de generalidade, e que ela ocorra em  $x_1^*$ , i.e.,

$$\begin{aligned} x_1^* &= f_1(x^*) = 0 \\ x_i^* &> f_i(x^*) = 0, \quad i = 2, \dots, n. \end{aligned} \tag{2.4}$$

---

<sup>1</sup>Portanto,  $x^*$  é solução do  $G(x) = 0$ .

Para as componentes não degeneradas consideraremos, sem perda de generalidade, a situação,  $x_i^* > f_i(x^*)$ , no outro caso a análise é análoga.

Observe, de (2.4), que  $x^*$  pode ser vista como a solução de dois sistemas não lineares

$$G_1(x) = 0 \quad \text{e} \quad G_2(x) = 0,$$

onde  $G_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $i = 1, 2$ , são duas funções definidas por:

$$G_1(x) = \begin{pmatrix} x_1 \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix}, \quad G_2(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix},$$

cujas matrizes Jacobianas na solução  $x^*$ , têm a forma da matriz  $B_*$ , considerada na hipótese **A3**.

No caso geral de  $m$  degenerações, ( $m \leq n$ ), i.e.,

$$\begin{aligned} x_i^* &= f_i(x^*) = 0 & i &= 1, \dots, m \\ x_i^* &> f_i(x^*) = 0, & i &= m+1, \dots, n, \end{aligned}$$

o vetor  $x^*$  pode ser visto como a solução de  $2^m$  sistemas de equações não lineares

$$G_i(x) = 0$$

onde, as  $G_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $i = 1, \dots, 2^m$ , são de classe  $C^1$ . Além disso:

- Para algum  $k = 1, \dots, 2^m$ ,  $G_k = F$ , e portanto a matriz Jacobiana da  $G_k$  em  $x$  coincide com a matriz Jacobiana de  $F$  em  $x$ .
- Para todo  $i = 1, \dots, 2^m$ ,  $i \neq k$ , as matrizes Jacobianas das  $G_i$  têm pelo menos uma linha da matriz identidade.

Assim, as matrizes Jacobianas das  $G_i$  numa solução degenerada  $x^*$ , têm a forma da matriz  $B_*$  considerada na hipótese **A3**.

Sob as seguintes hipóteses provaremos que as seqüências geradas pelo algoritmo quase-Newton genérico (2.2), estão bem definidas e convergem linearmente à solução de  $G(x) = 0$ .

## 2.1 Hipóteses Locais

**A1.**  $x^* \in \mathbb{R}^n$  é tal que  $G(x^*) = 0$ .

**A2.** Existem  $\gamma, \epsilon > 0$ , tais que:

$$\|F'(x) - F'(x^*)\| \leq \gamma \|x - x^*\|,$$

para todo  $x \in \mathcal{B}(x^*, \epsilon)$ .

**A3.** As matrizes  $B_*$  são não singulares.

## 2.2 Resultados de Convergência

**Teorema 2.1 :** (Das duas vizinhanças) *Dado qualquer  $r \in (0, 1)$  existem  $\epsilon, \delta > 0$  tais que,*

$$\text{se } \|x^0 - x^*\| \leq \epsilon \text{ e } \|A_k - F'(x^*)\| \leq \delta, \text{ para todo } k,$$

*qualquer seqüência  $\{x^k\}$  gerada por*

$$x^{k+1} = x^k - B_k^{-1} G(x^k),$$

*está bem definida, converge para  $x^*$  e satisfaz*

$$\|x^{k+1} - x^*\| \leq r \|x^k - x^*\|, \text{ para todo } k = 0, 1, 2, \dots$$

O seguinte **Lema** será útil na prova do Teorema das duas vizinhanças.

**Lema 2.1 :** *Para cada  $x \in \mathbb{R}^n, A \in \mathbb{R}^{n \times n}$ , seja*

$$\Gamma(x, A) = x - B^{-1}G(x), \quad (2.5)$$

onde

$$B = \begin{pmatrix} [B]_1 \\ \vdots \\ [B]_n \end{pmatrix}, \text{ com } [B]_i \in \{[A]_i, \mathbf{e}_i^T\}, \quad (2.6)$$

sendo que

$$[B]_i = \begin{cases} \mathbf{e}_i & \text{se } x_i < f_i(x), \\ [A]_i & \text{se } x_i > f_i(x). \end{cases} \quad (2.7)$$

Seja  $r \in (0, 1)$ . Então existem  $\epsilon_1$  e  $\delta_1$  tais que,

$$\text{se } \|x - x^*\| < \epsilon_1 \text{ e } \|A - F'(x^*)\| \leq \delta_1,$$

a função  $\Gamma(x, A)$  está bem definida, e satisfaz

$$\|\Gamma(x, A) - x^*\| \leq r \|x - x^*\|.$$

**Prova :**

Nesta prova se usa,  $\|\cdot\| = \|\cdot\|_\infty$ . Seja  $\epsilon_1 > 0$  tal que para todo  $i = 1, \dots, n$ ,

$$\text{se } f_i(x^*) > x_i^* \text{ então } f_i(x) > x_i,$$

$$\text{se } f_i(x^*) < x_i^* \text{ então } f_i(x) < x_i,$$

para todo  $x \in \mathcal{B}(x^*, \epsilon_1)$ . Claramente  $\epsilon_1$  existe pela continuidade da  $F$ .

Seja  $\delta_1 \leq \frac{r}{4\theta}$ , onde  $\theta$  é dado na hipótese **A3**, para  $\|\cdot\|_\infty$ .

Pela definição de  $\|\cdot\|_\infty$ , se  $x \in \mathcal{B}(x^*, \epsilon_1)$ ,  $A \in \mathcal{B}(F'(x^*), \delta_1)$ ,  $B$  é matriz associada a  $A$  pela regra (2.6)-(2.7), e  $B_* \in \partial_B G(x^*)$  é matriz associada a  $F'(x^*)$  pela mesma regra, temos que:

$$\|B - B_*\| \leq \delta_1,$$

logo, pelo **Lema de Banach**,  $B^{-1}$  existe e além disso,

$$\|B^{-1}\| \leq 2 \|B_*^{-1}\| \leq 2\theta.$$

Subtraindo  $x^*$  de (2.5) e aplicando o operador norma  $\|\cdot\|$  tem-se:

$$\begin{aligned} \|\Gamma(x, A) - x^*\| &= \|(x - x^*) - B^{-1}G(x) + B^{-1}B_*(x - x^*) - B^{-1}B_*(x - x^*)\| \\ &= \|(I - B^{-1}B^*)(x - x^*) - B^{-1}[G(x) - G(x^*) - B_*(x - x^*)]\| \\ &\leq \|B^{-1}\| [\|B - B^*\| \|x - x^*\| + \|G(x) - G(x^*) - B_*(x - x^*)\|] \\ &\leq 2\theta \left[ \delta_1 + \frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \right] \|x - x^*\| \end{aligned}$$

Mas,

$$G(x) - G(x^*) = \begin{pmatrix} g_1(x) - g_1(x^*) \\ \vdots \\ g_n(x) - g_n(x^*) \end{pmatrix}, \quad (2.8)$$

onde  $g_i(x) - g_i(x^*) \in \{x_i - x_i^*, f_i(x) - f_i(x^*)\}$ ,

$$g_i(x) - g_i(x^*) = \begin{cases} x_i - x_i^* & \text{se } x_i^* < f_i(x^*), \\ f_i(x) - f_i(x^*) & \text{se } x_i^* > f_i(x^*), \end{cases} \quad (2.9)$$

e

$$[B_*(x - x^*)]_i = \begin{cases} x_i - x_i^* & \text{se } x_i^* < f_i(x^*), \\ f'_i(x^*)(x - x^*) & \text{se } x_i^* > f_i(x^*). \end{cases} \quad (2.10)$$

Se  $x_i^* = f_i(x^*)$  qualquer das escolhas, tanto em (2.9) quanto em (2.10), serve, dado que neste caso  $\min\{x_i^*, f_i(x^*)\} = x_i^* = f_i(x^*) = 0$ .

Assim, de (2.7)-(2.9) e (2.10):

$$\frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} = \frac{\|\bar{F}(x)\|}{\|x - x^*\|} = \frac{\left\| \begin{pmatrix} \bar{f}_1(x) \\ \vdots \\ \bar{f}_n(x) \end{pmatrix} \right\|}{\|x - x^*\|},$$

onde,

$$\bar{f}_i(x) = \begin{cases} 0 & \text{se } x_i^* < f_i(x^*), \\ f_i(x) - f_i(x^*) - f'_i(x^*)(x - x^*) & \text{se } x_i^* > f_i(x^*). \end{cases}$$

Usando a definição da  $\|\cdot\|_\infty$  :

$$\begin{aligned} \frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} &= \frac{\|\bar{F}(x)\|}{\|x - x^*\|} \\ &= \max_{1 \leq i \leq n} \frac{|f_i(x) - f_i(x^*) - f'_i(x^*)(x - x^*)|}{\|x - x^*\|} \end{aligned}$$

Seja  $1 \leq k \leq n$  tal que:

$$\frac{|f_k(x) - f_k(x^*) - f'_k(x^*)(x - x^*)|}{\|x - x^*\|} = \max_{1 \leq i \leq n} \frac{|f_i(x) - f_i(x^*) - f'_i(x^*)(x - x^*)|}{\|x - x^*\|}$$

logo,

$$\frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} = \frac{|f_k(x) - f_k(x^*) - f'_k(x^*)(x - x^*)|}{\|x - x^*\|}$$

e pela diferenciabilidade da  $F$ , para qualquer  $\rho > 0$  existe  $\epsilon > 0$  tal que

$$\|x - x^*\| < \epsilon \Rightarrow \frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \leq \rho.$$

Em particular, dado  $\rho = \frac{r}{4\theta}$ , existe  $\epsilon_r > 0$  tal que

$$\|x - x^*\| < \epsilon_r \Rightarrow \frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \leq \rho.$$

Portanto:

$$\begin{aligned} \|\Gamma(x, A) - x^*\| &\leq 2\theta \left[ \frac{r}{4\theta} + \frac{r}{4\theta} \right] \|x - x^*\| \\ &= r \|x - x^*\|. \end{aligned}$$

Portanto, o resultado está provado. ■

### Prova do Teorema 2.1 :

Considerando a função  $\Gamma(x, A) = x - B^{-1}G(x)$  definida em (2.5), onde  $B$  é gerada por (2.6)-(2.7), tem-se

$$x^{k+1} = \Gamma(x^k, A_k), \quad k = 0, 1, 2, \dots$$

Sejam  $\epsilon_1$  e  $\delta_1$  os definidos no **Lema 2.1** e suponha  $\epsilon \in (0, \epsilon_1)$ ,  $\delta \in (0, \delta_1)$ .

Para  $k = 0$ ,

$$\text{se } \|x^0 - x^*\| \leq \epsilon \leq \epsilon_1 \text{ e } \|A_0 - F'(x^*)\| \leq \delta \leq \delta_1,$$

pelo **Lema 2.1**,  $x^1$  está bem definido e satisfaz

$$\|x^1 - x^*\| = \|\Gamma(x^0, A_0) - x^*\| \leq r \|x^0 - x^*\|.$$

Portanto, para  $k = 0$  a tese do **Teorema 2.1** é válida.

Suponha a hipótese indutiva para  $k - 1$ , i.e.,

$$\text{se } \|x^{k-1} - x^*\| \leq \epsilon \quad \text{como } \|A_{k-1} - F'(x^*)\| \leq \delta,$$

então  $x^k$  está bem definido e satisfaz

$$\|x^k - x^*\| = \|\Gamma(x^{k-1}, A_{k-1}) - x^*\| \leq r\|x^{k-1} - x^*\|.$$

Mas, da última desigualdade tem-se:

$$\|x^k - x^*\| \leq r\|x^{k-1} - x^*\| \leq r^k \epsilon < \epsilon < \epsilon_1,$$

Assim, pelo **Lema 2.1**,  $x^{k+1}$  está bem definido e satisfaz

$$\|x^{k+1} - x^*\| \leq r\|x^k - x^*\|,$$

com o qual se conclui a demonstração do **Teorema 2.1**. ■

Observe que na prova do **Teorema 2.1** usamos a norma infinito. Mas, se  $e_k = \|x^k - x^*\|$  é o erro em qualquer outra norma, então  $e_k \leq C r^k e^0$ , onde  $C$  é uma constante positiva que não depende de  $k$  e  $r$  é como no **Teorema 2.1**. Assim, temos a convergência  $r$ -linear em qualquer outra norma.

**Teorema 2.2:** *Suponha as hipóteses **A1**, **A2** e **A3** e que para algum  $x^0$  a seqüência de pontos gerada por*

$$x^{k+1} = x^k - B_k^{-1} G(x^k), \tag{2.11}$$

onde  $B_k$  é gerada como em (2.6), satisfaz  $\lim_{k \rightarrow \infty} x^k = x^*$ . Definimos  $A_k$ , como em (2.3) e  $s^k = x^{k+1} - x^k$ .

Se

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - F'(x^*)) s^k\|}{\|s^k\|} = 0$$

então a seqüência  $\{x^k\}$  converge superlinearmente a  $x^*$ .

O **Lema 2.2** será útil na prova do **Teorema 2.2**.

**Lema 2.2**: *Suponha as hipóteses A1, A2, e A3. Então existem  $\epsilon, \beta > 0$  tais que, se  $\|x - x^*\| < \epsilon$  então*

$$\|G(x)\| \geq \beta \|x - x^*\|.$$

**Prova** :

$$\begin{aligned} \|x - x^*\| &= \|(B_*)^{-1} B_*(x - x^*)\| \\ &\leq \|B_*^{-1}\| \|B_*(x - x^*)\| \end{aligned} \quad (2.12)$$

Pelo que foi feito na demonstração do **Lema 2.1**, para qualquer  $\rho > 0$  existe  $\epsilon > 0$  tal que se  $\|x - x^*\| < \epsilon$  então

$$\frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \leq \rho. \quad (2.13)$$

Assim ,

$$\lim_{x \rightarrow x^*} \frac{\|G(x) - G(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} = 0.$$

Por (2.12) e usando a hipótese **A1**:

$$\lim_{x \rightarrow x^*} \frac{\|G(x) - B_*(x - x^*)\|}{\|B_*^{-1}\| \|B_*(x - x^*)\|} = 0.$$

Logo,

$$\lim_{x \rightarrow x^*} \frac{\|G(x) - B_*(x - x^*)\|}{\|B_*(x - x^*)\|} = 0.$$

Mas,

$$\| \|G(x)\| - \|B_*(x - x^*)\| \| \leq \|G(x) - B_*(x - x^*)\|$$

e de (2.13), dado  $\rho = \frac{1}{2}$  existe  $\epsilon > 0$  tal que, se  $0 < \|x - x^*\| \leq \epsilon$  então

$$-\frac{1}{2} \leq \frac{\|G(x)\| - \|B_*(x - x^*)\|}{\|B_*(x - x^*)\|} \leq \frac{1}{2}, \quad (2.14)$$

ou seja,

$$\|G(x)\| \geq \frac{1}{2} \|B_*(x - x^*)\| \geq \frac{1}{2} \theta \|x - x^*\|, \quad (2.15)$$

onde, a última desigualdade obtém-se de (2.14) e da hipótese **A3**. Portanto, a tese do **lema 2.2** segue de (2.15), com  $\beta = \frac{1}{2} \theta$ . ■

Na prova do **Lema 2.2** usamos a norma infinito, mas o resultado é válido para qualquer outra norma, com uma mudança apropriada na constante  $\beta$ .

No seguinte **Teorema** provamos que, para a reformulação do PCN por meio de  $G(x) = 0$ , a condição do tipo Dennis-Moré-Walker garante convergência superlinear. Nesta prova usamos a norma infinito, mas lembremos que os resultados de convergência superlinear são independentes da norma.

### Prova do Teorema 2:

Sem perda de generalidade suponha que  $x^k \in \mathcal{B}(x^*, \epsilon_1)$ , para todo  $k$  e que  $\|\cdot\| = \|\cdot\|_\infty$ .

Seja

$$A_k = \begin{pmatrix} [A^k]_1 \\ \vdots \\ [A^k]_n \end{pmatrix};$$

com esta notação,

$$\|(A_k - F'(x^*)) s^k\| = \left\| \begin{pmatrix} [A^k]_1 - f'_1(x^*) \\ \vdots \\ [A^k]_n - f'_n(x^*) \end{pmatrix} s^k \right\|. \quad (2.16)$$

De (2.3) e pela hipótese **A3**, para cada  $B_k$  existe uma  $B_*$  tal que

$$\|(B_k - B_*) s^k\| = \left\| \begin{pmatrix} [B^k]_1 - [B^*]_1 \\ \vdots \\ [B^k]_n - [B^*]_n \end{pmatrix} s^k \right\|, \quad (2.17)$$

onde,

$$[B^k]_i - [B^*]_i = \begin{cases} 0 & \text{se } x_i^* < f_i(x^*), \\ [A^k]_i - f'_i(x^*) & \text{se } x_i^* > f_i(x^*), \end{cases}$$

Assim, de (2.16) e (2.17),

$$\|(B_k - B_*) s^k\| \leq \|(A_k - F'(x^*)) s^k\|.$$

De (2.11), tem-se

$$\begin{aligned} 0 &= B_k s^k + G(x^k) \\ -G(x^{k+1}) &= (B_k - B_*) s^k - G(x^{k+1}) + G(x^k) + B_* s^k \end{aligned} \quad (2.18)$$

Multiplicando (2.18) por  $\|s^k\|^{-1}$  e aplicando o operador norma  $\|\cdot\|$  tem-se:

$$\begin{aligned} \frac{\|G(x^{k+1})\|}{\|s^k\|} &\leq \frac{\|(B_k - B_*) s^k\|}{\|s^k\|} + \frac{\|-G(x^{k+1}) + G(x^k) + B_* s^k\|}{\|s^k\|} \\ &\leq \frac{\|(A_k - F'(x^*)) s^k\|}{\|s^k\|} + \frac{\|-G(x^{k+1}) + G(x^k) + B_* s^k\|}{\|s^k\|} \\ &\leq \frac{\|(A_k - F'(x^*)) s^k\|}{\|s^k\|} + \gamma \max\{\|x^{k+1} - x^*\|, \|x^k - x^*\|\}. \end{aligned}$$

Usando a hipótese de que:

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - F'(x^*)) s^k\|}{\|s^k\|} = 0$$

e

$$\lim_{k \rightarrow \infty} (x^k - x^*) = 0,$$

tem-se,

$$\lim_{k \rightarrow \infty} \frac{\|G(x^{k+1})\|}{\|s^k\|} = 0.$$

Mas, pelo **Lema 2.2**, existe uma constante positiva  $\beta$  tal que:

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\|G(x^{k+1})\|}{\|s^k\|} &\geq \lim_{k \rightarrow \infty} \beta \frac{\|x^{k+1} - x^*\|}{\|s^k\|} \\ 0 &\geq \lim_{k \rightarrow \infty} \beta \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\| + \|x^{k+1} - x^*\|} \\ &= \lim_{k \rightarrow \infty} \beta \frac{\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}}{1 + \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}}, \end{aligned}$$

o que implica

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

e portanto  $\{x^k\}$  converge a  $x^*$  superlinearmente. ■

### 2.3 Família LCSU para $G(x) = 0$ .

Na maior parte de algoritmos práticos a matriz  $A_{k+1}$  é calculada depois do cálculo de  $x^{k+1}$ , usando  $A_k$  e a informação disponível de  $x^k, x^{k+1}$ . Este é o caso de uma importante sub-classe de métodos quase-Newton: os métodos **Least Change Secant Update, (LCSU)** para resolver sistemas de equações não lineares. Nestes métodos, a atualização de  $A_k$  se faz minimizando alguma medida

razoável da mudança da aproximação da Jacobiana atual sujeita a que a nova aproximação satisfaça uma condição secante (e quiçá outras propriedades).

