ANÁLISE DE CONVERGÊNCIA E DE IMPLEMENTAÇÃO DE UM ALGORITMO PROXIMAL PARA O PROBLEMA DE COMPLEMENTARIDADE NÃO-LINEAR

Rodrigo Olimpio Costa

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS–GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:

Prof. Paulo Roberto Oliveira, Doucter.

Geci José Pereira da Silva

Prof. Geci José Pereira da Silva, D.Sc.

Prof. Carlos Humes Jr., Ph.D.

Maria Helena Cautiero Jardim

Prof. Maria Helena Cautiero Horta Jardim, D.Sc.

RIO DE JANEIRO, RJ - BRASIL
OUTUBRO DE 2004

COSTA, RODRIGO OLIMPIO

Análise de Convergência e de Implementação de um Algoritmo Proximal Para o Problema de Complementaridade Não-Linear [Rio de Janeiro] 2004

VIII, 47 p. 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia de Sistemas e Computação, 2004)

Tese – Universidade Federal do Rio de Janeiro, COPPE

1 - Algoritmo Proximal

2 - Taxa de convergência

3 - Método de Newton

I. COPPE/UFRJ II. Título (série)

"Ser capaz de colocar continuamente em questão as suas próprias opiniões - esta é, para mim, a condição preliminar de qualquer inteligência"
(Ítalo Calvino)

- Aos meus pais, a minha irmã e aos meus familiares.
- À memória de minha avó Mercedes Costa.

Agradecimentos

- Primeiramente a Deus, que me deu forças para terminar este trabalho, apesar das dificuldades.

- A meus pais, a minha irmã e a todos aqueles que me apoiaram.

- Ao professor Paulo Roberto Oliveira, pela orientação e a atenção que me deu durante a elaboração desta dissertação.

- Aos professores Carlos Humes Jr., Maria Helena Cautiero Horta Jardim e Geci José Pereira da Silva pelas sugestões dadas.

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

ANÁLISE DE CONVERGÊNCIA E DE IMPLEMENTAÇÃO DE UM ALGORITMO PROXIMAL PARA O PROBLEMA DE COMPLEMENTARIDADE NÃO-LINEAR

Rodrigo Olimpio Costa

Outubro/2004

Orientador: Paulo Roberto Oliveira

Programa: Engenharia de Sistemas e Computação

Neste trabalho, apresentamos um algoritmo proximal inexato para o problema de complementaridade não-linear usando P_0-funções que converge globalmente e analisamos questões sobre a sua taxa de convergência, através de um método de regularização com métrica variável.

Sobre a resolução do problema regularizado, apresentamos um método tipo Newton e hipóteses adequadas que possibilitam a implementação do algoritmo proximal de forma a obtermos uma solução aproximada para o problema .

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

CONVERGENCE AND IMPLEMENTATION ANALYSIS OF A PROXIMAL ALGORITHM FOR THE NONLINEAR COMPLEMENTARITY PROBLEM

Rodrigo Olimpio Costa

October/2004

Advisor: Paulo Roberto Oliveira

Department: Computing and Systems Engineering

In this work, we present a inexact proximal algorithm for the nonlinear complementarity problem using P_0-functions that converges globally and analysed questions on the its rate of convergence, through regularization method with variable metrics.

On the resolution of the regularized problem, we present a Newton-type method and suitable hyphotesis that to make possible the implementation of the proximal algorithm to obtain a aproximate solution for the problem.

Índice

1 Introdução **1**

2 Preliminares **5**

2.1 Introdução 5

2.2 Notações 5

2.3 Conceitos básicos 6

2.4 Diferenciabilidade, subdiferenciabilidade e conceitos relacionados . . . 7

2.5 Coercividade de uma função 10

2.6 Propriedades P e P_0 10

3 Reformulações equivalentes para o $PCN(F)$ **13**

3.1 A função de Fischer-Burmeister 13

4 Regularização do $PCN(F)$ e o Algoritmo proximal **21**

4.1 O problema regularizado 21

4.2 O operador Fischer-Burmeister regularizado 22

4.3 O Algoritmo proximal Inexato I 23

5 Análise de convergência e Escolha de parâmetros **29**

5.1 Análise de Convergência 29

5.2 Análise do problema regularizado 32

5.2.1 Escolha dos parâmetros 32

5.2.2 Como resolver subproblemas 33

6 O Método de Newton para o problema regularizado **35**

6.1 Método de Newton para o $PCN(F^k)$ 35

6.2 Algoritmo do Ponto Proximal II 38

Conclusões 42

Referências Bibliográficas 44

Capítulo 1

Introdução

Consideremos o problema de complementaridade não-linear (PCN(F)), que constitui em encontrar um vetor $x \in \Re^n$ que satisfaça as seguintes condições:

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

onde $F : \Re^n \rightarrow \Re^n$ é assumida como sendo uma P_0-função continuamente diferenciável.

Observemos que F é uma P_0-função se para quaisquer $x, y \in \Re^n$ com $x \neq y$, existe um índice i tal que

$$x_i \neq y_i \qquad \text{e} \qquad (x_i - y_i)(F_i(x) - F_i(y)) \geq 0;$$

Este problema tem atraído a atenção de muitos pesquisadores e suas aplicações podem ser encontradas em diversas áreas como engenharia, economia, otimização entre outras. As referências [6] e [14] são as mais indicadas para o conhecimento destas aplicações.