Alguns exemplos de atualizações LCSU são:

**Broyden**

$$A_{k+1} = A_k + \frac{(y^k - A_k s^k)}{(s^k)^t s^k} s^k,$$

**BFGS**

$$A_{k+1} = A_k + \frac{y^k (y^k)^t}{(y^k)^t s^k} - \frac{A_k s^k (s^k)^t A_k}{(s^k)^t A_k s^k},$$

onde,  $s^k = x^{k+1} - x^k, y^k = F(x^{k+1}) - F(x^k)$ .

Martínez ([15]), desenvolveu uma teoria de convergência local para métodos tipo **LCSU** que inclui a teoria de Dennis e Walker ([3]) e cujas regras serão usadas nesta seção, para gerar a família do tipo **LCSU** para o sistema  $G(x) = 0$ , com  $G$  dada por (2.1). Provaremos que esta família **LCSU** fornece algoritmos localmente convergentes e superlineares.

**Algoritmo 2.1**

Suponha que  $x^0$  e  $A_0$  são arbitrários. Para  $k = 1, 2 \dots, x^{k+1}$  e  $A_{k+1}$  são gerados como segue:

$$\begin{aligned} x^{k+1} &= x^k - B_k^{-1} G(x^k) \\ A_{k+1} &= P_{x^k, x^{k+1}}(A_k) \end{aligned} \tag{2.19}$$

Além das hipóteses gerais **A1**, **A2**, e **A3**, suponha como em ([15]) que:

**A4.** Existe  $\alpha_1 > 0$  tal que, para todo  $x, y \in \mathbb{R}^n$ , existe  $A \in V(x, y)$  que satisfaz

$$\|A - F'(x^*)\| \leq \alpha_1 \sigma(x, y), \tag{2.20}$$

onde,  $\sigma(x, y) = \max\{\|x - x^*\|, \|y - x^*\|\}$ .

**A5.** Existe  $\alpha_2 > 0$  tal que, para todo  $x, y \in \mathbb{R}^n$ ,  $A \in V(x, y)$

$$\|A\|_{x,y} \leq [1 + \alpha_2 \sigma(x, y)] \|A\|, \quad (2.21)$$

$$\|A\| \leq [1 + \alpha_2 \sigma(x, y)] \|A\|_{x,y} \quad (2.22)$$

Martínez ([15]) demonstrou que uma matriz  $A$  gerada pela regra (2.19) e que satisfaz as hipóteses **A4** e **A5**, pode deteriorar-se mas isto ocorre de forma controlada, o que é garantido pelos seguintes resultados:

**Lema 2.3:** *Suponha as hipóteses A1-A5. Existem  $\alpha_3, \alpha_4 > 0$  tais que para todo  $x, y \in \mathcal{B}(x^*, \epsilon_1)$ ,  $A \in V(x, y)$ ,*

$$\|P_{x,y}(A) - F'(x^*)\| \leq [1 + \sigma(x, y)] \|A - F'(x^*)\| + \alpha_3 \sigma(x, y).$$

**Prova:** (Ver [15].)

**Corolário 2.1:** *Existe  $\alpha_5 > 0$  tal que*

$$\|P_{x,y}(A) - F'(x^*)\| \leq \|A - F'(x^*)\| + \alpha_5 \|x - x^*\|, \quad (2.23)$$

*sempre que  $x, y \in \mathcal{B}(x^*, \epsilon_1)$ ,  $A \in \mathcal{B}(F'(x^*), \delta_1)$  e  $\|y - x^*\| \leq \|x - x^*\|$ .*

**Prova:** (Ver [15].)

Estes dois resultados, além das hipóteses **A1**, **A2** e **A3**, são fundamentais para garantir o seguinte Teorema:

**Teorema 2.3:** *Suponha as hipóteses A1-A5 e que a seqüência  $\{A_k\}$  está definida por (2.19). Dado  $r \in (0, 1)$  existem  $\bar{\epsilon}$  e  $\bar{\delta}$  tais que,*

$$\text{se } \|x^0 - x^*\| \leq \bar{\epsilon} \text{ e } \|A_0 - F'(x^*)\| \leq \bar{\delta},$$

*qualquer seqüência  $x_k$  gerada por*

$$x^{k+1} = x^k - B_k^{-1} G(x^k)$$

está bem definida, converge a  $x^*$  e para todo  $k = 0, 1, 2, \dots$

$$\|x^{k+1} - x^*\| \leq r \|x^k - x^*\|.$$

Além disso, para todo  $k, j = 0, 1, 2, \dots$  existem constantes positivas  $\alpha_6$  e  $\alpha_7$  tais que:

$$\|A_{k+j} - F'(x^*)\| \leq \|A_k - F'(x^*)\| + \alpha_6 \|x^k - x^*\|. \quad (2.24)$$

$$\|A_{k+j} - F'(x^*)\|^2 \leq \|A_k - F'(x^*)\|^2 + \alpha_7 \|x^k - x^*\| \quad (2.25)$$

### Prova:

Suponha  $\bar{\epsilon} \in (0, \epsilon_1)$ ,  $\bar{\delta} \in (0, \delta_1)$ , com  $\epsilon_1$  e  $\delta_1$  os definidos no **Lema 2.1**, e

$$\bar{\delta} + \alpha_5 \frac{\bar{\epsilon}}{1-r} < \delta_1$$

A prova é feita usando indução sobre  $k$ .

Para  $k = 0$ , se  $\|x^0 - x^*\| \leq \bar{\epsilon} \leq \epsilon_1$  e  $\|A_0 - F'(x^*)\| \leq \bar{\delta} \leq \delta_1$ . Pelo **Lema 2.1**,  $x^1$  está bem definido e satisfaz

$$\|x^1 - x^*\| \leq r \|x^0 - x^*\|.$$

Portanto, para  $k = 0$  a tese do teorema é verdadeira.

Suponha a hipótese indutiva para  $k - 1$ , i.e.,

$$\text{se } \|x^{k-1} - x^*\| \leq \bar{\epsilon} \text{ e } \|A_{k-1} - F'(x^*)\| \leq \bar{\delta},$$

então  $x^k$  está bem definido e satisfaz

$$\|x^k - x^*\| \leq r \|x^{k-1} - x^*\|.$$

Mas, da última desigualdade tem-se:

$$\|x^k - x^*\| \leq r \|x^{k-1} - x^*\| \leq r^k \bar{\epsilon} < \bar{\epsilon} < \epsilon_1,$$

e pelo **Corolário 2.1**,

$$\begin{aligned}
 \|A_k - F'(x^*)\| &\leq \|A_{k-1} - F'(x^*)\| + \alpha_5 \|x^{k-1} - x^*\| \\
 &\leq \bar{\delta} + \alpha_5 \bar{\epsilon} \sum_{j=0}^{k-1} r^j \\
 &\leq \bar{\delta} + \alpha_5 \bar{\epsilon} \sum_{j=0}^{\infty} r^j \\
 &\leq \bar{\delta} + \alpha_5 \frac{\bar{\epsilon}}{1-r} < \delta_1
 \end{aligned}$$

Assim, pelo **Lema 2.1**,  $x^{k+1}$  está bem definido e satisfaz

$$\|x^{k+1} - x^*\| \leq r \|x^k - x^*\|.$$

Provaremos por indução sobre  $j$  que

$$\|A_{k+j} - F'(x^*)\| \leq \|A_k - F'(x^*)\| + \alpha_5 \|x^k - x^*\| \sum_{m=0}^{j-1} r^m.$$

Para  $j = 0$  o resultado é imediato usando o fato  $\sum_{m=0}^{-1} = 0$ .

Suponha a desigualdade válida para  $j - 1$ . i.e.,

$$\|A_{k+j-1} - F'(x^*)\| \leq \|A_k - F'(x^*)\| + \alpha_5 \|x^k - x^*\| \sum_{m=0}^{j-2} r^m.$$

De (2.23), tem-se

$$\begin{aligned}
 \|A_{k+j} - F'(x^*)\| &\leq \|A_{k+j-1} - F'(x^*)\| + \alpha_5 \|x^{k+j-1} - x^*\| \\
 \|A_{k+j} - F'(x^*)\| &\leq \|A_{k+j-1} - F'(x^*)\| + \alpha_5 r^{j-1} \|x^k - x^*\|,
 \end{aligned}$$

Portanto,

$$\|A_{k+j} - F'(x^*)\| \leq \|A_k - F'(x^*)\| + \alpha_5 \|x^k - x^*\| \sum_{m=0}^{j-1} r^m.$$

Assim, (2.24) segue desta última desigualdade com

$$\alpha_6 = \frac{\alpha_5}{1-r},$$

e (2.25) se deduz de (2.24) usando a limitação de  $\|A_k - F'(x^*)\|$ . ■

**Teorema 2.4:** *Supondo as hipóteses do Teorema 2.3,*

$$\lim_{k \rightarrow \infty} \|A_{k+1} - A_k\| = 0. \quad (2.26)$$

**Prova:** É a mesma do **Teorema 3.3** em ([15].) No entanto, para maior clareza do texto a apresentamos no que segue. Para simplificar a notação,  $\|\cdot\|_k = \|\cdot\|_{x^k, x^{k+1}}$ .

Suponha que (2.26) não é verdadeira. Então existe um conjunto infinito  $K_1$  de índices tal que

$$\|A_{k+1} - A_k\| \geq \gamma > 0,$$

para todo  $k \in K_1$ . Portanto por (2.21),

$$[1 + \alpha_2 \sigma(x^k, x^{k+1})] \|A_{k+1} - A_k\|_k \geq \gamma,$$

para todo  $k \in K_1$ . Assim, para  $k$  suficientemente grande e  $k \in K_1$ ,

$$\|A_{k+1} - A_k\|_k \geq \frac{\gamma}{2}.$$

Logo,

$$\|A_{k+1} - A_k\|_k^2 \geq \frac{\gamma^2}{4} \quad (2.27)$$

para todo  $k$  em um conjunto infinito de índices  $K_2$ .

Sejam  $\hat{A}$  e  $\hat{A}_k$  as projeções de  $F'(x^*)$  sobre  $V(x^k, x^{k+1})$  com respeito às normas  $\|\cdot\|$  e  $\|\cdot\|_k$ , respectivamente.

De (2.20) e do **Teorema 2.3**, tem-se

$$\|\hat{A} - F'(x^*)\| \leq \alpha_1 \|x^k - x^*\|.$$

Assim, por (2.21),

$$\begin{aligned} \|\hat{A} - F'(x^*)\|_k &\leq (1 + \alpha_2 \|x^k - x^*\|) \alpha_1 \|x^k - x^*\| \\ &\leq (1 + \alpha_2 \epsilon) \alpha_1 \|x^k - x^*\| = \alpha_8 \|x^k - x^*\|, \end{aligned}$$

onde,  $\alpha_8 = (1 + \alpha_2 \epsilon) \alpha_1$ .

Portanto, pela definição de  $\hat{A}_k$ ,

$$\|\hat{A}_k - F'(x^*)\|_k \leq \alpha_8 \|x^k - x^*\|, \quad (2.28)$$

logo,

$$\|\hat{A}_k - F'(x^*)\|_k^2 \leq \alpha_8^2 \|x^k - x^*\|^2. \quad (2.29)$$

Seja  $k \in K_2$ . De (2.29) e pelo Teorema de Pitágoras tem-se:

$$\begin{aligned} \|A_{k+1} - F'(x^*)\|_k^2 &= \|A_{k+1} - \hat{A}_k\|_k^2 + \|\hat{A}_k - F'(x^*)\|_k^2 \\ &\leq \|A_{k+1} - \hat{A}_k\|_k^2 + \alpha_8^2 \|x^k - x^*\|^2 \\ &= \|A_k - \hat{A}_k\|_k^2 - \|A_{k+1} - A_k\|_k^2 + \alpha_8^2 \|x^k - x^*\|^2. \end{aligned}$$

Assim, por (2.27), (2.28), e (2.29),

$$\begin{aligned} \|A_{k+1} - F'(x^*)\|_k^2 &\leq \|A_k - \hat{A}_k\|_k^2 - \frac{\gamma^2}{4} + \alpha_8^2 \|x^k - x^*\|^2 \\ &\leq (\|A_k - F'(x^*)\|_k + \|F'(x^*) - \hat{A}_k\|_k)^2 - \frac{\gamma^2}{4} \\ &\quad + \alpha_8^2 \|x^k - x^*\|^2 \\ &\leq (\|A_k - F'(x^*)\|_k + \alpha_8 \|x^k - x^*\|)^2 - \frac{\gamma^2}{4} \\ &\quad + \alpha_8^2 \|x^k - x^*\|^2 \\ &\leq \|A_k - F'(x^*)\|_k + 2 \alpha_8^2 \|x^k - x^*\|^2 - \frac{\gamma^2}{4} + \\ &\quad 2 \alpha_8 (1 + \alpha_2 \|x^k - x^*\|) \|A_k - F'(x^*)\| \|x^k - x^*\| \\ &\leq \|A_k - F'(x^*)\|_k^2 + \alpha_9 \|x^k - x^*\| - \frac{\gamma^2}{4}, \end{aligned}$$

Onde,  $\alpha_9 = 2\alpha_8(1 + \alpha_2\epsilon)\delta + 2\alpha_8^2\epsilon$ .

Portanto, existe  $\bar{k}$  tal que para  $k \in K_2$  e  $k \geq \bar{k}$ ,

$$\|A_{k+1} - F'(x^*)\|_k^2 \leq \|A_k - F'(x^*)\|_k^2 - \frac{\gamma^2}{8}.$$

Assim, por (2.21), (2.22) e o **Teorema 2.3**, tem-se que para  $\bar{k}$  grande,

$$\begin{aligned} \|A_{k+1} - F'(x^*)\|^2 &\leq (1 + \alpha_2\|x^k - x^*\|)^2 \|A_{k+1} - F'(x^*)\|_k^2 \\ &\leq (1 + \alpha_2\|x^k - x^*\|)^2 \left( \|A_k - F'(x^*)\|_k^2 - \frac{\gamma^2}{8} \right) \\ &\leq (1 + \alpha_2\|x^k - x^*\|)^2 \left[ (1 + \right. \\ &\quad \left. \alpha_2\|x^k - x^*\|)^2 \|A_k - F'(x^*)\|^2 - \frac{\gamma^2}{8} \right] \\ &\leq \|A_k - F'(x^*)\| - \frac{\gamma^2}{16}, \end{aligned} \tag{2.30}$$

para  $k \in K_2, k \geq \bar{k}$ .

Seja  $k_0 \geq \bar{k}$  tal que, para todo  $k \geq k_0$ ,

$$\alpha_7 \|x^k - x^*\| \leq \frac{\gamma^2}{32}, \tag{2.31}$$

onde  $\alpha_7$  é definido por (2.25).

Defina:

$$K_3 = \{k \in K_2 : k \geq k_0\} = \{k_1, k_2, k_3, \dots\}, \quad k_i < k_{i+1}, \quad i = 1, 2, 3, \dots$$

Então para todo  $j = 1, 2, 3, \dots$ , tem-se, por (2.29),

$$\|A_{k_{j+1}} - F'(x^*)\|^2 \leq \|A_{k_j} - F'(x^*)\|^2 - \frac{\gamma^2}{16}. \tag{2.32}$$

Por (2.25), (2.31) e (2.32),

$$\begin{aligned} \|A_{k_{j+1}} - F'(x^*)\|^2 &\leq \|A_{k_{j+1}} - F'(x^*)\|^2 + \alpha_7 \|x^k - x^*\| & (2.33) \\ &\leq \|A_{k_j} - F'(x^*)\|^2 - \frac{\gamma^2}{32}. \end{aligned}$$

A desigualdade, (2.31) é satisfeita para todo  $j = 1, 2, 3, \dots$ . Portanto:

$$\|A_{k_j} - F'(x^*)\|^2 \leq \|A_{k_j} - F'(x^*)\|^2 - (j-1) \frac{\gamma^2}{32}. \quad (2.34)$$

Mas, de (2.32) tem-se  $\|A_{k_j} - F'(x^*)\|^2 < 0$  para  $j$  suficientemente grande, o que é uma contradição. ■

Com este resultado pode-se derivar uma condição suficiente para convergência superlinear, como mostra o **Teorema 2.5**.

**Teorema 2.5:** *Suponha as hipóteses A1-A5, as seqüências  $\{A_k\}$  e  $\{x^k\}$  geradas por (2.19) e por*

$$x^{k+1} = x^k - B_k^{-1} G(x^k), \quad (2.35)$$

*respectivamente, com  $B_k$  gerada como em (2.3) e  $\lim_{k \rightarrow \infty} x^k = x^*$ .*

*Além disso, se*

$$\lim_{k \rightarrow \infty} \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} = 0 \quad (2.36)$$

*então a seqüência  $\{x^k\}$  converge superlinearmente a  $x^*$ .*

**Prova :**

É uma aplicação direta dos **Teoremas 2.2** e **2.4**. i.e.,

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} &= \lim_{k \rightarrow \infty} \left[ \frac{\|(A_k - A_{k+1})s^k + (A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} \right] \\ &\leq \lim_{k \rightarrow \infty} \left[ \frac{\|(A_k - A_{k+1})s^k\|}{\|s^k\|} + \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} \right] \end{aligned}$$

Mas, do **Teorema 2.4** e da hipótese (2.34), tem-se que a expressão da direita da última desigualdade é zero. Assim,

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} = 0$$

e pelo **Teorema 2.2**, a seqüência  $\{x^k\}$  converge superlinearmente a  $x^*$ . ■

De acordo com o que foi feito nesta seção, a família LCSU gerada por (2.19) fornece algoritmos localmente convergentes e superlineares sob as hipóteses **A1-A5**. Isto é garantido no seguinte **Teorema**.

**Teorema 2.6:** *Suponha as hipóteses A1-A5 e que as seqüências  $\{x^k\}$  e  $\{A_k\}$  são geradas por*

$$x^{k+1} = x^k - B_k^{-1} G(x^k),$$

e (2.19) respectivamente, com  $B_k$  gerada como em (2.3).

Dado  $r \in (0, 1)$  existem  $\bar{\epsilon}$  e  $\bar{\delta}$  tais que, se

$$\|x^0 - x^*\| \leq \bar{\epsilon} \text{ e } \|A_0 - F'(x^*)\| \leq \bar{\delta},$$

a seqüência  $\{x^k\}$  converge local e linearmente a  $x^*$  e se

$$\lim_{k \rightarrow \infty} \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} = 0,$$

então a convergência é superlinear.

**Prova:** é uma aplicação direta dos **Teorema 3, 4 e 5**. ■

## Capítulo 3

# Reformulação do PCN usando a função de Fischer

É fácil ver que resolver o PCN é equivalente a resolver o sistema de equações não lineares

$$\Phi(x) = 0, \quad (3.1)$$

onde a função  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  é dada por:

$$\Phi(x) = \begin{pmatrix} \phi(x_1, f_1(x)) \\ \vdots \\ \phi(x_n, f_n(x)) \end{pmatrix}, \quad (3.2)$$

sendo que, para cada  $i = 1, \dots, n$ ,

$$\phi(x_i, f_i(x)) = \|(x_i, f_i(x))\|_2 - x_i - f_i(x).$$

Se para algum  $i$ ,  $1 \leq i \leq n$ ,  $x_i = f_i(x) = 0$ , a função  $\Phi$  não é diferenciável em  $x$ .

O objetivo deste capítulo é desenvolver e analisar métodos para resolver o PCN usando sua reformulação baseada na função de Fischer, (3.1).

Dado  $x^0 \in \mathbb{R}^n$  uma aproximação inicial, o algoritmo quase-Newton genérico estudado neste capítulo é dado por

$$x^{k+1} = x^k - B_k^{-1} \Phi(x^k),$$

onde,

$$B_k = \begin{pmatrix} [B^k]_1 \\ \vdots \\ [B^k]_n \end{pmatrix}, \quad (3.3)$$

sendo que

$$[B^k]_i = \begin{cases} \left( \frac{x_i^k}{\|(x_i^k, f_i(x^k))\|_2} - 1 \right) \mathbf{e}_i^T + \left( \frac{f_i(x^k)}{\|(x_i^k, f_i(x^k))\|_2} - 1 \right) [A^k]_i, & x_i^k \neq 0 \text{ ou} \\ & f_i(x^k) \neq 0 \\ \left( \frac{z_i^k}{\|(z_i^k, [A^k]_i z^k)\|_2} - 1 \right) \mathbf{e}_i^T + \left( \frac{[A^k]_i z^k}{\|(z_i^k, [A^k]_i z^k)\|_2} - 1 \right) [A^k]_i, & x_i^k = f_i(x^k) = 0, \end{cases}$$

onde,  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  é a base canônica de  $\mathbb{R}^n$ ,  $A_k = \begin{pmatrix} [A^k]_1 \\ \vdots \\ [A^k]_n \end{pmatrix}$  é uma aproximação qualquer da Jacobiana de  $F$  em  $x^k$  e  $z^k \in \mathbb{R}^n$  é tal que  $z_i^k \neq 0$ , se  $x_i^k = f_i(x^k) = 0$ .

Pela definição de  $[B^k]_i$  para o caso  $x_i^k \neq 0$  ou  $f_i(x^k) \neq 0$ , tem-se:

$$[B^k]_i = \begin{cases} -\mathbf{e}_i^T & \text{se } x_i^k = 0 \text{ e } f_i(x^k) > 0 \\ -[A^k]_i & \text{se } x_i^k > 0 \text{ e } f_i(x^k) = 0 \end{cases}$$

Analogamente ao que foi feito no **Capítulo 2** para o sistema  $G(x) = 0$ , neste capítulo desenvolvemos uma família do tipo **Least Change Secant Update (LCSU)**, para o sistema  $\Phi(x) = 0$ , que é gerada de acordo com as regras introduzidas por Martínez ([15]). Além disso, provamos que a família LCSU fornece algoritmos localmente convergentes e superlineares.

### 3.1 Preliminares

Como foi dito na introdução do capítulo, a função  $\Phi$  não é diferenciável em  $x$ , se  $x_i = f_i(x) = 0$ , para algum  $i$ ,  $1 \leq i \leq n$ .

Observe que a função  $\phi$  que define cada uma das componentes de  $\Phi(x)$  é diferenciável para qualquer  $(a, b) \neq (0, 0)$ . Além disso, a derivada de  $\phi$ ,  $\nabla\phi$ , é Lipschitz contínua i.e, existem  $\eta, \epsilon > 0$ , tais que:

$$\|\nabla\phi(a, b) - \nabla\phi(c, d)\| \leq \eta\|(a - b, c - d)\|, \quad (3.4)$$

para todo  $(c, d) \in \mathcal{B}((a, b); \epsilon)$ .

Se  $D_\Phi$  denota o conjunto de pontos onde  $\Phi$  é diferenciável então o B-Jacobiano generalizado,<sup>1</sup>  $\partial_B \Phi(x)$ , é definido da seguinte forma: para qualquer  $x \in \mathbb{R}^n$ ,

$$\partial_B \Phi(x) = \left\{ \lim_{x^k \rightarrow x} \Phi'(x^k) : x^k \in D_\Phi \right\}.$$

O conjunto  $\partial_B \Phi(x)$  é compacto e não vazio. ([21])

De Luca, Facchinei e Kanzow ([2]), dão um procedimento para calcular um elemento de  $\partial_B \Phi(x)$ , e provam que, o elemento que constroem pertence a  $\partial_B \Phi(x)$ . A idéia da prova é aplicar a definição de  $\partial_B \Phi(x)$ ; para isto eles constroem uma seqüência de pontos onde  $\Phi$  é diferenciável e tal que a seqüência das Jacobianas nesses pontos converge à matriz que eles construíram.

A forma da seqüência por eles proposta é:

$$\bar{y}^k = x + \epsilon^k z \quad (3.5)$$

onde  $\{\epsilon^k\}$  é uma seqüência de números positivos convergindo a zero e  $z$  é o vetor da definição de  $[B^k]_i$ .

---

<sup>1</sup>Este conceito foi introduzido em ([21]).

Usando a seqüência  $\{\bar{y}^k\}$ , eles provam que no limite a seqüência de Jacobianas  $\{\Phi'(\bar{y}^k)\}$  converge a

$$\bar{B} = \begin{pmatrix} [\bar{B}]_1 \\ \vdots \\ [\bar{B}]_n \end{pmatrix}, \quad (3.6)$$

sendo que

$$[\bar{B}]_i = \begin{cases} \left( \frac{x_i}{\|(x_i, f_i(x))\|_2} - 1 \right) \mathbf{e}_i^T + \left( \frac{f_i(x)}{\|(x_i, f_i(x))\|_2} - 1 \right) [f'_i(x)], & x_i \neq 0 \text{ ou} \\ & f_i(x) \neq 0 \\ \left( \frac{z_i}{\|(z_i, [f'_i(x)]z)\|_2} - 1 \right) \mathbf{e}_i^T + \left( \frac{[f'_i(x)]z}{\|(z_i, [f'_i(x)]z)\|_2} - 1 \right) [f'_i(x)], & x_i = f_i(x) = 0, \end{cases}$$

Assim, para cada  $1 \leq i \leq n$ ,

$$\lim_{k \rightarrow \infty} \nabla \phi(\bar{y}_i^k, f_i(\bar{y}^k)) = [\bar{B}]_i. \quad (3.7)$$

Seguindo o que foi feito em ([2]) e definindo a seqüência

$$y^k = x^* + \epsilon^k z,$$

onde  $x^*$  é solução do PCN, podemos provar que a matriz:

$$B_* = \begin{pmatrix} [B_*]_1 \\ \vdots \\ [B_*]_n \end{pmatrix}, \quad (3.8)$$

onde para todo  $i = 1, \dots, n$ :

$$[B_*]_i = \begin{cases} -\mathbf{e}_i & \text{se } 0 = x_i^* < f_i(x^*) \\ -f'_i(x^*) & \text{se } x_i^* > f_i(x^*) = 0 \\ (\alpha_i^* - 1)\mathbf{e}_i + (\beta_i^* - 1)f'_i(x^*) & \text{se } x_i^* = f_i(x^*) = 0 \end{cases}$$

$$\alpha_i^* = \frac{z_i}{\|(z_i, f'_i(x^*)z)\|_2} \quad \beta_i^* = \frac{f'_i(x^*)z}{\|(z_i, f'_i(x^*)z)\|_2},$$

é um elemento de  $\partial_B \Phi(x^*)$ . Assim, no limite, a seqüência  $\{\Phi'(y^k)\}$  converge a  $B_*$  e portanto, para cada  $1 \leq i \leq n$ ,

$$\lim_{k \rightarrow \infty} \nabla \phi(y_i^k, f_i(y^k)) = [B_*]_i. \quad (3.9)$$

Estas matrizes  $B_*$  formam um conjunto infinito já que existem infinitas possibilidades para o vetor  $z \in \mathbb{R}^n$ . Chamaremos tal conjunto de  $Z_*$ . Assim,

$$Z_* = \{B_*(z) : z_i \neq 0, \text{ se } x_i^* = f_i(x^*) = 0\}. \quad (3.10)$$

Provaremos a seguir que  $Z_*$  é um conjunto compacto. Como estamos trabalhando num espaço de dimensão finita então basta provar que  $Z_*$  é limitado e fechado.

1. Observe que  $Z_* \subset \partial_B \Phi(x^*)$ , e portanto  $Z_*$  é limitado.<sup>2</sup>
2. Vamos supor, sem perda de generalidade, que as primeiras  $m$  componentes de  $x^*$  são degeneradas.

Se  $z \in \mathbb{R}^n$  é tal que as primeiras  $m$  componentes são não nulas então existe uma seqüência  $\{z^r\}$  de pontos de  $\mathbb{R}^n$  tal que as primeiras  $m$  componentes de cada  $z^r$  são não nulas e

$$\lim_{r \rightarrow \infty} z^r = z. \quad (3.11)$$

Seja  $\{B_*^r\}$  uma seqüência infinita de matrizes contida em  $Z_*$  que converge para uma matriz  $\tilde{B}$ . Da definição das  $B_*$  tem-se que cada uma das linhas de  $B_*^r$  é uma função contínua. Então, denotando a  $i$ -ésima linha de  $B_*^r$  por  $[B_*^r]_i$ , para

---

<sup>2</sup> $\partial_B \Phi(x^*)$  é um conjunto compacto em  $\mathbb{R}^{n \times n}$ .

simplicidade de notação, temos:

$$\tilde{B} = \lim_{r \rightarrow \infty} B_*^r = \begin{pmatrix} \lim_{r \rightarrow \infty} [B_*^r]_1 \\ \vdots \\ \lim_{r \rightarrow \infty} [B_*^r]_m \\ \vdots \\ \lim_{r \rightarrow \infty} [B_*^r]_n \end{pmatrix},$$

onde, para  $i = 1, 2, \dots, m$ ,

$$\begin{aligned} [\tilde{B}]_i &= \lim_{r \rightarrow \infty} [B_*^r]_i \\ &= \lim_{r \rightarrow \infty} \left[ \left( \frac{z_i^r}{\|(z_i^r, f'_i(x^*)z^r)\|_2} - 1 \right) \mathbf{e}_i^T + \left( \frac{f'_i(x^*)z^r}{\|(z_i^r, f'_i(x^*)z^r)\|_2} - 1 \right) f'_i(x^*) \right] \\ &= \left[ \left( \frac{z_i}{\|(z_i, f'_i(x^*)z)\|_2} - 1 \right) \mathbf{e}_i^T + \left( \frac{f'_i(x^*)z}{\|(z_i, f'_i(x^*)z)\|_2} - 1 \right) f'_i(x^*) \right]. \end{aligned}$$

e, para  $i = m + 1, \dots, n$ ,

$$[\tilde{B}]_i = \lim_{r \rightarrow \infty} [B_*^r]_i = \begin{cases} -\mathbf{e}_i & \text{se } 0 = x_i^* < f_i(x^*) \\ -f'_i(x^*) & \text{se } 0 = f_i(x^*) < x_i^* \end{cases}$$

Assim, a matriz  $\tilde{B} \in Z_*$  e portanto  $Z_*$  é fechado.

Logo, por 1. e 2. o conjunto  $Z_*$  é compacto.

## 3.2 Hipóteses Locais

**H1.**  $x^* \in \mathbb{R}^n$  é tal que  $\Phi(x^*) = 0$ .

**H2.** Existem  $\gamma, \epsilon > 0$ , tais que:

$$\|F'(x) - F'(x^*)\| \leq \gamma \|x - x^*\|.$$

para todo  $x \in \mathcal{B}(x^*, \epsilon)$ .

**H3.** Todas as matrizes de  $Z_*$  são não singulares.

Da hipótese **H3** e pela compacidade de  $Z_*$ , tem-se que existe uma constante positiva  $\mu$  tal que para toda  $B_*(z) \in Z_*$ :

$$\|B_*^{-1}\| \leq \mu.$$

## 3.3 Resultados de Convergência

A seguir apresentamos dois **Lemas**, que serão de utilidade na demonstração do **Teorema 3.1**.

**Lema 3.1:** *Seja  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $F \in C^1$  cuja Jacobiana satisfaz a hipótese H2. Sejam  $\epsilon, \delta \geq 0$  dados e sejam  $B$  e  $\bar{B}$  definidas por (3.3) e (3.6), respectivamente. Então para cada  $x \in \mathcal{B}(x^*, \epsilon)$ ,  $A \in \mathcal{B}(F'(x^*), \delta)$ ,*

*existem constantes  $\tau \geq 2$  e  $\hat{\delta} \geq 0$  tais que*

$$\|B - \bar{B}\| \leq \tau \hat{\delta}.$$

## Prova

Assim como no **Capítulo 2** usaremos aqui,  $\|\cdot\| = \|\cdot\|_\infty$ .

Sejam,

$$\Omega_1 = \{i : x_i \neq 0 \text{ ou } f_i(x) \neq 0\}$$

$$\Omega_2 = \{i : x_i = f_i(x) = 0\}.$$

Usando (3.3) e (3.6), tem-se:

$$\|B - \bar{B}\| = \left\| \begin{pmatrix} [B]_1 - [\bar{B}]_1 \\ \vdots \\ [B]_n - [\bar{B}]_n \end{pmatrix} \right\|,$$

onde,

$$[B]_i - [\bar{B}]_i = \begin{cases} \left( \frac{f_i(x)}{\|(x_i, f_i(x))\|_2} - 1 \right) ([A]_i - f'_i(x)), & \text{se } i \in \Gamma_1 \\ (\alpha_i - \bar{\alpha}_i) \mathbf{e}_i^T + (\beta_i - 1)[A]_i - (\bar{\beta}_i - 1)f'_i(x), & \text{se } i \in \Gamma_2, \end{cases}$$

$$\alpha_i = \frac{z_i}{\|(z_i, [A]_i z)\|_2}, \quad \bar{\alpha}_i = \frac{z_i}{\|(z_i, f'_i(x)z)\|_2}$$

$$\beta_i = \frac{[A]_i z}{\|(z_i, [A]_i z)\|_2}, \quad \bar{\beta}_i = \frac{(f'_i(x)z)}{\|(z_i, f'_i(x)z)\|_2}.$$

Sejam,

$$\left\| \left( \frac{f_j(x)}{\|(x_j, f_j(x))\|_2} - 1 \right) ([A]_j - f'_j(x)) \right\|_1 = \max_{i \in \Omega_1} \{ \| [B]_i - [\bar{B}]_i \|_1 \},$$

$$\|(\alpha_k - \bar{\alpha}_k) \mathbf{e}_k^T + (\beta_k - 1)[A]_k - (\bar{\beta}_k - 1)f'_k(x)\|_1 = \max_{i \in \Omega_2} \{ \| [B]_i - [\bar{B}]_i \|_1 \},$$

Pela definição de  $\|\cdot\|_\infty$ , temos duas situações para analisar:

1.  $\|B - \bar{B}\| = \left\| \left( \frac{f_j(x)}{\|(x_j, f_j(x))\|_2} - 1 \right) ([A]_j - f'_j(x)) \right\|_1$ .
2.  $\|B - \bar{B}\| = \|(\alpha_k - \bar{\alpha}_k) \mathbf{e}_k^T + (\beta_k - 1)[A]_k - (\bar{\beta}_k - 1)f'_k(x)\|_1$ .

No primeiro caso usando propriedades das normas temos:

$$\begin{aligned}
 \|B - \bar{B}\| &\leq \frac{|f_j(x)|}{\|(x_j, f_j(x))\|_2} \|[A]_j - f'_j(x)\|_1 + \|[A]_j - f'_j(x)\|_1 \\
 &\leq 2 \|[A]_j - f'_j(x)\|_1 \\
 &\leq 2 \left( \|[A]_j - f'_j(x^*)\|_1 + \|f'_j(x^*) - f'_j(x)\|_1 \right) \\
 &\leq 2n \left( \|[A]_j - f'_j(x^*)\| + \|f'_j(x^*) - f'_j(x)\| \right) \\
 &\leq 2n(\delta + \gamma' \epsilon),
 \end{aligned}$$

onde,  $\gamma' \geq 1$  existe pela hipótese 2.