Existem vários métodos para a resolução do problema de complementaridade não-linear. Os algoritmos a serem considerados nesta dissertação pertencem à chamada classe dos **métodos de regularização**, que são designados para resolver problemas mal postos. De fato, esses métodos têm sido recentemente usados com muito sucesso para melhorar a robustez de certos problemas de complementaridade, sob a avaliação de vários testes.

Para uma discussão detalhada sobre problemas mal postos em programação matemática, uma referência é [8].

Os métodos de regularização tentam contornar a dificuldade que encontramos na resolução de problemas robustos, substituindo a solução do problema original pela solução de uma sequência de problemas bem postos, isto é, uma sequência de problemas perturbados que possuem melhores condições de serem resolvidos.

No contexto de problemas de complementaridade, Facchinei e Kanzow [11] utilizaram a regularização de Tikhonov para a classe de P_0-funções, que é uma técnica que consiste em resolver uma sequência de problemas de complementaridade $PCN(F^k)$, onde $F^k(x) = F(x) + c_k x$ e c_k é um parâmetro positivo convergindo a zero.

Também para P_0-funções, Yamashita, Imai e Fukushima [30] usaram uma regularização proximal, que é dada por $F^k(x) = F(x) + c_k(x - x^k)$, onde x^k é o iterado e c_k é um parâmetro positivo que não necessariamente converge a zero. Este tipo de regularização foi bem estudado por Rockafellar [28] para operadores monótonos.

Nesses dois casos citados, o subproblema $PCN(F^k)$ é melhor tratável que o problema original. Isto se explica pelo fato de que se F é uma P_0-função, então F^k é uma P-função, isto é, para todo $x, y \in \Re^n$ com $x \neq y$, existe um índice i tal que

$$x_i \neq y_i \qquad (x_i - y_i)(F_i(x) - F_i(y)) > 0;$$

Esta propriedade garante que o subproblema tem no máximo uma solução.

Já Pereira [23] apresentou uma nova regularização proximal com métrica variável. Considerando o problema de complementaridade com P_0-funções, essa regularização se dá da seguinte forma: dado $x^k > 0$, temos

$$F^k(x) = F(x) + c_k(X^k)^{-r}(x - x^k)$$

onde $(X^k)^{-r}$ é definida por $(X^k)^{-r} = diag\{(x_1^k)^{-r}, ..., (x_n^k)^{-r}\}$ e c_k é uma parâmetro positivo.

Em uma linha similar, o recente trabalho de Oliveira, G.L e Oliveira, P.R [20] apresenta métodos do tipo proximal com métrica variável onde o núcleo quadrático é substituído pela métrica $G(x) = X^{-r}$.

Assim, usando essa regularização proximal, Pereira [23] apresentou um algoritmo inexato e provou que ele converge globalmente, sob hipóteses adequadas sobre os parâmetros $c_k{}'s$ e sobre o conjunto solução.

Nesta dissertação, veremos de que forma reformulamos um problema de complementariedade não-linear, utilizando a regularização de Pereira [23]. Para isso, faremos uso da função de Fischer-Burmeister, que possui propriedades fundamentais no estudo dos métodos de regularização. Apresentamos os principais conceitos e propriedades relacionados a esta função. Apresentamos também um algoritmo proximal inexato (Algoritmo 4.3.1) como o apresentado em [23] que converge globalmente para a solução do problema original $PCN(F)$, mas agora com um novo critério e com novas hipóteses sobre o parâmetro c_k e analisaremos questões sobre sua taxa convergência.

Será apresentado um resultado que mostra que o algoritmo proximal para P_0-funções possui taxa de convergência superlinear fazendo uso da regularização proximal de Yamashita, Imai e Fukushima [30]. A partir deste resultado, discutiremos se o mesmo pode ser aplicado à regularização com métrica variável de Pereira [23]. A análise da taxa de convergência do algoritmo será feita usando a hipótese de que o conjunto solução do $PCN(F)$ é não-vazio e limitado e que tem solução única. Veremos sob que condições essa unicidade acontece. Questões relacionadas à implementação do algoritmo também serão abordadas.

Com relação aos parâmetros $c_k{}'s$, apresentaremos uma forma de escolhê-los de maneira adequada às nossas hipóteses e de forma que a solução do subproblema $PCN(F^k)$ se aproxime da solução do problema $PCN(F)$.

Já sobre a resolução do problema regularizado $PCN(F^k)$, será apresentado um método tipo Newton de modo adaptado ao nosso problema e que tenha por objetivo minimizar a função de mérito regularizada. Lembramos que uma função $f : \Re^n \rightarrow \Re$

é uma função de mérito para o $PCN(F^k)$ se satisfaz a seguinte condição:

$$f(x) \geq 0 \quad \text{e} \quad f(x) = 0 \Longleftrightarrow x \text{ é solução do } PCN(F^k)$$

Métodos tipo Newton têm sido utilizados em vários campos da programação matemática, em particular em problemas de Complementaridade e de Desigualdades Variacionais.

A partir deste método, apresentamos uma versão específica do Algoritmo 4.3.1 que possibilita encontrar uma solução aproximada a do problema original $PCN(F)$.

A dissertação está organizada da seguinte maneira: No capítulo 2, daremos os conceitos básicos e algumas propriedades que serão necessárias em análises subseqüentes. No capítulo 3, veremos de que forma podemos reformular o problema de complementaridade de modo a torná-lo melhor condicionado, usando a função de Fischer-Burmeister. Veremos também importantes resultados relacionados a esta função. No capítulo 4, apresentamos o problema regularizado e um algoritmo proximal inexato para o mesmo. No capítulo 5 abordamos questões relacionadas à taxa de convergência e escolha de parâmetros. E finalmente, no capítulo 6, apresentamos o Método de Newton para a resolução do problema regularizado $PCN(F^k)$ e uma versão específica do algoritmo proximal inexato.