Portanto, existem constantes  $\tau = 2$  e  $\hat{\delta} = n(\delta + \gamma' \epsilon)$ , tais que:

$$\|B - \bar{B}\| \leq \tau \hat{\delta}.$$

No segundo caso, usando a desigualdade triangular, temos

$$\|B - \bar{B}\| \leq \|(\alpha_k - \bar{\alpha}_k) \mathbf{e}_k^T\|_1 + \|(\beta_k - 1)[A]_k - (\bar{\beta}_k - 1)f'_k(x)\|_1. \quad (3.12)$$

O primeiro termo do lado direito de (3.12) pode limitar-se assim:

$$\begin{aligned}
 \|(\alpha_k - \alpha_k^*) \mathbf{e}_k^T\|_1 &= \left\| \left( \frac{z_k}{\|(z_k, [A]_k z)\|_2} - \frac{z_k}{\|(z_k, f'_k(x)z)\|_2} \right) \mathbf{e}_k^T \right\|_1 \\
 &\leq |z_k| \left| \frac{\|(z_k, f'_k(x)z)\|_2 - \|(z_k, [A]_k z)\|_2}{\|(z_k, [A]_k z)\|_2 \|(z_k, f'_k(x)z)\|_2} \right| \|\mathbf{e}_k^T\|_1 \\
 &\leq |z_k| \frac{\|(0, f'_k(x)z - [A]_k z)\|_2}{\|(z_k, [A]_k z)\|_2 \|(z_k, f'_k(x)z)\|_2} \\
 &= |z_k| \frac{|[A]_k z - f'_k(x)z|}{\|(z_k, [A]_k z)\|_2 \|(z_k, f'_k(x)z)\|_2} \\
 &\leq \frac{|([A]_k - f'_k(x))z|}{\|(z_k, [A]_k z)\|_2} \\
 &\leq \frac{\|[A]_k - f'_k(x)\|_2 \|z\|_2}{\|(z_k, [A]_k z)\|_2} \\
 &\leq \frac{\|[A]_k - f'_k(x)\|_1 \|z\|_2}{\|(z_k, [A]_k z)\|_2}
 \end{aligned}$$

logo,

$$\|(\alpha_k - \bar{\alpha}_k) \mathbf{e}_k^T\|_1 \leq n(\delta + \gamma' \epsilon) \frac{\|z\|_2}{\|(z_k, [A]_k z)\|_2}. \quad (3.13)$$

Agora, o segundo termo do lado direito de (3.12):

$$\begin{aligned}
 \|(\beta_k - 1)[A]_k - (\bar{\beta}_k - 1)f'_k(x)\|_1 &= \|\beta_k[A]_k - \bar{\beta}_k f'_k(x) - ([A]_k - f'_k(x))\|_1 = \\
 &\left\| \frac{[A]_k z}{\|(z_k, [A]_k z)\|_2} [A]_k - \frac{f'_k(x)z}{\|(z_k, f'_k(x)z)\|_2} f'_k(x) - ([A]_k - f'_k(x)) \right\|_1. \quad (3.14)
 \end{aligned}$$

Mas,  $\|[A]_k - f'_k(x)\|_1 \leq n(\delta + \gamma' \epsilon)$ , onde,  $\gamma' \geq 1$  existe pela hipótese **H2**. Logo, se  $\varepsilon_k = [A]_k - f'_k(x)$  então  $[A]_k = f'_k(x) + \varepsilon_k$ , com  $\|\varepsilon\|_1 \leq n(\delta + \gamma' \epsilon)$  e pela

continuidade do produto interno:

$$[A]_k z = f'_k(x)z + \varepsilon^t z \quad (3.15)$$

então, (3.14) pode escrever-se assim:

$$\begin{aligned} & \left\| \frac{(f'_k(x)z + \varepsilon_k z)}{\|(z_k, [A]_k z)\|_2} [A]_k - \frac{f'_k(x)z}{\|(z_k, f'_k(x)z)\|_2} ([A]_k - \varepsilon_k) - \varepsilon_k \right\|_1 = \\ & \left\| f'_k(x)z \left[ \frac{1}{\|(z_k, [A]_k z)\|_2} - \frac{1}{\|(z_k, f'_k(x)z)\|_2} \right] [A]_k \right. \\ & \left. + \frac{\varepsilon_k z}{\|(z_k, [A]_k z)\|_2} [A]_k + \frac{f'_k(x)z}{\|(z_k, f'_k(x)z)\|_2} \varepsilon_k - \varepsilon_k \right\|_1 \leq \\ & |f'_k(x)z| \left| \frac{\|(z_k, f'_k(x)z)\|_2 - \|(z_k, [A]_k z)\|_2}{\|(z_k, [A]_k z)\|_2 \|(z_k, f'_k(x)z)\|_2} \right| \|[A]_k\|_1 \\ & + \frac{|\varepsilon_k z|}{\|(z_k, [A]_k z)\|_2} \|[A]_k\|_1 + \frac{|f'_k(x)z|}{\|(z_k, f'_k(x)z)\|_2} \|\varepsilon_k\|_1 + \|\varepsilon_k\|_1 \leq \\ & 2n(\delta + \gamma'\epsilon) \left( \frac{\|z\|_2}{\|(z_k, [A]_k z)\|_2} \|[A]_k\|_1 + 1 \right). \end{aligned}$$

Logo,

$$\|(\beta_k - 1)[A]_k - (\bar{\beta}_k - 1)f'_k(x)\|_1 \leq 2 \left( \frac{\|z\|_2}{\|(z_k, [A]_k z)\|_2} \|[A]_k\|_1 + 1 \right) n(\delta + \gamma'\epsilon).$$

Por (3.12), (3.13) e da última desigualdade tem-se:

$$\|B - \bar{B}\| \leq \left[ \frac{\|z\|_2}{\|(z_k, [A]_k z)\|_2} + 2 \left( \frac{\|z\|_2}{\|(z_k, [A]_k z)\|_2} \|[A]_k\|_1 + 1 \right) \right] (n\delta + \gamma'\epsilon).$$

Assim, existem

$$\tau = \left[ \frac{\|z\|_2}{\|(z_k, [A]_k z)\|_2} + 2 \left( \frac{\|z\|_2}{\|(z_k, [A]_k z)\|_2} \|[A]_k\|_1 + 1 \right) \right] \geq 2,$$

$$\bar{\delta} = n(\delta + \gamma'\epsilon) \geq 0$$

tais que,

$$\|B - \bar{B}\| \leq \tau \bar{\delta},$$

com o que fica provado o **Lema 3.1**. ■

**Lema 3.2:** Para cada  $x \in \mathbb{R}^n$ ,  $A \in \mathbb{R}^{n \times n}$ , seja

$$\Psi(x, A) = x - B^{-1} \Phi(x), \tag{3.16}$$

onde

$$B = \begin{pmatrix} [B]_1 \\ \vdots \\ [B]_n \end{pmatrix},$$

sendo que

$$[B]_i = \begin{cases} \left( \frac{x_i}{\|(x_i, f_i(x))\|_2} - 1 \right) \mathbf{e}_i^T + \left( \frac{f_i(x)}{\|(x_i, f_i(x))\|_2} - 1 \right) [A]_i, & x_i \neq 0 \text{ ou} \\ & f_i(x) \neq 0 \\ \left( \frac{z_i}{\|(z_i, [A]_i z)\|_2} - 1 \right) \mathbf{e}_i^T + \left( \frac{[A]_i z}{\|(z_i, [A]_i z)\|_2} - 1 \right) [A]_i, & x_i = f_i(x) = 0, \end{cases}$$

Seja  $r \in (0, 1)$ . Então existem  $\epsilon_1, \delta_1 \geq 0$  tais que,

$$\text{se } \|x - x^*\| < \epsilon_1 \text{ e } \|A - F'(x^*)\| \leq \delta_1,$$

a função  $\Psi(x, A)$  está bem definida, e satisfaz

$$\|\Psi(x, A) - x^*\| \leq r \|x - x^*\|.$$

## Prova

Nesta prova se usa  $\|\cdot\| = \|\cdot\|_\infty$ . Seja  $\epsilon_1 > 0$  tal que para todo  $i = 1, \dots, n$ ,

$$\text{se } f_i(x^*) > x_i^* \text{ então } f_i(x) > x_i,$$

$$\text{se } f_i(x^*) < x_i^* \text{ então } f_i(x) < x_i,$$

para todo  $x \in \mathcal{B}(x^*, \epsilon_1)$ . Claramente  $\epsilon_1$  existe pela continuidade da  $F$ . Em particular consideraremos

$$\epsilon_1 \leq \frac{\tau}{16\gamma'\Gamma\tau\mu},$$

onde  $\mu$  é dado na hipótese **H3** para  $\|\cdot\|_\infty$ ,  $\gamma' \geq 1$  e  $\tau$  são as constantes do **Lema 3.1**. Seja  $\Gamma = n\eta'(1 + \kappa')$ , com  $\eta' \geq 1$ ,  $\kappa' \geq 1$ , que existem por (3.4) e porque  $F \in C^1$ , respectivamente.

Seja ainda  $\delta_1 \leq \frac{\tau}{16\tau\mu}$ . Sejam  $x \in \mathcal{B}(x^*, \epsilon_1)$ ,  $A \in \mathcal{B}(F'(x^*), \delta_1)$ ,  $B$  matriz associada a  $A$  pela regra (3.3),  $B_*$  matriz associada a  $F'(x^*)$  pela mesma regra e  $\bar{B}$  definida no **Lema 3.1**. Temos que:

$$\|B - B_*\| \leq \|B - \bar{B}\| + \|\bar{B} - B_*\|. \quad (3.17)$$

Usando o **Lema 3.2**, o primeiro termo do lado direito de (3.17) pode limitar-se da seguinte forma:

$$\|B - \bar{B}\| \leq \tau(\delta_1 + \gamma'\epsilon_1) \leq \frac{\tau}{8\mu}. \quad (3.18)$$

Para limitar o segundo termo de (3.17), observemos o seguinte:

Pela definição de  $\|\cdot\| = \|\cdot\|_\infty$ ,

$$\|\bar{B} - B_*\| = \max_{1 \leq i \leq n} \{ \|[ \bar{B} ]_i - [ B_* ]_i \|_1 \}.$$

Para cada  $1 \leq i \leq n$ , usando (3.7), (3.9), a continuidade da norma e a continuidade de  $F$ , assim como a continuidade Lipschitz de  $\nabla\phi$ , temos:

$$\begin{aligned} \|[\bar{B}]_i - [B^*]_i\|_1 &\leq n \|[\bar{B}]_i - [B^*]_i\| \\ &= n \lim_{k \rightarrow \infty} \|\nabla\phi(\bar{y}_i^k, f_i(\bar{y}^k)) - \nabla\phi(y_i^k, f_i(y^k))\| \\ &\leq \lim_{k \rightarrow \infty} \eta' \|(\bar{y}_i^k - y_i^k, f_i(\bar{y}^k) - f_i(y^k))\| \\ &\leq n \eta' (1 + \kappa') \|x - x^*\| = \Gamma \epsilon_1, \end{aligned}$$

Assim,

$$\|\bar{B} - B^*\| \leq \frac{r}{16\mu}. \quad (3.19)$$

Logo, de (3.17), (3.18), e (3.19) temos que:

$$\|B - B^*\| \leq \frac{3r}{16\mu} < \frac{1}{2\mu} \quad (3.20)$$

Portanto, pelo **Lema de Banach**,  $B^{-1}$  existe e além disso,

$$\|B^{-1}\| \leq 2 \|B^*\| \leq 2\mu.$$

Subtraindo  $x^*$  de (3.16) e aplicando o operador norma  $\|\cdot\|$  tem-se:

$$\begin{aligned} \|\Psi(x, A) - x^*\| &= \|(x - x^*) - B^{-1}\Phi(x) + B^{-1}B_*(x - x^*) - B^{-1}B_*(x - x^*)\| \\ &= \|(I - B^{-1}B^*)(x - x^*) - B^{-1}[\Phi(x) - \Phi(x^*) - B_*(x - x^*)]\| \\ &\leq \|B^{-1}\| [\|B - B^*\| \|x - x^*\| + \|\Phi(x) - \Phi(x^*) - B_*(x - x^*)\|] \end{aligned}$$

$$\leq 2\mu \left[ \frac{3r}{16\mu} + \frac{\|\Phi(x) - \Phi(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \right] \|x - x^*\| \quad (3.21)$$

Mas,

$$\frac{\|\Phi(x) - \Phi(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \leq \frac{\|\Phi(x) - \Phi(x^*) - \bar{B}(x - x^*)\|}{\|x - x^*\|} + \|\bar{B} - B_*\|$$

onde,  $\bar{B} \in \partial_B \Phi(x)$ . Assim<sup>3</sup> para qualquer  $\rho > 0$  existe  $\epsilon > 0$  tal que

$$\|x - x^*\| \leq \epsilon \Rightarrow \frac{\|\Phi(x) - \Phi(x^*) - \bar{B}(x - x^*)\|}{\|x - x^*\|} \leq \rho.$$

Em particular, dado  $\rho = \frac{r}{4\mu}$ , existe  $\epsilon_r > 0$  tal que

$$\|x - x^*\| \leq \epsilon_r \Rightarrow \frac{\|\Phi(x) - \Phi(x^*) - \bar{B}(x - x^*)\|}{\|x - x^*\|} \leq \frac{r}{4\mu}.$$

Logo,

$$\frac{\|\Phi(x) - \Phi(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \leq \frac{r}{4\mu} + \|\bar{B} - B_*\| \leq \frac{r}{4\mu} + \frac{r}{16\mu} = \frac{5}{16\mu}$$

Finalmente, de (3.21) e da última desigualdade:

$$\begin{aligned} \|\Psi(x, A) - x^*\| &\leq 2\mu \left[ \frac{3r}{16\mu} + \frac{5r}{16\mu} \right] \|x - x^*\| \\ &= r \|x - x^*\|, \end{aligned}$$

---

<sup>3</sup>Qi [1994], provou que se  $H$  é semisuave em  $x$ ,

$$\lim_{(x+h) \rightarrow x} \frac{\|H(x+h) - H(x) - \bar{B}h\|}{\|h\|} = 0,$$

para toda matriz  $\bar{B} \in \partial_B H(x)$ .

com o que fica provado o **Lema 3.2**. ■

**Teorema 3.1 : (Das duas vizinhanças)** *Dado qualquer  $r \in (0, 1)$  existem  $\epsilon, \delta > 0$  tais que,*

$$\text{se } \|x^0 - x^*\| \leq \epsilon \text{ e } \|A_k - F'(x^*)\| \leq \delta, \text{ para todo } k,$$

*a seqüência  $\{x^k\}$  gerada por*

$$x^{k+1} = x^k - B_k^{-1} \Phi(x^k),$$

*está bem definida, converge a  $x^*$  e satisfaz*

$$\|x^{k+1} - x^*\| \leq r \|x^k - x^*\|, \text{ para todo } k = 0, 1, 2, \dots,$$

### Prova do Teorema 3.1

Considerando a função  $\Psi(x, A) = x - B^{-1}\Phi(x)$  definida em (3.16), onde  $B$  é gerada por (3.3), tem-se:

$$x^{k+1} = \Psi(x^k, A_k), \quad k = 0, 1, 2, \dots$$

Sejam  $\epsilon_1$  e  $\delta_1$  os mesmos definidos no **Lema 3.2** e suponha  $\epsilon \in (0, \epsilon_1), \delta \in (0, \delta_1)$ .

Para  $k = 0$ , se  $\|x^0 - x^*\| \leq \epsilon \leq \epsilon_1$  e  $\|A_0 - F'(x^*)\| \leq \delta \leq \delta_1$ , pelo **Lema 3.1**,  $x^1$  está bem definido e satisfaz

$$\|x^1 - x^*\| = \|\Psi(x^0, A_0) - x^*\| \leq r \|x^0 - x^*\|.$$

Portanto, para  $k = 0$  a tese do **Teorema 3.1** é válida.

Suponha a hipótese indutiva para  $k - 1$ , i.e.,

$$\text{se } \|x^{k-1} - x^*\| \leq \epsilon, \text{ como } \|A_{k-1} - F'(x^*)\| \leq \delta,$$

então  $x^k$  está bem definido e satisfaz

$$\|x^k - x^*\| = \|\Psi(x^{k-1}, A_{k-1}) - x^*\| \leq r\|x^{k-1} - x^*\|.$$

Da última desigualdade tem-se:

$$\|x^k - x^*\| \leq r\|x^{k-1} - x^*\| \leq r^k \epsilon < \epsilon < \epsilon_1,$$

Assim, pelo **Lema 3.2**,  $x^{k+1}$  está bem definido e satisfaz

$$\|x^{k+1} - x^*\| \leq r\|x^k - x^*\|,$$

com o que fica provado o **Teorema 3.1.** ■

**Teorema 3.2** *Suponha as hipóteses **H1**, **H2** e **H3** e que para algum  $x^0$  qualquer seqüência de pontos gerada por*

$$x^{k+1} = x^k - B_k^{-1} \Phi(x^k), \tag{3.22}$$

*onde  $B_k$  é gerada como em (3.3), satisfaz  $\lim_{k \rightarrow \infty} x^k = x^*$ . Definimos  $A_k$ , como em (3.3) e  $s^k = x^{k+1} - x^k$ .*

Se

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - F'(x^*)) s^k\|}{\|s^k\|} = 0$$

*então a seqüência  $\{x^k\}$  converge superlinearmente a  $x^*$ .*

O seguinte **Lema** será útil na prova do **Teorema 3.2**.