Capítulo 2

Preliminares

2.1 Introdução

Neste capítulo veremos alguns conceitos matemáticos e propriedades básicas que serão usadas nos capítulos subseqüentes. As referências [5], [6], [11], [24] e [30] foram essenciais para estes preliminares.

2.2 Notações

É importante apresentarmos algumas notações que serão usadas no decorrer desta dissertação. As principais são as seguintes:

$\Re^n = \{x = (x_1, x_2, ..., x_n); x_i \in \Re, i = 1, ..., n\}$: espaço vetorial euclideano;

$\Re^n_+ = \{x = (x_1, x_2, ..., x_n); x_i \geq 0, i = 1, ..., n\}$;

$\Re^n_{++} = \{x = (x_1, x_2, ..., x_n); x_i > 0, i = 1, ..., n\}$;

$\langle x, y \rangle = \sum_{k=1}^n x_i y_i =$ produto interno em $\Re^n$ para quaisquer vetores $x, y \in \Re^n$

$\|x\| = \sqrt{\langle x, x \rangle}$: norma euclideana em $\Re^n$,para algum vetor $x \in \Re^n$

$G'(ou \nabla G)$: matriz jacobiana de uma função $G : \Re^n \to \Re^n$

Se G é direcionalmente diferenciável, denotamos a derivada direcional de G em $x \in \Re^n$ por $G'(x; d)$.

2.3 Conceitos básicos

Primeiramente veremos algumas definições em relação a uma aplicação $G : \Re^n \to \Re^n$. G é dita ser uma

(a) função monótona se

$$\Big\langle x - y, G(x) - G(y) \Big\rangle \geq 0 \text{ para todo } x, y \in \Re^n,$$

(b) função fortemente monótona com módulo μ se

$$\Big\langle x - y, G(x) - G(y) \Big\rangle \geq \mu \|x - y\|^2 \text{ para todo } x, y \in \Re^n,$$

(c) P_0-função se para quaisquer $x, y \in \Re^n$, existe um índice i tal que

$$x_i \neq y_i \qquad (x_i - y_i)(G_i(x) - G_i(y)) \geq 0,$$

(d) P-função se para quaisquer $x, y \in \Re^n$, existe um índice i tal que

$$x_i \neq y_i \qquad (x_i - y_i)(G_i(x) - G_i(y)) > 0.$$

(e) função Lipschitz contínua se existe uma constante $L > 0$ tal que

$$\|G(x) - G(x')\| \leq L\|x - x'\|,$$

para todo x, $x' \in \Re^n$.

A partir desta definição, torna-se claro que uma função fortemente monótona é uma função monótona e uma P-função também é uma P_0-função. Além do mais, toda P-função é injetiva. Se G é uma função monótona ela também é uma P_0- função e se a mesma for uma função monótona diferenciável, então $\nabla G(x)$ é semidefinida positiva para todo $x \in \Re^n$.

2.4 Diferenciabilidade, subdiferenciabilidade e conceitos relacionados

Nesta seção, daremos as definições e algumas propriedades relacionadas à diferenciabilidade e à subdiferenciabilidade de funções que serão importantes na análise de propriedades que serão vistas mais à frente. Primeiramente veremos o conceito de suavidade.

Definição 2.4.1 . *Dizemos que uma função é suave em um ponto $x \in \Re^n$ se ela é continuamente diferenciável neste mesmo ponto. Caso contrário, dizemos que ela é não-suave.*

Definição 2.4.2 . *Podemos dizer que uma função $G : X \to Y$, X e Y espaços de dimensão finita (válida para espaços de Banach), admite uma derivada estrita em x, um elemento do espaço de aplicações lineares contínuas $L(X,Y)$ denotado $D_sG(x)$, se para cada $v \in X$, se tem:*

$$\lim_{x' \to x, t \downarrow 0} \frac{G(x' + tv) - G(x')}{t} = \left\langle D_sG(x), v \right\rangle$$

e a convergência é uniforme para v em conjuntos compactos.(Esta última condição é automática se G é Lipschitziana em uma vizinhança de x).

Agora veremos alguns conceitos de diferenciais generalizados.

Seja $G : \Re^n \to \Re^n$ uma função localmente Lipschitziana, então G é diferenciável em quase todo o $\Re^n$. Agora, seja $D_G \subseteq \Re^n$ o conjunto de pontos onde G é diferenciável. Então o B-subdiferencial de G em x é definido como

$$\partial_B G(x) = \{V \in \Re^{n \times n} | \exists \{x^k\} \subseteq D_G \text{ tq } lim_{x^k \to x} G'(x^k) = V\}$$

O subdiferencial de Clarke de G em x é definido por $\partial G(x) = co\partial_B G(x)$, onde *co* denota o envelope convexo de um conjunto.

Definição 2.4.3 *Seja $G : \Re^n \to \Re^n$. A derivada direcional de G em x na direção v, denotada por $G'(x, v)$, é definida como*

$$G'(x, v) = \lim_{t \downarrow 0} \frac{G(x + tv) - G(x)}{t},$$

quando este limite existe.

E a derivada direcional generalizada de G em x na direção v, denotada por $G^0(x, v)$, é definida como

$$G^0(x, v) = \limsup_{y \to x, t \downarrow 0} \frac{G(y + tv) - G(y)}{t},$$

onde $y \in \Re^n$ e t escalar positivo.

E mais, se $G'(x, v) = G^0(x, v)$, dizemos que G é regular em x.

Definição 2.4.4 *Seja $G : \Re^n \to \Re^n$ localmente Lipschitziana em $x \in \Re^n$. Dizemos que G é semi-suave em x se*

$$\lim_{d' \to d, t \downarrow 0} Hd'$$

existe para toda direção $d \in \Re^n$, com $H \in \partial G(x + td')$.

Definição 2.4.5 *Uma função semi-suave $G : \Re^n \to \Re^n$ é dita ser fortemente semi-suave em $x \in \Re^n$ se para qualquer $H \in \partial G(x + d)$ e para qualquer $d \to 0$,*

$$Hd - G'(x; d) = O(\|d\|^2)$$

Notemos que se G é semisuave em x, então é também direcionalmente diferenciável no mesmo e sua derivada direcional na direção v é dada pela definição (2.4.4).

As prova das duas proposições seguintes podem ser encontrada em [10].

Proposição 2.4.6 *Se $G : \Re^n \to \Re^n$ é semi-suave em x, então:*

$$\lim_{h \to 0} \frac{\|G(x + h) - G(x) - Hh\|}{\|h\|} = 0$$

onde $H \in \partial G(x + h)$.

Proposição 2.4.7 *Se $G : \Re^n \to \Re^n$ é fortemente semi-suave em $x \in \Re^n$ e direcionalmente diferenciável em uma vizinhança de x, então*

$$\limsup_{h \to 0} \frac{\|G(x+h) - G(x) - Hh\|}{\|h\|^2} < \infty$$

onde $H \in \partial G(x+h)$.

Vimos conceitos de B-diferenciabilidade e semi-suavidade. Entretanto não há regras de cálculo exato para a B-derivada de funções em geral. Em 1996, Qi [24] propôs o seguinte conceito de diferencial de importante papel no cálculo exato de diferenciais.

Definição 2.4.8 *Seja $G : \Re^n \to \Re^n$ uma função-vetor contínua e seja $T : \Re^n \to \Re^n \times \Re^n$ um operador ponto-conjunto. Então G é dita ser C-diferenciável em $x \in \Re^n$ com T e T é dito ser um operador C-diferencial de G se*

(i) $T(y)$ é não-vazio e compacto para qualquer y em uma vizinhança de x;

(ii) T é semicontínua superior em x, i.e, $T(x) \geq \limsup_k T(x_k)$ para toda sequência $\{x_k\} \subset \Re^n$ convergindo a x.

(iii) para qualquer $V \in T(x+d)$,

$$G(x+d) = G(x) + Vd + O(\|d\|).$$

Se, além do mais

(iv) para qualquer $V \in T(x+d)$,

$$G(x+d) = G(x) + Vd + O(\|d\|^2),$$

dizemos então que G é fortemente C-diferenciável em x com T e T é chamado um operador C-diferencial forte de G.

A demonstração da próxima proposição segue da Definição 2.4.5. Mas sua prova pode ser encontrada em [24, Teorema 2.1].

Proposição 2.4.9 *Sejam $F : \Re^n \to \Re^l$ e $G : \Re^l \to \Re^n$ subdiferenciáveis e fortemente semi-suaves. Então $H \equiv GoF : \Re^n \to \Re^n$ é fortemente C-diferenciável com o operador T definido por*

$$T(x) \equiv \{AB \in \Re^{n\times n} | A \in \partial G(F(x)), B \in \partial F(x)\},$$

isto é, existe uma constante positiva k tal que, para todo $d \in \Re^n$ suficientemente pequeno e qualquer $V \in T(x+d)$,

$$\|H(x+d) - H(x) - Vd\| \leq k\|d\|^2$$

Outras propriedades sobre semi-suavidade podem ser encontradas em [24].

Definição 2.4.10 *Seja $G : \Re^n \to \Re^n$ uma função localmente Lipschitziana. Então um ponto $x \in \Re^n$ é BD-regular com G se todos os elementos em $\partial_B G(x)$ são não-singulares.*

Notemos que x^* é uma solução isolada da equação $G(x) = 0$ se x^* é regular com G e $G(x^*) = 0$ como pode ser visto em [22, Proposição 3].

2.5 Coercividade de uma função

A coercividade tem um papel fundamental na prova da existência de solução do problema regularizado e também é muito importante na análise do método tipo Newton que será apresentado mais adiante.

Definição 2.5.1 *Dizemos que uma função $G : \Re^n \to \Re^n$ é coerciva se qualquer sequência $\{v^k\}$ com $\|v^k\| \to \infty$ implicar que $\|G(v^k)\| \to \infty$.*

2.6 Propriedades P e P_0

Definição 2.6.1 *. Uma matriz $M \in \Re^{n\times n}$ é uma*

(i) P_0-matriz se todos os seus menores principais são não-negativos

(ii) P-matriz se todos os seus menores principais são positivos.

É óbvio que toda P-matriz é uma P_0-matriz [6]. A seguinte caracterização sobre P_0-matrizes é bastante útil. Sua demonstração pode ser encontrada em [12].

Proposição 2.6.2 *A matriz $M \in \Re^{n\times n}$ é uma P_0-matriz se e somente se para todo vetor x não-nulo existe um índice i tal que $x_i \neq 0$ e $x_i(Mx)_i \geq 0$.*

Lembremos que uma função $G : \Re^n \to \Re^n$ é uma P_0-função se para quaisquer x, y, existe um índice i tal que

$$x_i \neq y_i \qquad (x_i - y_i)(G_i(x) - G_i(y)) \geq 0,$$

e uma P-função se, para quaisquer x, y , existe um índice i tal que

$$x_i \neq y_i \qquad (x_i - y_i)(G_i(x) - G_i(y)) > 0.$$

O lema a seguir é útil na prova da existência de solução do problema regularizado.