**Lema 3.3** *Suponha as hipóteses **H1**, **H2**, e **H3**. Então existem  $\epsilon, \eta > 0$  tais que, se  $\|x - x^*\| < \epsilon$  então*

$$\|\Phi(x)\| \geq \eta\|x - x^*\|.$$

**Prova :**

$$\begin{aligned} \|x - x^*\| &= \|(B_*)^{-1} B_*(x - x^*)\| \\ &\leq \|B_*^{-1}\| \|B_*(x - x^*)\| \end{aligned} \quad (3.23)$$

Conforme à demonstração do **Lema 3.1**, para qualquer  $\rho > 0$  existe  $\epsilon > 0$  tal que se  $\|x - x^*\| < \epsilon$  então

$$\frac{\|\Phi(x) - \Phi(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} \leq \rho. \quad (3.24)$$

Assim ,

$$\lim_{x \rightarrow x^*} \frac{\|\Phi(x) - \Phi(x^*) - B_*(x - x^*)\|}{\|x - x^*\|} = 0.$$

Por (3.24) e usando a hipótese **H1**:

$$\lim_{x \rightarrow x^*} \frac{\|\Phi(x) - B_*(x - x^*)\|}{\|B_*^{-1}\| \|B_*(x - x^*)\|} = 0.$$

Logo,

$$\lim_{x \rightarrow x^*} \frac{\|\Phi(x) - B_*(x - x^*)\|}{\|B_*(x - x^*)\|} = 0.$$

Mas,

$$|\|\Phi(x)\| - \|B_*(x - x^*)\|| \leq \|\Phi(x) - B_*(x - x^*)\|;$$

de (3.24), dado  $\rho = \frac{1}{2}$  existe  $\epsilon > 0$  tal que, se  $0 < \|x - x^*\| \leq \epsilon$  então

$$-\frac{1}{2} \leq \frac{\|\Phi(x)\| - \|B_*(x - x^*)\|}{\|B_*(x - x^*)\|} \leq \frac{1}{2}, \quad (3.25)$$

ou seja,

$$\|\Phi(x)\| \geq \frac{1}{2} \|B_*(x - x^*)\| \geq \frac{1}{2} \mu \|x - x^*\|, \quad (3.26)$$

onde, a última desigualdade obtém-se de (3.25) e da hipótese **H3**.

Portanto, a tese do **Lema 3.3** segue de (3.26), com  $\eta = \frac{1}{2}\mu$ . ■

**Prova do Teorema 3.2:** Suporemos que  $x^k \in \mathcal{B}(x^*, \epsilon_1)$ , para todo  $k$  e que  $\|\cdot\| = \|\cdot\|_\infty$ .

Seja

$$A_k = \begin{pmatrix} [A^k]_1 \\ \vdots \\ [A^k]_n \end{pmatrix};$$

com esta notação

$$\|(A_k - F'(x^*))s^k\| = \left\| \begin{pmatrix} [A^k]_1 - f'_1(x^*) \\ \vdots \\ [A^k]_n - f'_n(x^*) \end{pmatrix} s^k \right\|, \quad (3.27)$$

De (3.3) e pela hipótese **H3**, para cada  $B_k$  existe uma  $B_*$  tal que

$$\|(B_k - B_*)s^k\| = \left\| \begin{pmatrix} [B^k]_1 - [B^*]_1 \\ \vdots \\ [B^k]_n^t - [B^*]_n \end{pmatrix} s^k \right\|, \quad (3.28)$$

onde,

$$[B]_1 - [B^*]_1 = \begin{cases} 0 & x_i^* = 0 \text{ e } f_i(x^*) > 0 \\ -( [A]_i - f'_i(x^*) ) & x_i^* > 0 \text{ e } f_i(x^*) = 0 \\ (\alpha_i - \alpha_i^*)\mathbf{e}_i^T + (\beta_i - 1)[A]_i - (\beta_i^* - 1)f'_i(x^*) & x_i^* = f_i(x^*) = 0, \end{cases}$$

$$\alpha_i = \frac{z_i}{\|(z_i, [A]_i z)\|_2}, \quad \alpha_i^* = \frac{z_i}{\|(z_i, f'_i(x^*)z)\|_2}$$

$$\beta_i = \frac{[A]_i z}{\|(z_i, [A]_i z)\|_2}, \quad \beta_i^* = \frac{(f'_i(x^*)z)}{\|(z_i, f'_i(x^*)z)\|_2}.$$

Assim, de (3.27) e (3.28), conforme a demonstração do **Lema 3.1** :

$$\|(B_k - B_*^k) s^k\| \leq \|(A_k - F'(x_*)) s^k\|.$$

De (3.22), tem-se

$$\begin{aligned} 0 &= B_k s^k + \Phi(x^k) \\ -\Phi(x^{k+1}) &= (B_k - B_*^k) s^k - \Phi(x^{k+1}) + \Phi(x^k) + B_*^k s^k. \end{aligned} \quad (3.29)$$

Multiplicando (3.29) por  $\|s^k\|^{-1}$  e aplicando o operador norma  $\|\cdot\|$  tem-se:

$$\begin{aligned} \frac{\|\Phi(x^{k+1})\|}{\|s^k\|} &\leq \frac{\|(B_k - B_*^k) s^k\|}{\|s^k\|} + \frac{\|-\Phi(x^{k+1}) + \Phi(x^k) + B_*^k s^k\|}{\|s^k\|} \\ &\leq \frac{\|(A_k - F'(x_*)) s^k\|}{\|s^k\|} + \frac{\|-\Phi(x^{k+1}) + \Phi(x^k) + B_*^k s^k\|}{\|s^k\|} \\ &\leq \frac{\|(A_k - F'(x_*)) s^k\|}{\|s^k\|} + \gamma \max\{\|x^{k+1} - x^*\|, \|x^k - x^*\|\}. \end{aligned}$$

Usando a hipótese

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - F'(x_*)) s^k\|}{\|s^k\|} = 0,$$

e

$$\lim_{k \rightarrow \infty} (x^k - x^*) = 0,$$

tem-se,

$$\lim_{k \rightarrow \infty} \frac{\|\Phi(x^{k+1})\|}{\|s^k\|} = 0.$$

Mas, pelo **Lema 3.3**, existe uma constante positiva  $\eta$  tal que:

$$\lim_{k \rightarrow \infty} \frac{\|\Phi(x^{k+1})\|}{\|s^k\|} \geq \lim_{k \rightarrow \infty} \eta \frac{\|x^{k+1} - x^*\|}{\|s^k\|}.$$

Então,

$$\begin{aligned} 0 &\geq \lim_{k \rightarrow \infty} \eta \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\| + \|x^{k+1} - x^*\|} \\ &= \lim_{k \rightarrow \infty} \eta \frac{\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}}{1 + \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|}}, \end{aligned}$$

o que implica

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

e portanto  $\{x^k\}$  converge a  $x^*$  superlinearmente. ■

### 3.4 Família LCSU para $\Phi(x) = 0$ .

De forma análoga ao que fizemos no **Capítulo 2** para o sistema  $G(x) = 0$ , geraremos uma família do tipo **LCSU** para o sistema  $\Phi(x) = 0$ , com  $\Phi$  dada por (3.1). Para isto, usaremos as regras gerais da teoria de convergência local desenvolvida por Martínez [rfmml] para métodos tipo **LCSU**. Além disso, provaremos que a família **LCSU** fornece algoritmos localmente convergentes e superlineares.

#### Algoritmo 3.1

Suponha que  $x^0$  e  $A_0$  são arbitrários. Para  $k = 1, 2, \dots$ ,  $x^{k+1}$  e  $A_{k+1}$  são gerados como segue:

$$\begin{aligned} x^{k+1} &= x^k - B_k^{-1} \Phi(x^k) \\ A_{k+1} &= P_{x^k, x^{k+1}}(A_k) \end{aligned} \tag{3.30}$$

Além das hipóteses gerais **H1**, **H2**, e **H3**, suponha como em [15] que:

**H4.** Existe  $\alpha_1 > 0$  tal que, para todo  $x, y \in \mathbb{R}^n$ , existe  $A \in V(x, y)$  que satisfaz

$$\|A - F'(x^*)\| \leq \alpha_1 \sigma(x, y), \quad (3.31)$$

onde,  $\sigma(x, y) = \max\{\|x - x^*\|, \|y - x^*\|\}$ .

**H5.** Existe  $\alpha_2 > 0$  tal que, para todo  $x, y \in \mathbb{R}^n$ ,  $A \in V(x, y)$

$$\|A\|_{x,y} \leq [1 + \alpha_2 \sigma(x, y)] \|A\|. \quad (3.32)$$

$$\|A\| \leq [1 + \alpha_2 \sigma(x, y)] \|A\|_{x,y} \quad (3.33)$$

Martínez ([15]), demonstrou que uma matriz  $A$  gerada pela regra (3.30) e que satisfaz as hipóteses **H4** e **H5**, pode deteriorar-se mas isto ocorre de forma controlada, o que é garantido pelos seguintes resultados:

**Lema 3.4** *Suponha as hipóteses H1-H5. Existem  $\alpha_3, \alpha_4 > 0$  tais que para todo  $x, y \in \mathcal{B}(x^*, \epsilon_1)$ ,  $A \in V(x, y)$ ,*

$$\|P_{x,y}(A) - F'(x^*)\| \leq [1 + \sigma(x, y)] \|A - F'(x^*)\| + \alpha_3 \sigma(x, y).$$

**Prova:** (Ver [15]).

**Corolário 3.1:** *Existe  $\alpha_5 > 0$  tal que*

$$\|P_{x,y}(A) - F'(x^*)\| \leq \|A - F'(x^*)\| + \alpha_5 \|x - x^*\|, \quad (3.34)$$

*sempre que  $x, y \in \mathcal{B}(x^*, \epsilon_1)$ ,  $A \in \mathcal{B}(F'(x^*), \delta_1)$  e  $\|y - x^*\| \leq \|x - x^*\|$ .*

**Prova:** (Ver [15]).

Estes dois resultados, além das hipóteses **H1**, **H2** e **H3**, são fundamentais para garantir o seguinte **Teorema**:

**Teorema 3.3:** *Suponha as hipóteses H1-H5 e que a seqüência  $\{A_k\}$  está definida por (3.30). Dado  $r \in (0, 1)$  existem  $\bar{\epsilon}$  e  $\bar{\delta}$  tais que,*

$$\text{se } \|x^0 - x^*\| \leq \bar{\epsilon} \text{ e } \|A_0 - F'(x^*)\| \leq \bar{\delta},$$

qualquer seqüência  $x_k$  gerada por

$$x^{k+1} = x^k - B_k^{-1} \Phi(x^k),$$

está bem definida, converge a  $x^*$  e para todo  $k = 0, 1, 2, \dots$

$$\|x^{k+1} - x^*\| \leq r \|x^k - x^*\|.$$

Além disso, para todo  $k, j = 0, 1, 2, \dots$  existem constantes positivas  $\alpha_6$  e  $\alpha_7$  tais que:

$$\|A_{k+j} - F'(x^*)\| \leq \|A_k - F'(x^*)\| + \alpha_6 \|x^k - x^*\|. \quad (3.35)$$

$$\|A_{k+j} - F'(x^*)\|^2 \leq \|A_k - F'(x^*)\|^2 + \alpha_7 \|x^k - x^*\| \quad (3.36)$$

**Prova:** Suponha  $\bar{\epsilon} \in (0, \epsilon_1)$ ,  $\bar{\delta} \in (0, \delta_1)$ , com  $\epsilon_1$  e  $\delta_1$  os definidos no **Lema 3.2**, e

$$\bar{\delta} + \alpha_5 \frac{\bar{\epsilon}}{1-r} < \delta_1$$

A prova é feita usando indução sobre  $k$ .

Para  $k = 0$ ,

$$\text{se } \|x^0 - x^*\| \leq \bar{\epsilon} \leq \epsilon_1 \quad \text{e} \quad \|A_0 - F'(x^*)\| \leq \bar{\delta} \leq \delta_1,$$

então, pelo **Lema 3.2**,  $x^1$  está bem definido e satisfaz

$$\|x^1 - x^*\| \leq r \|x^0 - x^*\|.$$

Portanto, para  $k = 0$  a tese do teorema é verdadeira.

Suponha a hipótese indutiva para  $k - 1$ , i.e.,

$$\text{se } \|x^{k-1} - x^*\| \leq \bar{\epsilon} \quad \text{e} \quad \|A_{k-1} - F'(x^*)\| \leq \bar{\delta},$$

então  $x^k$  está bem definido e satisfaz

$$\|x^k - x^*\| \leq r \|x^{k-1} - x^*\|.$$

Mas, da última desigualdade tem-se:

$$\|x^k - x^*\| \leq r\|x^{k-1} - x^*\| \leq r^k \epsilon < \epsilon < \epsilon_1,$$

e pelo **Corolário 3**,

$$\begin{aligned} \|A_k - F'(x^*)\| &\leq \|A_{k-1} - F'(x^*)\| + \alpha_5 \|x^{k-1} - x^*\| \\ &\leq \bar{\delta} + \alpha_5 \bar{\epsilon} \sum_{j=0}^{k-1} r^j \\ &\leq \bar{\delta} + \alpha_5 \bar{\epsilon} \sum_{j=0}^{\infty} r^j \\ &\leq \bar{\delta} + \alpha_5 \frac{\bar{\epsilon}}{1-r} < \delta_1 \end{aligned}$$

Assim, pelo **Lema 3.2**,  $x^{k+1}$  está bem definido e satisfaz

$$\|x^{k+1} - x^*\| \leq r\|x^k - x^*\|.$$

Provaremos por indução sobre  $j$  que

$$\|A_{k+j} - F'(x^*)\| \leq \|A_k - F'(x^*)\| + \alpha_5 \|x^k - x^*\| \sum_{m=0}^{j-1} r^m.$$

Para  $j = 0$  o resultado é imediato usando o fato de que  $\sum_{m=0}^{-1} = 0$ .

Suponha a desigualdade válida para  $j - 1$ , i.e.,

$$\|A_{k+j-1} - F'(x^*)\| \leq \|A_k - F'(x^*)\| + \alpha_5 \|x^k - x^*\| \sum_{m=0}^{j-2} r^m.$$

De (2.35), tem-se

$$\begin{aligned} \|A_{k+j} - F'(x^*)\| &\leq \|A_{k+j-1} - F'(x^*)\| + \alpha_5 \|x^{k+j-1} - x^*\| \\ \|A_{k+j} - F'(x^*)\| &\leq \|A_{k+j-1} - F'(x^*)\| + \alpha_5 r^{j-1} \|x^k - x^*\|, \end{aligned}$$

Portanto,

$$\|A_{k+j} - F'(x^*)\| \leq \|A_k - F'(x^*)\| + \alpha_5 \|x^k - x^*\| \sum_{m=0}^{j-1} r^m.$$

Assim, (3.35) segue desta última desigualdade com

$$\alpha_6 = \frac{\alpha_5}{1 - r},$$

e (3.36) se deduz de (3.35) usando a limitação de  $\|A_k - F'(x^*)\|$ . ■

**Teorema 3.4 :** *Sob as hipóteses do Teorema 3.3,*

$$\lim_{k \rightarrow \infty} \|A_{k+1} - A_k\| = 0. \quad (3.37)$$

**Prova:** É a mesma do **Teorema 3.3.3** de Martínez [1990].

Com este resultado pode-se derivar uma condição suficiente para convergência superlinear, como mostra o seguinte **Teorema**:

**Teorema 3.5 :** *Suponha as hipóteses H1-H5, as seqüências  $\{A_k\}$  e  $\{x^k\}$  geradas por (3.30) e por*

$$x^{k+1} = x^k - B_k^{-1} \Phi(x^k), \quad (3.38)$$

*respectivamente, com  $B_k$  gerada como em (1.3) e  $\lim_{k \rightarrow \infty} x^k = x^*$ .*

*Além disso, se*

$$\lim_{k \rightarrow \infty} \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} = 0 \quad (3.39)$$

*então a seqüência  $\{x^k\}$  converge superlinearmente a  $x^*$ .*

**Prova :**

É uma aplicação direta dos **Teorema 3.2** e **3.4**. i.e.,

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} &= \lim_{k \rightarrow \infty} \left[ \frac{\|(A_k - A_{k+1})s^k + (A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} \right] \\ &\leq \lim_{k \rightarrow \infty} \left[ \frac{\|(A_k - A_{k+1})s^k\|}{\|s^k\|} + \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} \right]. \end{aligned}$$

Mas, do **Teorema 3.4** e da hipótese (3.39), tem-se que a expressão da direita da última desigualdade é zero. Assim,

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - F'(x^*))s^k\|}{\|s^k\|} = 0$$

e pelo **Teorema 3.2**, a seqüência  $\{x^k\}$  converge superlinearmente a  $x^*$ .

De acordo com o que fizemos nesta seção, a família **LCSU** gerada por (3.30) fornece algoritmos localmente convergentes e superlineares sob as hipóteses **H1-H5**. Isto é garantido no seguinte **Teorema**:

**Teorema 3.6:** *Suponha as hipóteses **H1-H5** e que as seqüências  $\{x^k\}$  e  $\{A_k\}$  são geradas por*

$$x^{k+1} = x^k - B_k^{-1} \Phi(x^k), \quad (3.40)$$

e (3.30) respectivamente, com  $B_k$  gerada como em (3.3).

Dado  $r \in (0, 1)$  existem  $\bar{\epsilon}$  e  $\bar{\delta}$  tais que, se

$$\|x^0 - x^*\| \leq \bar{\epsilon} \text{ e } \|A_0 - F'(x^*)\| \leq \bar{\delta},$$

a seqüência  $\{x^k\}$  converge local e linearmente a  $x^*$  e se

$$\lim_{k \rightarrow \infty} \frac{\|(A_{k+1} - F'(x^*))s^k\|}{\|s^k\|} = 0, \quad (3.41)$$

então a convergência é superlinear.

**Prova:**

É uma aplicação direta dos **Teorema 3.3, 3.4 e 3.5**.

# Capítulo 4

## A hipótese de não singularidade

Uma das hipóteses sob a qual foi desenvolvida a teoria de convergência local, nos **Capítulos 2 e 3**, é a não singularidade das matrizes de  $\partial_B G(x^*)$  e  $Z_* \subset \partial_B \Phi(x^*)$ , respectivamente. Esta hipótese é conhecida como a hipótese de *BD-regularidade das funções  $G$  e  $\Phi$  na solução  $x^*$* .

Observe que estamos supondo *BD-regularidade* da função  $\Phi$  só num subconjunto do  $\partial_B \Phi(x^*)$ , que logicamente é uma hipótese mais fraca que pedir *BD-regularidade* em todo o  $\partial_B \Phi(x^*)$ .

Neste capítulo mostraremos que sob a hipótese de *BD-regularidade de  $G$  em  $x^*$*  é possível encontrar elementos singulares em  $Z_* \subset \partial_B \Phi(x^*)$ . Esta situação é ilustrada com o primeiro exemplo do capítulo e generalizada para o caso  $n = 2$ . No caso de  $n$  qualquer, sob certas hipóteses, apresentamos uma condição suficiente para ter elementos singulares em  $Z_* \subset \partial_B \Phi(x^*)$ .