Lema 2.6.3 *Seja $G : \Re^n \to \Re^n$ uma P_0-função contínua. Consideremos uma sequência $\{a^k\}$ em $\Re^n$ tal que $\|a^k\| \to \infty$. Seja $j \in \{1, ...n\}$. Então existe uma subsequência, que podemos assumir sem perda de generalidade como $\{a_j^k\}$ e um índice j tal que uma das condições abaixo ocorre:*

(i) $a_j^k \to +\infty$ e $\{G_j(a^k)\}$ é limitada inferiormente;

(ii) $a_j^k \to -\infty$ e $\{G_j(a^k)\}$ é limitada superiormente.

Demonstração. Por hipótese $\{a^k\}$ é ilimitada. Desse modo, temos que o conjunto de índices J definido como

$$J := \{i \in \{1, ..., n\} | \{a_i^k\} \text{ é ilimitada}\},$$

é não-vazio. Assim, podemos supor sem perda de generalidade que $|a_j^k| \to \infty$ para todo $j \in J$. Definimos agora a sequência limitada $\{b^k\}$ por

$$b_i^k := \begin{cases} 0 & \text{se} \quad i \in J \\ a_i^k & \text{se} \quad i \notin J \end{cases}$$

Como $\{a^k\} \neq \{b^k\}$ e G é uma P_0-função, existe um $j \in J$ tal que

$$a_j^k[G_j(a^k) - G_j(b^k)] = (a_j^k - b_j^k)[G_j(a^k) - G_j(b^k)] \geq 0$$

Agora observamos que, quando $|a_j^k| \to \infty$, temos dois possíveis casos a serem analisados:

Caso 1. $a_j^k \to +\infty$

Temos da desigualdade anterior (considerando que estamos assumindo que $a_j^k > 0$) que $G_j(a^k)$ é limitada inferiormente, já que G_j é contínua e $\{b^k\}$ é limitada. Desse modo, provamos a parte **(i)**.

Caso 2. $a_j^k \to -\infty$

Temos da desigualdade anterior (considerando que estamos assumindo que $a_j^k < 0$) que $G_j(a^k)$ é limitada superiormente. Desse modo, provamos a parte **(ii)**. ■

Já a proposição seguinte será utilizada na demonstração da unicidade do problema regularizado.

Proposição 2.6.4 *Seja $G : \Re^n \to \Re^n$ uma P-função. Então o $PCN(G)$ tem no máximo uma solução.*

Demonstração. Visto que G é uma P-função, temos que ela é injetiva. Logo se o $PCN(G)$ tem solução, ela é única. ■

Capítulo 3

Reformulações equivalentes para o $PCN(F)$

Neste capítulo apresentaremos uma reformulação do problema de complementaridade não-linear equivalente ao problema original.

3.1 A função de Fischer-Burmeister

Definição 3.1.1 *Uma função $\varphi : \Re^2 \longrightarrow \Re$ é dita ser uma função-PCN se satisfaz a seguinte condição:*

$$\varphi(a,b) = 0 \Longleftrightarrow a \geq 0, b \geq 0 \text{ e } ab = 0.$$

O problema de complementaridade não-linear pode ser reformulado como um sistema de equações por diversos caminhos. Nesta dissertação, faremos uso da função de Fischer-Burmeister que é uma função $\psi : \Re^2 \longrightarrow \Re$ definida por

$$\psi(a,b) = \sqrt{a^2 + b^2} - a - b$$

Esta função tem atraído muita atenção na área da Complementaridade não-linear e das Desigualdades variacionais. Ela foi usada por Facchinei e Kanzow [11] no estudo de métodos de regularização, depois de ser introduzida por Fischer em 1992. O resultado a seguir deixa claro a importância desta função.

Lema 3.1.2 *A função de Fischer-Burmeister é uma função-PCN.*

Demonstração. Suponhamos que $\psi(a,b) = 0$. Então temos que

$$0 \leq \sqrt{a^2 + b^2} = a + b$$

Assim, elevando ambos os lados ao quadrado, temos que

$$(a+b)^2 = a^2 + b^2 \Longrightarrow ab = 0$$

A prova segue usando o fato que $a + b \geq 0$. Reciprocamente, suponhamos que $a \geq 0, b \geq 0$ e $ab = 0$. Usando o fato que $ab = 0$, temos que

$$\sqrt{a^2 + b^2} = |a + b|$$

Assim, como $a \geq 0$ e $b \geq 0$ temos que $\psi(a,b) = \sqrt{a^2 + b^2} - a - b = 0$. ∎

Notemos que ψ é continuamente diferenciável em todo o espaço, exceto na origem. Com isso, podemos definir o seguinte operador:

$$\Theta(x) = \begin{bmatrix} \psi(x_1, F_1(x)) \\ \psi(x_2, F_2(x)) \\ \vdots \\ \psi(x_n, F_n(x)) \end{bmatrix}$$

com $F : \Re^n \to \Re^n$.

O próximo lema apresenta uma reformulação vista do problema de complementaridade não-linear que se constitui em encontrar um vetor $x \in \Re^n$ que satisfaça as seguintes condições:

$$x \geq 0 \qquad F(x) \geq 0 \qquad x^T F(x) = 0$$

onde $F : \Re^n \to \Re^n$ é assumida como sendo uma P_0-função continuamente diferenciável.