Iniciamos dando resposta à seguinte pergunta:

*Se todas as matrizes de  $\partial_B G(x^*)$  são não singulares então todas as matrizes em  $Z_* \subset \partial_B \Phi(x^*)$  são não singulares. ?*

A resposta é: *não necessariamente*. Ilustramos este fato nos seguintes exemplos:

**Exemplo 1:** Seja:

$$\begin{aligned} F : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto F(x) = \begin{pmatrix} x_1 + 3x_2 - 1 \\ x_1 + x_2 - 1 \end{pmatrix} \end{aligned}$$

$x^* = (1, 0)^T$  é solução degenerada do PCN associado a  $F$ , já que

$$\begin{aligned} x_1^* &> f_1(x^*) = 0, \\ x_2^* &= f_2(x^*) = 0. \end{aligned}$$

Assim,

$$\partial_B G(x^*) = \left\{ \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} \right\}.$$

As matrizes em  $\partial_B G(x^*)$  são não singulares. Pelo que foi feito no **Capítulo 3**, sabemos que um elemento de  $Z_* \subset \partial_B \Phi(x^*)$  tem a forma:

$$B_*(z) = \begin{pmatrix} -1 & -3 \\ (\alpha - 1)\mathbf{e}_2^T + (\beta - 1)(1, 1) \end{pmatrix},$$

com  $\alpha^2 + \beta^2 = 1$ , sendo que

$$\begin{aligned} \alpha &= \frac{z_2}{\|(z_2, f_2'(x^*)z)\|_2} = \frac{z_2}{\|(z_2, z_1 + z_2)\|_2} \\ \beta &= \frac{f_2'(x^*)z}{\|(z_2, f_2'(x^*)z)\|_2} = \frac{z_1 + z_2}{\|(z_2, z_1 + z_2)\|_2} \end{aligned}$$

Seja  $z \in \mathbb{R}^2$  tal que  $z_2 \neq 0$  já que  $x_2^* = f_2(x^*) = 0$ .

Seja  $z^T = (1, 3)$ ; então  $\|(z_2, z_1 + z_2)\|_2 = 5$  e

$$B_*(z) = \begin{pmatrix} -1 & -3 \\ (\frac{3}{5} - 1)\mathbf{e}_2^T + (\frac{4}{5} - 1)(1, 1) \end{pmatrix},$$

ou seja

$$B_*(z) = \begin{pmatrix} -1 & -3 \\ -\frac{1}{5} & -\frac{3}{5} \end{pmatrix},$$

que é uma matriz singular.

Portanto, *existe pelo menos um elemento singular em  $Z_* \subset \partial_B \Phi(x^*)$ .*

O segundo exemplo se enquadra nos problemas para os quais sob a hipótese de não singularidade dos elementos de  $\partial_B G(x^*)$ , temos todos os elementos de  $Z_* \subset \partial_B \Phi(x^*)$  também não singulares.

**Exemplo 2:** Seja:

$$\begin{aligned} F : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto F(x) = \begin{pmatrix} 2x_1 + x_2 - 2 \\ x_1 + x_2 - 1 \end{pmatrix} \end{aligned}$$

$x^* = (1, 0)^T$  é solução degenerada do PCN associado a  $F$ , já que:

$$\begin{aligned} x_1^* &> f_1(x^*) = 0 \\ x_2^* &= f_2(x^*) = 0 \end{aligned}$$

Assim:

$$\partial_B G(x^*) = \left\{ \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \right\},$$

Observe que as matrizes em  $\partial_B G(x^*)$  são não singulares.

Novamente temos que toda  $B_*(z) \in Z_* \subset \partial_B \Phi(x^*)$  é da forma:

$$B_*(z) = \begin{pmatrix} -2 & -1 \\ (\alpha - 1)\mathbf{e}_2^T + (\beta - 1)(1, 1) \end{pmatrix},$$

com  $\alpha^2 + \beta^2 = 1$ , sendo

$$\alpha = \frac{z_2}{\|(z_2, f_2'(x^*)z)\|_2} = \frac{z_2}{\|(z_2, z_1 + z_2)\|_2}$$

$$\beta = \frac{f_2'(x^*)z}{\|(z_2, f_2'(x^*)z)\|_2} = \frac{z_1 + z_2}{\|(z_2, z_1 + z_2)\|_2}$$

e  $z \in \mathbb{R}^2$ , tal que  $z_2 \neq 0$  já que  $x_2^* = f_2(x^*) = 0$ .

Seja  $z^T = (1, 3)$ ; então  $\|(z_2, z_1 + z_2)\|_2 = 5$  e

$$B_*(z) = \begin{pmatrix} -2 & -1 \\ (\frac{3}{5} - 1)\mathbf{e}_2^T + (\frac{4}{5} - 1)(1, 1) \end{pmatrix}$$

$$B_*(z) = \begin{pmatrix} -2 & -1 \\ -\frac{1}{5} & -\frac{3}{5} \end{pmatrix},$$

que é uma matriz não singular.

Mas, será que  $B_*(z)$  é não singular para toda  $B_*(z) \in Z_* \subset \partial_B \Phi(x^*)$ ? A resposta é sim!

Felizmente, no caso  $n = 2$  podemos dar uma condição necessária e suficiente que nos permite saber se existem elementos singulares ou não em  $Z_* \subset \partial_B \Phi(x^*)$ , dado que todos os elementos de  $\partial_B G(x^*)$  sejam não singulares.

Seja  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $F \in C^1$  e  $x^*$  uma solução degenerada do PCN. Em particular e sem perda de generalidade, consideraremos a situação:

$$\begin{aligned} x_1^* &> f_1(x^*) = 0 \\ x_2^* &= f_2(x^*) = 0, \end{aligned}$$

para a qual,  $G(x^*) = 0$  e  $\Phi(x^*) = 0$ .

As funções :  $G, \Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  estão definidas por:

$$G(x) = \begin{pmatrix} \min \{x_1, f_1(x)\} \\ \min \{x_2, f_2(x)\} \end{pmatrix}. \quad (4.1)$$

e

$$\Phi(x) = \begin{pmatrix} \phi(x_1, f_1(x)) \\ \phi(x_2, f_2(x)) \end{pmatrix}, \quad (4.2)$$

sendo que, para  $i = 1, 2$ ,

$$\phi(x_i, f_i(x)) = \|(x_i, f_i(x))\|_2 - x_i - f_i(x).$$

Assim:

$$\partial_B G(x^*) = \left\{ \left( \begin{pmatrix} \frac{\partial f_1(x^*)}{\partial x_1} & \frac{\partial f_1(x^*)}{\partial x_2} \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \frac{\partial f_1(x^*)}{\partial x_1} & \frac{\partial f_1(x^*)}{\partial x_2} \\ \frac{\partial f_2(x^*)}{\partial x_1} & \frac{\partial f_2(x^*)}{\partial x_2} \end{pmatrix} \right) \right\}. \quad (4.3)$$

Para simplificar a notação chamaremos um elemento de  $\partial_B G(x^*)$  por:

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

**Teorema 4.1 :** *Sejam  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $F \in C^1$ ,  $G$  e  $\Phi$  definidas como em (4.1) e (4.2). Se todas as matrizes de  $\partial_B G(x^*)$  são não singulares então existem matrizes singulares em  $Z_* \subset \partial_B \Phi(x^*)$  se e somente se*

1.  $b_{11}$  e  $\det(B)$  têm sinais opostos.
2.  $(b_{12}, b_{22}) = p(0, 1)$  com  $p < 0$ .

### Prova:

Pela hipótese de não singularidade dos elementos de  $\partial_B G(x^*)$  temos que  $\det(B) \neq 0$ , o que implica:

$$\begin{aligned} b_{11} &\neq 0 \\ b_{11} b_{22} &\neq b_{12} b_{21}. \end{aligned}$$

Agora,

$$\bar{B} = \begin{pmatrix} -b_{11} & -b_{12} \\ (\alpha - 1)\mathbf{e}_2^T + (\beta - 1)(b_{21}, b_{22}) \end{pmatrix} \in Z_* \subset \partial_B \Phi(x^*),$$

com  $\alpha^2 + \beta^2 = 1$ .

Nosso interesse é, sob a hipótese de não singularidade dos elementos de  $\partial_B G(x^*)$ , encontrar um critério que permita decidir se existem ou não elementos *singulares* em  $Z_* \subset \partial_B \Phi(x^*)$ . Podemos então perguntar o seguinte:

*Existirá  $\lambda \in \mathbb{R}$  tal que*

$$(\alpha - 1)\mathbf{e}_2^T + (\beta - 1)(b_{21}, b_{22}) = \lambda(-b_{11}, -b_{12})? \quad (4.4)$$

Em primeiro lugar, suponhamos  $(b_{21}, b_{22}) \neq p(0, 1)$ . com  $p \neq 0$ .

De (4.4), tem-se:

$$\beta = 1 - \frac{b_{11}}{b_{21}} \lambda \quad (4.5)$$

$$\alpha = 1 + \left( \frac{b_{11} b_{22}}{b_{21}} - b_{12} \right) \lambda. \quad (4.6)$$

De  $\alpha^2 + \beta^2 = 1$ , (4.5) e (4.6):

$$\left( 1 - \frac{b_{11}}{b_{21}} \lambda \right)^2 + \left( 1 + \left( \frac{b_{11} b_{22}}{b_{21}} - b_{12} \right) \lambda \right)^2 = 1.$$

Assim, temos um polinômio em  $\lambda$  dado por:

$$p(\lambda) = a \lambda^2 + b \lambda + c = 0,$$

onde,

$$a = \left( \frac{b_{11}}{b_{21}} \right)^2 + \left( \frac{b_{11} b_{22}}{b_{21}} - b_{12} \right)^2$$

$$b = -2 \frac{b_{11}}{b_{21}} + 2 \left( \frac{b_{11} b_{22}}{b_{21}} - b_{12} \right)$$

$$c = 1.$$

O valor do discriminante de  $p(\lambda)$  é

$$b^2 - 4ac = 8 \left( \frac{b_{11} b_{12}}{b_{21}} - \frac{b_{11}^2 b_{22}}{b_{21}^2} \right) \quad (4.7)$$

$$= -8 \frac{b_{11}}{b_{21}^2} \det(B). \quad (4.8)$$

Portanto, se  $b_{11}$  e  $\det(B)$  têm sinais opostos<sup>1</sup> existem valores reais de  $\lambda$  que satisfazem (4.4) i.e., existem elementos singulares em  $Z_* \subset \partial_B \Phi(x^*)$ .

---

<sup>1</sup> $\det B \neq 0$  pela hipótese de não singularidade dos elementos de  $\partial_B G(x^*)$ .

Agora, suponhamos  $(b_{21}, b_{22}) = p(0, 1)$ ,  $p \neq 0$ . Logo,

$$B_*(z) = \begin{pmatrix} -b_{11} & -b_{12} \\ 0 & (\alpha - 1) + (\beta - 1)p \end{pmatrix},$$

assim,  $\det(B_*(z)) = b_{11}(1 - \alpha - \beta p + p)$ .

Assim  $\det(B_*(z)) = 0$  se e somente se<sup>2</sup>:

$$(1 - \alpha - \beta p + p) = 0,$$

i.e., se e somente se

$$\alpha = (1 + p) - \beta p. \quad (4.9)$$

Substituindo esta expressão em  $\alpha^2 + \beta^2 = 1$ , temos um polinômio de segundo grau em  $\beta$  cujo discriminante é  $-8p$ . Portanto,  $\det(B_*(z)) = 0$  se e somente se  $p < 0$ , o que conclui a prova do **Teorema 4.1**. ■

Mas, BD-regularidade de  $\Phi$  em  $x^*$  implica BD-regularidade de  $G$  em  $x^*$ , isto é formalizado no seguinte **Lema**.

**Lema 4.1:** *Sejam  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $F \in C^1$ ,  $G$  e  $\Phi$  definidas como em (4.1) e (4.2). Se todas as matrizes de  $Z^*$  são não singulares então todas as matrizes em  $\partial_B G(x^*)$  são não singulares.*

**Prova:** Dado que qualquer matriz  $B \in \partial_B G(x^*)$  pode ser vista como  $-B_*(z) \in Z_*$  se tomamos  $\alpha = 1, \beta = 0$  ou  $\alpha = 0, \beta = 1$  então  $\partial_B G(x^*) \subset Z_*$  e assim, se todos os elementos de  $Z_*$  são não singulares, então  $G$  é BD-regular em  $x^*$ . ■

---

<sup>2</sup>Já que  $b_{11} \neq 0$ .

Voltando ao exemplo 2 observamos que para

$$B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \in \partial_B G(x^*),$$

tem-se que  $b_{11}$  e  $\det(B)$  têm o mesmo sinal. Assim, nesse exemplo, todas as matrizes em  $Z_* \subset \partial_B \Phi(x^*)$  são não singulares.

Na situação geral, i.e., para  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $F \in C^1$ , seja  $x^*$  uma solução degenerada do PCN. Suponhamos que  $x^*$  tem  $m$  degenerações. Sem perda de generalidade suporemos que as degenerações se apresentam nas últimas  $m$  componentes:

$$\begin{aligned} x_i^* &> f_i(x^*) = 0 \quad i = 1, \dots, n - m, \\ x_i^* &= f_i(x^*) = 0 \quad i = n - m + 1, \dots, n. \end{aligned}$$

Logo,  $\partial_B G(x^*)$  tem no máximo  $2^m$  elementos, que vamos supor *não singulares*.

Seja,

$$B = \begin{pmatrix} [B]_1 \\ \vdots \\ [B]_n \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ij} & \cdots & b_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{nn} \end{pmatrix} \in \partial_B G(x^*).$$

Assim, as componentes de  $B_*(z) \in Z_* \subset \partial_B \Phi(x^*)$ , são da forma:

$$[B_*(z)]_i = \begin{cases} -[B]_i & \text{se } i = 1, \dots, n - m \\ (\alpha_i - 1)e_i^T + (\beta_i - 1)[B]_i & \text{se } i = n - m + 1, \dots, n. \end{cases}$$

Como no caso  $n = 2$ , vamos perguntar se existe  $\lambda \in \mathbb{R}$  tal que:

$$(\alpha_j - 1)e_j^T + (\beta_j - 1)[B]_j = -\lambda[B]_i, \tag{4.10}$$

para alguns  $i = 1, \dots, n - m$  e  $j = (n - m + 1), \dots, n$  fixos.

Se  $[B]_j \neq p e_j^T$ ,  $p \neq 0$  então  $b_{ik} \neq 0$  para pelo menos um  $k$ ,  $k = 1, \dots, n$ .

Sem perda de generalidade suponha que  $[B]_j$  tem as primeiras  $j$  componentes não nulas. Assim, de (4.10) temos:

$$\alpha_j = 1 + \frac{1}{j-1} \left( \sum_{k=1}^{j-1} \frac{b_{ik} b_{jj}}{b_{jk}} - (j-1)b_{ij} \right) \lambda. \quad (4.11)$$

$$\beta_j = 1 - \frac{1}{j-1} \left( \sum_{k=1}^{j-1} \frac{b_{ik}}{b_{jk}} \right) \lambda \quad (4.12)$$

$$0 = -b_{ik} \lambda, \quad k = j+1, \dots, n. \quad (4.13)$$

Substituindo (4.11) e (4.12) em  $\alpha_j^2 + \beta_j^2 = 1$ , obtemos o seguinte polinômio em  $\lambda$ :

$$p(\lambda) = a \lambda^2 + b \lambda + c = 0,$$

onde,

$$a = \frac{1}{(j-1)^2} \left[ \left( \sum_{k=1}^{j-1} \frac{b_{ik} b_{jj}}{b_{jk}} - (j-1)b_{ij} \right)^2 + \left( \sum_{k=1}^{j-1} \frac{b_{ik}}{b_{jk}} \right)^2 \right]$$

$$b = \frac{2}{(j-1)^2} \left[ \left( \sum_{k=1}^n \frac{b_{ik} b_{jj}}{b_{jk}} - (j-1)b_{ij} \right) - \left( \sum_{k=1}^{j-1} \frac{b_{ik}}{b_{jk}} \right) \right]$$

$$c = 1.$$

O valor do discriminante,  $b^2 - 4ac$ , de  $p(\lambda)$  é

$$-\frac{8}{(n-1)^2} \left( \sum_{k=1}^{j-1} \frac{\begin{vmatrix} b_{ik} & b_{ij} \\ b_{jk} & b_{jj} \end{vmatrix}}{b_{jk}} \right) \left( \sum_{k=1, k \neq j}^n \frac{b_{ik}}{b_{jk}} \right).$$

Portanto, se as expressões entre parênteses têm sinais opostos, então existem valores reais de  $\lambda$  que satisfazem (4.10) i.e., existem elementos singulares em  $Z_* \subset \partial_B \Phi(x^*)$ .

Observe que, em particular, se  $[B]_j$  não tem componentes nulas, tem-se:

$$\alpha_j = 1 + \frac{1}{n-1} \left( \sum_{k=1, k \neq j}^n \frac{b_{ik} b_{jj}}{b_{jk}} - (n-1)b_{ij} \right) \lambda. \quad (4.14)$$

$$\beta_j = 1 - \frac{1}{n-1} \left( \sum_{k=1, k \neq j}^n \frac{b_{ik}}{b_{jk}} \right) \lambda. \quad (4.15)$$

Substituindo (4.14) e (4.15) em  $\alpha_j^2 + \beta_j^2 = 1$ , obtemos o seguinte polinômio em  $\lambda$ :

$$p(\lambda) = a \lambda^2 + b \lambda + c = 0,$$

onde,

$$a = \frac{1}{(n-1)^2} \left[ \left( \sum_{k=1, k \neq j}^n \frac{b_{ik} b_{jj}}{b_{jk}} - (n-1)b_{ij} \right)^2 + \left( \sum_{k=1, k \neq j}^n \frac{b_{ik}}{b_{jk}} \right)^2 \right]$$

$$b = \frac{2}{(n-1)^2} \left[ \left( \sum_{k=1, k \neq j}^n \frac{b_{ik} b_{jj}}{b_{jk}} - (n-1)b_{ij} \right) - \left( \sum_{k=1, k \neq j}^n \frac{b_{ik}}{b_{jk}} \right) \right]$$

$$c = 1.$$

Logo, o valor do discriminante,  $b^2 - 4ac$ , de  $p(\lambda)$  é

$$-\frac{8}{(n-1)^2} \left( \sum_{k=1}^{j-1} \frac{\begin{vmatrix} b_{ik} & b_{ij} \\ b_{jk} & b_{jj} \end{vmatrix}}{b_{jk}} - \sum_{k=j+1}^n \frac{\begin{vmatrix} b_{ik} & b_{ij} \\ b_{jk} & b_{jj} \end{vmatrix}}{b_{jk}} \right) \left( \sum_{k=1, k \neq j}^n \frac{b_{ik}}{b_{jk}} \right).$$

Portanto, se as expressões entre parênteses têm sinais opostos, então existem elementos singulares em  $Z_* \subset \partial_B \Phi(x^*)$ .

Se  $[B]_j = p \mathbf{e}_j^T$ ,  $p \neq 0$  então (4.10) é equivalente a:

$$[(\alpha_j - 1) + (\beta_j - 1)p] \mathbf{e}_j^T = -\lambda [B]_i, \quad (4.16)$$

para alguns  $i = 1, \dots, n - m$  e  $j = (n - m + 1), \dots, n$  fixos.

Assim de (4.16), existirá um valor real de  $\lambda$  que satisfaça (4.10) se e somente se

$$\lambda = -\frac{(\alpha_j - 1) + (\beta_j - 1)p}{b_{ij}},$$

e  $[B]_i = \mathbf{e}_j^T$ . Neste caso existirão elementos singulares em  $Z_*$ .

# Capítulo 5

## Testes numéricos

Neste capítulo analisamos a sensibilidade das funções Mínimo e Fischer a degenerações na solução do PCN e o comportamento local dos algoritmos propostos nos **Capítulos 2 e 3**, respectivamente. Para isto consideramos a função  $F$  definida em ([8]), i.e.,

$$\begin{aligned} F : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto F(x) = (f_1(x), \dots, f_n(x))^T \end{aligned}$$

onde,

$$f_i(x) = \begin{cases} h_i(x) - h_i(x^*) & \text{se } i \text{ é ímpar ou } i > n/2, \\ h_i(x) - h_i(x^*) + 1 & \text{em outro caso,} \end{cases}$$

sendo que,

$$x^* = (1, 0, 1, 0 \dots)^T \in \mathbb{R}^n.$$

Para  $i = 1, \dots, n$ ,  $h_i$  são as 17 funções definidas em ([14]).

Neste caso a função  $F$  não é diferenciável em  $x^*$ , visto que, para  $i$  par e  $i > n/2$  tem-se:

$$f_i(x^*) = x_i^* = 0.$$

Assim,  $x^*$  é solução do PCN associado a  $F$  e portanto  $x^*$  é solução degenerada dos sistemas não lineares  $G(x) = 0$  e  $\Phi(x) = 0$  com  $G$  e  $\Phi$  definidas respectivamente por (2.1) e (3.2).

## 5.1 Sensibilidade em soluções degeneradas

Para  $F$ ,  $G$  e  $\Phi$  definidas desta forma, usando as 17 funções definidas em ([14]) e com  $n$  fixo, fizemos os seguintes testes:

1. Calculamos o número de condição de cada matriz de  $\partial_B G(x^*)$  e em cada caso determinamos o número de condição mínimo e o máximo.
2. Para uma amostra aleatória de vetores  $z$ , calculamos os números de condição das matrizes  $B_*(z) \in Z_* \subset \partial_B \Phi(x^*)$  e em cada caso, determinamos o mínimo e o máximo número de condição.
3. Maximizamos o número de condição dos elementos de  $Z_*$  como função de  $z$ .

Todos os testes foram feitos usando o MATLAB e os resultados são apresentados nas seguintes tabelas:

A **Tabela 1** resume os resultados dos itens 1. e 2. Nas duas primeiras colunas fazemos referência ao número e à dimensão, respectivamente, do problema considerado. (**Prob**, **Dim**)

Prob.	Dim.	Mínimo		Fischer	
		<i>min.</i>	<i>max.</i>	<i>min.</i>	<i>max.</i>
1	6	18.18	28.37	17.72	43.61
2	6	15.45	38.45	19.09	54.45
3	10	2.82	5.50	5.50	157.71
4	6	14.36	15.63	14.98	21.52
5	7	$\infty$	$\infty$	$\infty$	$\infty$
6	6	34.73	$\infty$	33.53	$\infty$
7	6	32.13	33.00	32.35	35.69
8	6	34.03	35.01	34.03	46.62
9	6	32.79	34.18	36.16	71.31
10	6	24.37	35.37	23.93	48.52
11	6	25.01	50.08	32.64	543.21
12	8	116.46	256.20	117.09	846.31
13	8	22.29	76.90	38.03	256.80
14	6	6.15	38.47	5.93	645.89
15	6	23.36	25.58	24.27	27.58
16	6	3.87	7.23	5.49	13.14
17	6	9.52	10.93	19.86	24.79

**Tabela 1:** Comparação entre números de condição de matrizes de  $\partial_B G(x^*)$  e de  $Z_*$ .

A Tabela 2 resume os resultados do item 3.

Prob.	Dim.	Mínimo	Fischer
1	6	28.37	$\infty$
2	6	38.45	$\infty$
3	10	5.50	$\infty$
4	6	15.63	22.34
5	7	$\infty$	$\infty$
6	6	$\infty$	$\infty$
7	6	33.00	$\infty$
8	6	35.01	$\infty$
9	6	34.18	$1.83 \times 10^{18}$
10	6	35.37	48.55
11	6	50.08	$\infty$
12	8	256.20	$\infty$
13	8	76.90	253.86
14	6	38.47	$2.19 \times 10^{19}$
15	6	25.58	31.89
16	6	7.23	13.23
17	6	10.93	$2.98 \times 10^{18}$

**Tabela 2:** O máximo número de condição das matrizes de  $\partial_B G(x^*)$  para a função Mínimo e a maximização do número de condição das matrizes em  $Z_*$  para a função de Fischer.

Nesta última tabela encontramos os resultados mais interessantes, os quais nos permitem concluir que a função de Fischer é muito mais sensível a degenerações na solução  $x^*$  que a função Mínimo. Isto afetará a convergência local do método que usa a função de Fischer, i.e., em problemas degenerados, se para ambas as reformulações do PCN tem-se convergência, a convergência da reformulação que usa a função de Fischer pode ser mais lenta que a reformulação que usa a função Mínimo.

## 5.2 Comportamento local dos algoritmos 2.1 e 3.1

No que segue analisamos o comportamento local dos algoritmos propostos nos Capítulos 2 e 3, respectivamente. Para isto, implementamos os algoritmos 2.1 e 3.1, em MATLAB, e os testamos em uma estação SUN SPARC 2 usando os 17 problemas teste gerados e sugeridos em [8]. Para ambos os casos, testamos e comparamos o método de Newton generalizado ([21]) e os algoritmos 2.1 e 3.1, usando a fórmula esparsa de Schubert para atualizar a matriz  $A_k$ . ([22])

Declaramos convergência em  $x^k$  sempre que:

$$\|G(x^k)\|_2 < \sqrt{n}10^{-5} \quad \text{para o Algoritmo 2.1,}$$

$$\|\Phi(x^k)\|_2 < \sqrt{n}10^{-5} \quad \text{para o Algoritmo 3.1,}$$

e divergência se  $k > 100$  ou

$$\|G(x^k)\|_\infty > 10^{20} \quad \text{para o Algoritmo 2.1,}$$

$$\|\Phi(x^k)\|_\infty > 10^{20} \quad \text{para o Algoritmo 3.1.}$$

Para todos os problemas o vetor inicial foi  $x^0 = (0.9, 0.1, \dots, \dots)^T \in \mathbb{R}^n$ . Os resultados deste conjunto de experimentos são resumidos na **Tabela 3** onde **Prob** denota o número do problema considerado, **Dim** é a dimensão do problema e cada uma das outras colunas resume o número total de iterações efetuadas por cada método.

Se houve convergência para uma solução do PCN diferente de  $x^*$  isto é indicado com um \* depois do número de iterações. A divergência é indicada com - .

O problema 6 do conjunto de testes não satisfaz a hipótese de não singularidade (Capítulos 2 e 3), nem para a função Mínimo nem para a função de Fischer e por isso este problema não é considerado no conjunto de experimentos.

Prob	Dim	Newton Mínimo	Newton Fischer	Alg. 2.1 Mínimo	Alg. 3.1 Fischer
1	100	3	4	3	6
2	100	4	5	3	6
3	100	-	-	-	-
4	100	3	4	4	4
5	101	-	-	-	-
7	100	4	4	4	6
8	100	3	4	6	-
9	100	5	4	6	6
10	100	4	4	7	7
11	100	1*	-	1*	-
12	100	1*	6*	-	-
13	100	-	-	-	-
14	100	-	-	-	-
15	100	15	17	-	-
16	100	2	4	2	4
17	100	5	5	6	6

**Tabela 3:** Comportamento local do método de Newton generalizado e dos Algoritmos 2.1 e 3.1 usando a fórmula esparsa de Schubert.

Os resultados da **Tabela 3** mostram que, para este conjunto de experimentos, o comportamento local do método que usa a função Mínimo é ligeiramente melhor do que aquele que usa a função de Fischer. Esta característica também foi observada recentemente em ([12]) através de experimentos numéricos e os autores usam esta observação para introduzir uma estratégia de globalização para resolver o PCN.

# Capítulo 6

## Estratégia de globalização

Baseados no bom comportamento local da função Mínimo, provado tanto em forma teórica como numérica nos capítulos anteriores, e no fato que a soma de quadrados do sistema relacionado com a função de Fischer é diferenciável,<sup>1</sup> apresentamos neste **Capítulo** um algoritmo híbrido global para resolver o PCN.

A idéia do algoritmo é a seguinte: iniciamos com o método local que usa a função Mínimo e continuamos usando o método se  $\|G(x)\|$  está decrescendo em um certo sentido. Caso contrário usamos um algoritmo global de minimização, o qual usa a função de Fischer. Mas, em cada iteração deste algoritmo de minimização estamos controlando o comportamento monótono de  $\|G(x)\|$ .

O algoritmo global de minimização<sup>2</sup> que estamos usando e que usa a função de Fischer é proposto em ([2]). Para maior clareza incluímos o algoritmo no texto.

Apresentamos alguns resultados numéricos do comportamento do algoritmo híbrido e comparamos estes resultados com os obtidos usando o algoritmo proposto em ([2]).

---

<sup>1</sup>métodos globalmente convergentes relacionados com sua minimização podem ser desenvolvidos. Infelizmente, esse não é o caso da função Mínimo.

<sup>2</sup>E que resolve também o PCN.

## 6.1 Iterações ordinárias e especiais

No que segue apresentamos os dois algoritmos que combinamos formando um algoritmo híbrido para resolver o PCN.

O **Algoritmo 6.1**, é um algoritmo local tipo Newton Generalizado, i. e, as matrizes  $B_k \in \partial_B G(x^k)$ .

### Algoritmo 6.1

Sendo dado um ponto inicial  $x^0$ , para  $k = 1, 2 \dots, x^{k+1}$  é gerado como segue:

$$x^{k+1} = x^k - B_k^{-1} G(x^k),$$

onde,  $B_k \in \partial_B G(x^k)$ , é calculada como em (2.3).

Uma iteração produzida por este algoritmo será chamada de **iteração ordinária**.

O **Algoritmo global de minimização** que usamos em nosso **Algoritmo híbrido** é o proposto em [2] e usa somente a função de Fischer. Esse algoritmo resolve o PCN resolvendo o problema equivalente:

$$\min \Psi(x)$$

onde,

$$\Psi(x) = \frac{1}{2} \|\Phi(x)\|_2^2 = \frac{1}{2} \sum_{i=1}^n \phi(x_i, f_i(x))^2.$$

Observe-se que  $\Psi \in C^1$ . Além disto, para cada matriz  $V \in \partial_B \Phi(x^k)$ ,

$$\nabla \Psi(x^k) = V^T \Phi(x^k).$$

Ver ([2]).

Para maior clareza, apresentamos o algoritmo a seguir.

## Algoritmo 6.2

**P.0** : (Inicialização)  $x^0 \in \mathbb{R}^n$ ,  $\rho > 0$ ,  $p > 2$ ,  $\sigma \in (0, 1/2)$ ,  $\beta \in (0, 1)$ ,  
 $\varepsilon \geq 0$ ,  $k = 0$

**P.1** : (Critério de parada)  $\|\nabla\Psi(x^k)\| \leq \varepsilon$ .

**P.2** : (Cálculo da direção de descida)

Escolha um elemento  $H_k \in \partial_B \Phi(x^k)$ . Encontre a solução  $d^k$  do sistema:

$$H_k d = -\Phi(x^k).$$

Se o sistema não tem solução ou se a condição

$$\nabla\Psi(x^k)^T d^k \leq -\rho \|d^k\|$$

não é satisfeita, fixe  $d^k = -\nabla\Psi(x^k)$ .

**P.3** : (Busca linear) Encontre  $t_k := \max\{\beta^l : l = 0, 1, 2, \dots\}$  tal que

$$\Psi(x^k + t_k d^k) \leq \Psi(x^k) + \sigma t_k \nabla\Psi(x^k)^T d^k.$$

**P.4** : (Atualização)

$$x^{k+1} = x^k + t_k d^k,$$

$k \leftarrow k + 1$ , e volte ao passo P.1.

Uma iteração gerada por este algoritmo será chamada de **iteração especial**.

## 6.2 Algoritmo híbrido

**Iterações ordinárias e iterações especiais**, definidas na seção anterior são combinadas no **Algoritmo híbrido 6.3**, que é um algoritmo tolerante no sentido que permite uma seqüência de iterações locais até que o decréscimo em  $\|G(x^k)\|$  não seja satisfatório. Os parâmetros  $q$  e  $\gamma$ , no algoritmo definirão dita tolerancia.

Iniciamos o **Algoritmo híbrido** com uma iteração do **Algoritmo 6.1** e a cada iteração, fazemos decrescer  $\|G(x^k)\|$ . Quando isto não é possível usando somente o **algoritmo 6.1** então usamos o **Algoritmo de minimização 6.2**.

Definimos, para todo  $k \in N$ ,

$$w^k = \operatorname{argmin}\{\|G(x^0)\| \cdots \|G(x^k)\|\}.$$

Para completar, definimos  $\|G(w^k)\| = \|G(x^0)\|$  se  $k < q$ .

### Algoritmo híbrido 6.3 . (Mínimo - Fischer )

**P . 0** :  $k = 0$ ,  $x^0 \in \mathbb{R}^n$ ,  $FLAG = 1$ . Sejam  $q \geq 0$  um inteiro,  $\gamma \in (0, 1)$ .

**P . 1** : Se  $FLAG = 1$  obtenha  $x^{k+1}$  usando a **iteração ordinária**.

Caso contrário obtenha  $x^{k+1}$  usando a **iteração especial**.

**P . 2** : Se

$$\|G(x^{k+1})\| \leq \gamma \|G(w^{(k-q)})\|$$

então  $FLAG = 1$ ,  $k \leftarrow k + 1$  e volte ao passo **P.1**.

Caso contrário redefine  $x^{k+1} \leftarrow w^{k+1}$ ,

$FLAG = -1$ ,  $k \leftarrow k + 1$  e volte ao passo **P.1**.

## 6.3 Testes numéricos

O **Algoritmo 6.3** foi implementado em MATLAB e testado numa estação SUN SPARC 2 usando os 17 problemas teste gerados e sugeridos em ([8]) com os pontos iniciais propostos em ([14]).

Declaramos convergência em  $x^k$  sempre que  $\|G(x^k)\|_2 < \sqrt{n} 10^{-5}$  e declaramos divergência se o número de iterações é maior que 100 ou  $t_k < t_{min}$ .

Usamos os seguintes valores para nossos parâmetros:

$$\begin{aligned} \gamma &= 0.9 & q &= 5 & \rho &= 10^{-8} & \beta &= 0.5 \\ \sigma &= 10^{-4} & p &= 2.1 & t_{min} &= 10^{-12} \end{aligned}$$

Seguindo o que foi feito em ([2]), na implementação do **Algoritmo 6.2**, incorporamos em **P.2** a estratégia de busca linear não monótona proposta em ([9]).

Os resultados deste conjunto de experimentos estão resumidos na **Tabela 4** onde apresentamos o comportamento do método global que usa as funções Mínimo e Fischer (**Algoritmo híbrido**), com o algoritmo global que usa somente a função de Fischer (**Algoritmo 6.2**)

Para cada teste reportamos o número do problema (**Prob**), a dimensão do problema (**Dim**) e o número total de iterações efetuado por cada método. Para o caso do **Algoritmo Mínimo- Fischer** reportamos o número de iterações seguido de (O, E) onde O é o número total de iterações ordinárias e E representa o número total de iterações especiais.

Se a convergência ocorreu numa solução do PCN diferente de  $x^*$  isto é indicado com um \* depois do número de iterações. A divergência é indicada com -.

Da mesma forma que no **Capítulo 5**, não incluímos o problema 6 do conjunto de testes porque ele não satisfaz a hipótese 3 da teoria de convergência desenvolvida nos **Capítulos 2, 3** respectivamente.

Prob.	Dim.	Mínimo-Fischer	Fischer
1	n=100	4 (4,0)	6
2	n= 100	6 (1,5)	29
3	n=100	-	-
4	n=100	10* (7,3)	13*
5	n=101	-	-
7	n=100	-	-
8	n=100	-	-
9	n=100	-	-
10	n=100	6 (6,0)	7
11	n=100	1 (1,0)	4
12	n=100	23 (9,14)	13
13	n=100	11 (6,5)	11
14	n=100	12 (7,5)	-
15	n=100	-	-
16	n=100	4* (4,0)	6*
17	n=100	7 (7,0)	7

**Tabela 4:** Comparação entre o Algoritmo híbrido 6.3 e Algoritmo global 6.2

Observamos na **Tabela 4** que, na maioria dos casos de convergência de ambos os algoritmos, o **Algoritmo 6.3** efetua menos iterações que o **Algoritmo 6.2**. No caso de convergência a outra solução, ambos os algoritmos convergem ao mesmo ponto.