Lema 3.1.3 *Seja F uma função de $\Re^n$ em $\Re^n$. Então $x^* \in \Re^n$ é solução do $PCN(F)$ se, e somente se, x^* é solução do sistema de equações $\Theta(x) = 0$.*

Demonstração. Se $x^* \in \Re^n$ é uma solução do $PCN(F)$, então para todo i=1,...n temos que

$$x_i^* \geq 0 \qquad F_i(x^*) \geq 0 \qquad x_i^* F_i(x^*) = 0$$

Portanto,

$$(x_i^* + F_i(x^*))^2 = (x_i^*)^2 + (F_i(x^*))^2$$

implicando que

$$\psi(x_i, F_i(x)) = 0$$

para todo i=1,...n.

Reciprocamente, suponha que $\Theta(x) = 0$. Então

$$\psi(x_i, F_i(x)) = 0$$

para todo i=1,...n. Desse modo, segue da definição 3.1.1 que

$$x_i^* \geq 0 \qquad F_i(x^*) \geq 0 \qquad x_i^* F_i(x^*) = 0$$

para todo i=1,...n. Logo x^* é solução do $PCN(F)$. ■

Percebemos que Θ não é diferenciável em x se $x_i = F(x_i) = 0$ para algum índice i, mas é localmente Lipschitz contínua. Com este operador, podemos definir a seguinte função de mérito $\phi(x) : \Re^n \rightarrow \Re$ correspondente:

$$\phi(x) = \frac{1}{2}\|\Theta(x)\|^2. \tag{3.1}$$

Veremos a seguir alguns resultados relacionados ao subdiferencial de uma função e em seguida alguns conceitos relacionados ao cálculo do subdiferencial de Θ. Lembrando que estamos assumindo que F é uma função continuamente diferenciável e que isso implica que Θ é localmente Lipschitziana. As demonstrações dos cinco próximos resultados podem ser encontradas em [2].

Proposição 3.1.4 *Se f é estritamente diferenciável em x, então f é Lipschitziana em uma vizinhança de x e $\partial f(x) = \{D_s f(x)\}$. Reciprocamente, se f é Lipschitziana em uma vizinhança de x e $\partial f(x)$ se reduz a um vetor único ξ, então f é estritamente diferenciável em x e $D_s f(x) = \xi$.*

Corolário 3.1.5 *Se f é Lipschitziana em uma vizinhança de x e X é um espaço de dimensão finita, então $\partial f(x')$ se reduz a um único elemento para todo x' em $x + \varepsilon B$ se e somente se f é continuamente diferenciável sobre $x + \varepsilon B$, onde B denota a bola unitária aberta de X.*

Teorema 3.1.6 (Regra da cadeia) *Seja $f = goh$, onde $h : X \to \Re^n$ e $g : \Re^n \to \Re$ (X espaço de dimensão finita). Assumimos que cada h_i é Lipschitziana em uma vizinhança de x e que g é Lipschitziana em uma vizinhança de $h(x)$, que implica que f é Lipschitziana em uma vizinhança de x. Então*

$$\partial f(x) \subset co\Big\{ \sum \alpha_i \xi_i : \xi_i \in \partial h_i(x), \alpha \in \partial g(h(x)) \Big\}$$

e a igualdade está assegurada sob qualquer uma das seguintes hipóteses:

(i) g é regular em $h(x)$, cada h_i é regular em x, e todo elemento α de $\partial g(h(x))$ tem componentes não negativos.(Neste caso segue que f é regular em x.)

(ii) g é estritamente diferenciável em $h(x)$ e $n = 1$.

(iii) g é regular em $h(x)$ e h é estritamente diferenciável em x(Neste caso, segue que f é regular em x e co na expressão acima é supérfluo).

Proposição 3.1.7 *Seja $G : \Re^n \to \Re^n$ Lipschitziana em uma vizinhança de x. Então:*

(a) $\partial G(x)$ é um subconjunto compacto, convexo e não-vazio de $\Re^{n \times n}$.

(b) ∂G é fechado em x, isto é, se $x_i \to x$, $Z_i \in \partial G(x_i)$ e $Z_i \to Z$, então $Z \in \partial G(x)$.

(c) ∂G é semicontínua superior em x, isto é, para qualquer $\varepsilon > 0$ existe um $\delta > 0$ tal que, para todo y em $x + \delta B$ (onde B denota a bola aberta unitária em $\Re^n$) ,

$$\partial G(y) \subset \partial G(x) + \varepsilon B_{n \times n}.$$

(d) Se cada componente g^i de G é Lipschitziana de posto K_i em x, então G é Lipschitziana em x de posto $K = |(K_1, K_2, ..., K_m)|$ e $\partial G \subset K\overline{B}_{n\times n}$.

(e) $\partial G(x) \subset \partial g^1(x) \times ... \times \partial g^n(x)$, onde o último(posterior) denota o conjunto de todas as matrizes cuja i-ésima linha pertence a $\partial g^i(x)$ para cada i. Se $n = 1$, então $\partial G(x) = \partial g^1(x)$.

Teorema 3.1.8 *Seja $g = hoG$, onde $G : \Re^n \longrightarrow \Re^n$ é Lipschitziana em uma vizinhança de x e $h : \Re^n \longrightarrow \Re$ é Lipschitziana em uma vizinhança de $G(x)$. Então g é Lipschitziana em uma vizinhaça de x e*

$$\partial g \subset co\{\partial h(G(x))\partial G(x)\}.$$

Se h for também estritamente diferenciável em x, então a igualdade está assegurada.

Agora, veremos um resultado relacionado ao cálculo do subdiferencial de Θ.

Proposição 3.1.9 *$\partial\Theta(x)^T \subseteq (A(x) - I) + \nabla F(x)(B(x) - I)$ onde I é a matriz identidade $n \times n$ e $A(x)$ e $B(x)$ são matrizes diagonais $n \times n$ (multivalorada) cujo i-ésimo elemento diagonal é dado por*

$$A_{ii}(x) = \frac{x_i}{\|(x_i, F_i(x))\|}, B_{ii}(x) = \frac{F_i(x)}{\|(x_i, F_i(x))\|}$$

se $(x_i, F_i(x)) \neq 0$ e por

$$A_{ii}(x) = \xi_{ii}, B_{ii}(x) = \rho_{ii},$$

para todo $(\xi_i, \rho_i) \in \Re^2$ tq $\|(\xi, \rho_i)\| \leq 1$
se $(x_i, F_i(x)) = 0$.

. **Demonstração.** Das regras de cálculo do jacobiano generalizado temos da Proposição 3.1.7 que

$$\partial\Theta(x)^T \subseteq \partial\Theta_1(x) \times ...\partial\Theta_n(x).$$

Se i é tal que $(x_i, F_i(x)) \neq 0$, $\Theta_i(x)$ é diferenciável. Como $\psi(x_i, F_i(x)) = \sqrt{x_i^2 + F_i^2(x)} - x_i - F_i(x)$ temos então que

$$
\begin{aligned}
\nabla\Theta_i(x) &= \nabla\psi(x_i, F_i(x)) \\
&= \frac{1}{2\sqrt{x_i^2 + F_i^2(x)}}(2x_i e_i + 2F_i(x)\nabla F_i(x)) - e_i - \nabla F_i(x) \\
&= \frac{x_i e_i}{\sqrt{x_i^2 + F_i^2(x)}} + \frac{F_i(x)\nabla F_i(x)}{\sqrt{x_i^2 + F_i^2(x)}} - e_i - \nabla F_i(x) \\
&= \Big(\frac{x_i e_i}{\sqrt{x_i^2 + F_i^2(x)}} - e_i\Big) + \Big(\frac{F_i(x)\nabla F_i(x)}{\sqrt{x_i^2 + F_i^2(x)}} - \nabla F_i(x)\Big) \\
&= \Big(\frac{x_i}{\|x_i, F_i(x)\|} - 1\Big)e_i + \nabla F_i(x)\Big(\frac{F_i(x)}{\|x_i, F_i(x)\|} - 1\Big).
\end{aligned}
$$

Assim,

$$
\nabla\Theta_i(x) = (\frac{x_i}{\|x_i, F_i(x)\|} - 1)e_i + \nabla F_i(x)(\frac{F_i(x)}{\|x_i, F_i(x)\|} - 1).
$$

Agora, analisaremos o caso em que $(x_i, F_i(x)) = 0$. Por definição

$$
\partial\|0, 0\| = \{(\xi_i, \rho_i) : \|\xi_i, \rho_i\| \leq 1\}.
$$

Usando a Definição 2.4.3, temos que ψ é regular em $F(x)$, i.e, ψ é direcionalmente diferenciável em $F(x)$, pois ψ é semi-suave em F já que F é continuamente diferenciável (Veja Definição 2.4.4). Sabemos também que F é estritamente diferenciável em x. Como Θ é uma composição de ψ e F, temos pelo item (iii) do Teorema 3.1.6 (que é o Teorema sobre gradiente generalizado de uma função composta) que Θ é regular em x e seu subdiferencial é dado da seguinte forma :

$$
\partial\Theta_i(x) = \partial\psi(x_i, F_i(x)) = (\xi_i - 1)e_i + \nabla F_i(x)(\rho_i - 1)
$$

Destas igualdades, segue diretamente a proposição. ■

A proposição a seguir nos dá algumas propriedades úteis para as funções vistas.

Proposição 3.1.10 *Os seguintes enunciados são verdadeiros :*

(a) A função Θ é semi-suave.

(b) Se ∇F é Lipschitz contínua, então Θ é fortemente semi-suave em todo o espaço $\Re^n$.

(c) A função de mérito ϕ é continuamente diferenciável em todo o espaço $\Re^n$(F diferenciável).

(d) Se F é uma P_0-função, então todo ponto estacionário de ϕ é tal que $\phi(x) = 0$.

Demonstração. Veja [30, Proposição 2.2].

A demonstração do próximo lema pode ser encontrada em [18].

Lema 3.1.11 *Sejam $\{a^k\}$ e $\{b^k\}$ duas sequências em $\Re$. Se uma das condições ocorrer*

(i) $a^k \to +\infty$ e $b^k \to +\infty$

(ii) $a^k \to -\infty$

(iii) $b^k \to -\infty$

Então $\|\psi(a^k, b^k)\| \to +\infty$.