Nos problemas onde as *iterações especiais* não foram acionadas (1, 10, 11, 16 e 17), era de se esperar que o número de iterações fosse menor, porque nesta situação tem-se Mínimo x fischer.

Dos 16 problemas, 6 não convergiram. Dos 10 restantes em apenas 5 as *iterações especiais* foram acionadas. Chaman atenção os problemas 2, 12, e 14. Em 14 somente o híbrido converge, indicando uma possibilidade deste algoritmo ser mais robusto. Contudo, o fato de não ter conseguido convergência em 6 problemas do conjunto de 16 impossibilita uma conclusão mais segura no sentido de robustez.

## 6.4 Uma aplicação

O Problema de Complementaridade Não Linear (PCN) tem tido, nas últimas três décadas, um papel central na modelagem matemática de problemas em diferentes áreas. A importância do PCN em Engenharia e Economia deve-se ao fato de que o conceito de complementaridade é sinônimo da noção de **sistema em equilíbrio**. Recentemente Ferris e Pang (ver [5]), fizeram uma documentação extensa de aplicações de Problemas de Complementaridade Não Linear de dimensão finita a problemas de Engenharia e Economia, com o objetivo de estimular a investigação em aproximações de complementaridade para resolver esses e outros problemas.

A seguir consideramos um desses problemas, conhecido como **problema de valor de capital invariante** ([5], [19], [10]). Nosso interesse é aplicar a teoria desenvolvida nos capítulos anteriores, (em particular testar o algoritmo híbrido, proposto no capítulo 4) a tal problema.

No **problema de valor de capital invariante**, o interesse é encontrar um capital inicial  $z$  que seja invariante sob o modelo de otimização (6.1). Para isto, segundo ([19]), se supõe que uma economia se desenvolve em um número infinito de períodos de tempo com produção e recursos constantes em cada período. No início de cada período de tempo  $t$ , a economia investe seus **bens de capital**  $z_t$  no processo de produção os quais produzem **bens de capital** e **bens de consumo**. O capital produzido no período  $t$ ,  $z_{t+1}$ , será investido no período seguinte, enquanto os **bens de consumo**  $x_t$ , produzidos nesse período vão determinar a **utilidade** do investimento. A **utilidade total** do investimento é uma soma ponderada, isto é, a **utilidade** ganha por consumo no tempo  $t$  é ponderada pelo fator  $\alpha_t \in (0, 1)$ .

O problema pode ser formulado da seguinte forma: se cada  $t = 1, 2, \dots$ ,  $A$  é a matriz de investimento de capital,  $B$  é a matriz de capital final nesse período,  $C$  é a matriz de investimento de recursos, e  $w$  é um vetor de disponibilidade de recursos, que é assumido constante para todos os períodos de tempo, então

$$\begin{aligned}
 \max \quad & \sum_{t=0}^{\infty} \alpha_t \mathcal{V}(x_t) \\
 \text{s.a} \quad & Ax_t \leq z_t \\
 & Bx_t \geq z_{t+1} \\
 & Cx_t \leq w \\
 & x_t \geq 0,
 \end{aligned} \tag{6.1}$$

onde,  $\mathcal{V}(x)$  é a utilidade proveniente do consumo especificado por  $x$ . Além disso, assume-se que  $\mathcal{V}$  é **côncava** e continuamente diferenciável.

Portanto, as condições de factibilidade em (6.1) conectam o vetor de capital inicial do  $t$ -ésimo período,  $z_t$ , o vetor de capital final nesse período,  $z_{t+1}$  e o vetor disponibilidade de recursos  $w$ , com o vetor de consumo  $x_t$ . Se quer encontrar um valor  $z_0 = z$  para o valor de capital inicial que forneça como saída um **caminho de crescimento constante ótimo**  $\{(x_t, z_t)\}$  que seja tal que  $x_t = x$  e  $z_t = z$ , para todo  $t$ .

Um **valor de capital**  $z \geq 0$  é chamado *reproduzível* se as condições de factibilidade em (6.1) para  $t = 0$  permitem que no final de um único período  $z_1 = z_0 = z$ .

Hansen e Koopmans ([10]), demonstram que existe uma equivalência entre o problema de encontrar um **valor de capital invariante** e o problema com período único, porém com uma restrição adicional, o que é descrito a seguir em dois passos:

**P1** : Dado um valor  $z$  *reproduzível*, resolver o problema:

$$\begin{aligned}
 \max \quad & \mathcal{V}(x) \\
 \text{s.a} \quad & Ax \leq z
 \end{aligned}$$

$$\begin{aligned} Bx &\geq z \\ Cx &\leq w \\ x &\geq 0, \end{aligned} \tag{6.2}$$

**P2** : Dado um fator de ponderação  $\alpha \in (0, 1)$ , escolha um valor reproduzível  $z$  em **P1** de tal forma que exista uma solução  $x(z)$  de (6.2) com multiplicadores de Lagrange associados,  $y$  e  $u$ , que satisfaçam  $u = \alpha y$ .

O valor  $z$  satisfazendo **P1** e **P2** é chamado de **valor de capital ótimo invariante**. Assim, usando essa informação tem-se que um valor de capital inicial  $z_0$  dá um **caminho de crescimento constante ótimo**  $\{(z_t, x_t)\}$  se e somente se existe  $(x, y, u)$  resolvendo o PCN:

$$\begin{array}{rcl} 0 \leq x & \perp & -\nabla\mathcal{V}(x) + (A - \alpha B)^T y + C^T u \geq 0 \\ 0 \leq y & \perp & (B - A)x \geq 0 \\ 0 \leq u & \perp & -Cx + w \geq 0 \end{array}$$

com  $x_t = x$ ,  $z_t = Ax$ , para todo  $t$ .

Implementamos o exemplo dado em ([19]) no qual se considera uma economia com 3 **bens de consumo**, 2 **bens de capital**, e 2 **recursos** (mão de obra qualificada e não qualificada). Cada **bem** pode ser produzido por dois processos alternativos. Assim, tem-se um total de 10 processos, dos quais os processos 1 até 6 produzem bens de consumo.

O modelo de *valor de capital invariante* pode ser formulado como o PCN com a função  $F : \mathbb{R}_+^{10} \times \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^{10} \times \mathbb{R}^2 \times \mathbb{R}^2$  definida por

$$F(x, y, u) = \begin{pmatrix} -\nabla\mathcal{V}(x) + (A - \alpha B)^T y + C^T u \\ (B - A)x \\ -Cx + w \end{pmatrix},$$

onde, a função utilidade  $\mathcal{V} : \mathbb{R}_+^6 \rightarrow \mathbb{R}$  definida da seguinte forma:

$$\mathcal{V}(x) = (x_1 + 2.5x_2)^{0.2} (2.5x_3 + x_4)^{0.2} (2x_5 + 3x_6)^{0.2},$$

que é côncava e continuamente diferenciável.

As matrizes  $A$ ,  $B$  e  $C$  são dadas por: (ver [10])

$$A = \begin{pmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 2 & 2 & 1 & 1 & 1 & 0.5 & 1 & 0.5 \end{pmatrix}$$

$$B = \begin{pmatrix} 1.5 & 1.5 & 1.5 & 1.5 & 1.5 & 1.5 & 4 & 3 & 1.5 & 1.5 \\ 2.7 & 2.7 & 1.8 & 1.8 & 0.9 & 0.9 & 0.9 & 0.4 & 2 & 1.5 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.5 & 1.5 & 1.5 & 0.5 & 0.5 & 1.5 & 1.5 & 0.5 & 0.5 & 1.5 \end{pmatrix}$$

e a disponibilidade de recursos dada por:

$$w = (0.8, 0.8).$$

O **Algoritmo híbrido 6.3** proposto na seção anterior foi testado usando como valores de desconto  $\alpha : 0.7, 0.8, 0.9$ , sugeridos em ([19]) e os pontos iniciais como em ([10]):

$$a = (0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0, 0, 0, 0)$$

$$b = (0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0, 0, 0, 0)$$

O número de iterações efetuadas pelo **Algoritmo híbrido**, o valor da **utilidade** obtida na solução e o **valor de capital invariante ótimo** são apresentados na seguinte tabela para cada um dos valores de  $\alpha$  mencionados anteriormente.

Desconto $\alpha$	Utilidade $\mathcal{V}(x^*)$	Iter. $a$ $(t_i, n_i)$	Iter. $b$ $(t_i, n_i)$	Valor capital $Ax^*$
0.7	0.48855	11 (5, 6)	12 (5, 7)	$(1.0667, 0.7794)^T$
0.8	0.52216	10 (4, 6)	10 (3, 7)	$(1.1789, 0.9258)^T$
0.9	0.55935	8 (4, 4)	11 (4, 7)	$(1.6000, 1.1706)^T$

Esta tabela permite observar que, a medida em que o valor de  $\alpha$  aumenta, a **utilidade atingida**, e o **valor de capital ótimo** também aumentam.

Os valores obtidos são análogos aos obtidos em [10], mas eles não reportam o número de iterações tomadas por seu algoritmo.

Em [19], para o mesmo exemplo e com  $\alpha = 0.7$ , o algoritmo converge em 10 iterações quando o ponto inicial é dado por  $a$  e 11 iterações quando partem de  $b$ .

# Capítulo 7

## Comentários finais

A técnica de reduzir Problemas de Complementaridade Não Linear a sistemas de equações não lineares é muito importante para resolver esse tipo de problemas porque, desta forma, o trabalho principal na maioria das iterações é a resolução de um único sistema linear. Na aproximação Newtoniana a matriz deste sistema é a matriz Jacobiana e se exige o cálculo da solução exata, enquanto que no método de Newton-Inexato somente se precisa de uma solução aproximada. Neste trabalho consideramos a aproximação Quase-Newton, que pode ser muito útil quando as derivadas do sistema são muito caras ou difíceis de obter.

O fato de que, no fundo uma iteração consista na resolução de um sistema linear junto com somente uma avaliação da função está associado à possibilidade de obter taxas de convergência altas (geralmente convergência superlinear) de métodos locais puros. Procedimentos de globalização são usualmente feitos de tal forma que iterações globais coincidam com iterações locais numa vizinhança da solução, mantendo desta forma convergência local rápida.

No entanto, convergência local rápida depende usualmente das características do problema, a principal das quais é a não singularidade dos Jacobianos (generalizados) do sistema não linear na solução considerada. De Luca, Facchinei e Kanzow ([2]) mostraram que, quando se usa o sistema não linear induzido pela aplicação da função de Fischer componente a componente, a não singularidade

dos Jacobianos Generalizados não está associada diretamente a degenerações na solução. Em outros termos, Jacobianos Generalizados não singulares podem ser encontrados em soluções degeneradas, e ainda nestes casos pode esperar-se convergência rápida. É fácil provar que os mesmos resultados se verificam para a clássica função Mínimo considerada nesta tese e para generalizações da função de Fischer introduzidas em ([12]). A soma de quadrados do sistema relacionado com a função de Fischer é suave, portanto métodos globalmente convergentes relacionados com sua minimização podem ser desenvolvidos. Infelizmente, esse não é o caso da função Mínimo.

Por outra parte foi provado nesta tese que Jacobianos Generalizados singulares podem aparecer em soluções degeneradas do Problema de Complementaridade Não Linear quando se usa a reformulação relativa à função de Fischer como resultado da forma algébrica dessa função de duas variáveis. Para entender geometricamente porque isto acontece, considere o Problema de Complementaridade em  $\mathbb{R}^2$ , definido por  $f_1(x_1, x_2) = x_1 + 2x_2 - 1$  e  $f_2(x_1, x_2) = x_1 + x_2 - 1$ , o qual tem solução não degenerada  $(0, 1)$  e solução degenerada em  $(1, 0)$ . Numa vizinhança de  $(1, 0)$  a reformulação relativa à função Mínimo é formada pela linha  $x_1 + 2x_2 - 1 = 0$  e a "curva" linear por partes  $\min\{x_2, x_1 + x_2 - 1\} = 0$ . Portanto, as curvas de nível das duas funções de Fischer envolvidas são aproximações suaves das curvas de nível de  $(x_1 + x_2 - 1)$  e  $\min\{x_2, x_1 + x_2 - 1\} = 0$ . É fácil ver, geometricamente, que o conjunto de pontos nos quais as curvas de nível da primeira função de Fischer são tangentes às curvas de nível da segunda formam uma curva contínua que emana de  $(1, 0)$ . Evidentemente, a Jacobiana da reformulação relativa à função de Fischer é singular em todos os pontos dessa curva e, assim, existe um Jacobiano Generalizado singular em  $(1, 0)$ . Claramente, esse fenômeno não ocorre no caso da função Mínimo. (Para a reformulação relativa à função Mínimo, o conjunto de Jacobianos Generalizados é formado por duas matrizes não singulares).

As observações anteriores parecem indicar que o desenvolvimento de teorias locais tipo-Newton para o sistema que usa a função Mínimo podem ser uma ferramenta útil para entender o comportamento de métodos práticos. Nesta tese desenvolvemos uma teoria **Least Change Secant Update** para métodos quase Newton baseados em projeções tipo secante. (As teorias de Newton e Newton-Inexato são simples aplicações de teorias existentes para sistemas semi-suaves.)

Finalmente, a associação entre métodos local e globalmente convergentes deveria ser considerada. Com este fim, sugerimos combinar estratégias locais baseadas na função Mínimo com estratégias globais baseadas na função de Fischer como as desenvolvidas em ([8]). Resultados numéricos preliminares mostraram que esta combinação vale a pena. No entanto muita investigação é necessária ao longo dessas linhas tanto do ponto de vista teórico como prático.

# Bibliografia

- [1] Clarke, F. H. *Optimization and Nonsmooth Analysis*. John Wiley & Sons, New York, 1983. ( Reprint by SIAM, Philadelphia, PA, 1990).
- [2] De Luca, T., Facchinei, F. and Kanzow, C. *A Semismooth Equation Approach to the Solution of Nonlinear Complementarity Problems*. *Mathematical Programming*, 75, pp 407-439, 1996.
- [3] Dennis, J. E. Jr and Walker, H. F. *Convergence Theorems for Least-Change Secant Update Methods*. *Siam Journal on Numerical Analysis*, 18, pp 949-987, 1981.
- [4] Facchinei, F. and Kanzow, C. *A Nonsmooth Inexact Newton Method for the Solution of Large Scale Nonlinear Complementarity Problems*. *Mathematical Programming*, 76, pp 493-512, 1997.
- [5] Ferris, M. C. and Pang, J-S. *Engineering and Economic Applications of Complementarity Problems* . Technical Report 95-07, Computer Science Department, University of Wisconsin, Madison, WI, 1995.
- [6] Fischer, A. *An NCP-Function and its Use for the Solution of Complementarity Problems*. *Recent Advances in Nonsmooth Optimization*. Editors: D-Z Du, L. Qi and R. Womersley, World Scientific, pp 88-105, 1995. .
- [7] Friesz, T. L., Tobin, R. L., Smith, T. E. and Harker, P. T. *A Nonlinear Complementarity Formulation and Solution Procedure for the General Derived Demand Network Equilibrium Problem*. *Journal of Reg. Sci.*, 23, pp 337-359, 1983.

- [8] Gomes-Ruggiero, M. A., Martínez J. M. and Santos, S. A. *Solving Nonsmooth Equations by Means of Quasi-Newton Methods with Globalization*. Recent Advances in Nonsmooth Optimization. Editors: D-Z Du, L. Qi and R. Womersley, World Scientific, pp 121-140, 1995.
- [9] Grippo, L., Lampariello, F. and Lucidi, S. *A Nonmonotone Linesearch Technique for Newton's method*. Siam Journal on Numerical Analysis 23, pp 707-716, 1986.
- [10] Hansen, T. and Koopmans, T. C. *On The Definition and Computation of a Capital Stock Invariant Under Optimization*. Journal of Economics Theory 5, pp 487-523, 1972.
- [11] Harker, P. T. and Pang, J. S. *Finite-dimensional Variational Inequality and Nonlinear Complementarity Problems: a Survey of Theory, Algorithms and Applications*. Mathematical Programming, 48, pp 161-220, 1990.
- [12] Kanzow, C. and Kleimichel, H. A. *A New Class of Semismooth Newton-Type Methods for Nonlinear Complementarity Problems*. Reihe A, Preprint 118, Hamburger Beiträge zur Angewandten Mathematik, 1997.
- [13] Lopes, V. L. R. and Martínez J. M. *On the Convergence of Quasi-Newton Methods for Nonsmooth Problems*. Numerical Functional Analysis and Optimization, 16, pp 1193-1209, 1995.
- [14] Lukšan, L. *Inexact Trust Region Method for Large Sparse Systems of Nonlinear Equations*. Journal of Optimization Theory and Applications, 81, pp 569-590, 1994.
- [15] Martínez, J. M. *Local Convergence Theory of Inexact Newton Methods Based on Structured Least Change Updates*. Mathematics of Computation, 55, pp 143-168, 1990.
- [16] Martínez, J. M. *On the Relation Between Two Local Convergence Theories of Least-Change Secant Update Methods*. Mathematics of Computation, 59, pp 457-481, 1992.
- [17] Martínez, J. M. and Zambaldi, M. C. *Least-Change Update Methods for Nonlinear Systems with Nondifferentiable Terms*. Numerical Functional Analysis and Optimization, 14, pp 405-415, 1993.

- [18] Oh, K. P. *The Formulation of the Mixed Lubrication Problem as a generalized Nonlinear Complementarity Problem*. Transactions of the ASME, 108, pp 598-604, 1986.
- [19] Pang, J-S. and Qi, L. *Nonsmooth Equations: Motivation and Algorithms*. SIAM Journal of Optimization, 3, N. 3, pp 443-465, 1993.
- [20] Pang, J-S. and Gabriel S. A. *NE/SQP: A Robust Algorithm for the Nonlinear Complementarity Problem*. Mathematical Programming 60, pp 295-337, 1993.
- [21] Qi, L. *Convergence Analysis of Some Algorithms for Solving Nonsmooth Equations*. Mathematics of Operational Research, 18, N. 1, pp 227-244, 1993.
- [22] Schubert, L. K. *Modification of a Quasi-Newton Method for Nonlinear Equations with a Sparse Jacobian*. Mathematics of Computation, 24, 1970, pp 27-30.
- [23] Sun, D. and Han, J. *Newton and Quasi-Newton Methods for a Class of Nonsmooth Equations and Related Problems*. SIAM Journal of Optimization, 7, N. 2, pp 463-480, 1997.