Proposição 3.1.12 *Seja $F : \Re^n \to \Re^n$ uma P-função(P_0-função). Então Θ é uma P-função(P_0-função).*

Demonstração. Primeiramente, assumimos que F é uma P-função. Assim, para $x \neq y$, existe um índice i tal que

$$(x_i - y_i)[F_i(x) - F_i(y)] > 0.$$

Queremos mostrar que para o mesmo índice i,

$$(x_i - y_i)[\Theta_i(x) - \Theta_i(y)] > 0.$$

Assumimos, sem perda de generalidade, que $y_i > x_i$ e mostraremos que $\Theta_i(y) > \Theta_i(x)$. Por contradição, suponha que $\Theta_i(y) \leq \Theta_i(x)$. Temos que $y_i > x_i$ e $F_i(y) > F_i(x)$. Agora, $\Theta_i(y) \leq \Theta_i(x)$ implica que

$$(x_i - y_i) + (F_i(x) - F_i(y)) \geq \sqrt{x_i^2 + F_i(x)^2} - \sqrt{y_i^2 + F_i(y)^2}.$$

Assim, elevando ambos os lados ao quadrado, simplificando, e usando a seguinte desigualdade

$$(x_i - y_i)(F_i(x) - F_i(y)) > 0,$$

temos que

$$\sqrt{x_i^2 + F_i(x)^2}\sqrt{y_i^2 + F_i(y)^2} < x_i y_i + F_i(x)F_i(y).$$

Se elevarmos ao quadrado a desigualdade acima, segue que

$$(x_i F_i(y) - y_i F_i(x))^2 < 0.$$

E isto é uma contradição. Desse modo, temos que Θ é uma P-função. A propriedade P_0 pode ser provada de forma análoga. ∎

Notemos que essa proposição depende fortemente da função de Fischer-Burmeister. Logo seria importante tentar usar uma outra função-PCN para a qual esse resultado também seje válido.

Capítulo 4

Regularização do $PCN(F)$ e o Algoritmo proximal

Neste capítulo, será apresentado o problema de complementaridade com uma regularização para a função F e introduziremos um algoritmo usando essa regularização.

4.1 O problema regularizado

Para regularizarmos o problema de complementaridade não-linear, usaremos a regularização proximal introduzida por Pereira [23]. Assim, dada uma função $F : \Re^n \to \Re^n, x^k \in \Re^n, x^k > 0, r \in \Re, r \geq 1$, temos a seguinte regularização $F^k : \Re^n \to \Re^n$

$$F^k(x) = F(x) + c_k(X^k)^{-r}(x - x^k) \tag{4.1}$$

onde $(X^k)^{-r}$ é definida por $(X^k)^{-r} = diag\{(x_1^k)^{-r}, ..., (x_n^k)^{-r}\}$ e c_k é uma parâmetro positivo.

Assim, o problema regularizado $PCN(F^k)$ é definido como

$$x \geq 0 \qquad F^k(x) \geq 0 \qquad x^T F^k(x) = 0$$

onde $F^k : \Re^n \to \Re^n$ é uma P-função, como será demonstrado. Para resolvermos este problema utilizaremos um método tipo Newton que será visto mais adiante.

4.2 O operador Fischer-Burmeister regularizado

Agora, para o problema regularizado, definimos o operador de Fischer-Burmeister correspondente como

$$\Theta^k(x) = \begin{bmatrix} \psi(x_1, F_1^k(x)) \\ \psi(x_2, F_2^k(x)) \\ \vdots \\ \psi(x_n, F_n^k(x)) \end{bmatrix}$$

e sua função de mérito associada como

$$\phi^k(x)=\frac{1}{2}||\Theta^k(x)||^2$$

onde F^k é a função regularizada e ψ é a função de Fischer-Burmeister.

A proposição a seguir mostra que F^k é uma regularização de F.

Proposição 4.2.1 *Se F é uma P_0-função, então F^k é uma P-função.*

Demonstração. Como F é uma P_0-função, para quaisquer $x \neq y \in \Re^n$, existe um índice i tal que

$$(x_i - y_i)(F_i(x) - F_i(y)) \geq 0 \tag{4.2}$$

O objetivo é mostrarmos que

$$(x_i - y_i)(F_i^k(x) - F_i^k(y)) > 0.$$

Usando a definição de F^k temos que

$$(x_i - y_i)(F_i^k(x) - F_i^k(y)) \quad = \quad (x_i - y_i)(F_i(x) - F_i(y)) + c_k(x_i^k)^{-r}(x_i - y_i)^2$$

Agora, usando (4.2) e os fatos que $x_i \neq y_i, x_i^k > 0$ e que $c_k > 0$ temos que

$$(x_i - y_i)(F_i^k(x) - F_i(y)^k) > 0$$

■

A próxima proposição nos dá propriedades úteis para o problema regularizado.

Proposição 4.2.2 *As seguintes afirmativas são verdadeiras:*

(a) O operador regularizado Θ^k é continuamente diferenciável sobre $\Re^n_{++} \times \Re^n$.

(b) Os conjuntos de nível de ϕ^k são limitados.

Demonstração. (a) Imediato, visto que ψ é continumente diferenciável sobre $\Re^n_{++} \times \Re^n$.

(b) Evidente, levando-se em consideração a Proposição 5.2.1 que mostra que ϕ^k é coerciva. ■

4.3 O Algoritmo proximal Inexato I

Nesta seção será exibido um algoritmo do tipo proximal inexato que resolve o problema de complementaridade não-linear, onde $F : \Re^n \to \Re^n$ é uma P_0-função continuamente diferenciável. Também daremos hipóteses suficientes para a convergência global do algoritmo. Sabendo que $F^k(x) = F(x) + c_k(X^k)^{-r}(x - x^k)$,

$$\Theta^k(x) = \begin{bmatrix} \psi(x_1, F_1^k(x)) \\ \psi(x_2, F_2^k(x)) \\ \vdots \\ \psi(x_n, F_n^k(x)) \end{bmatrix},$$

e

$$\phi^k(x)=\frac{1}{2}\|\Theta^k(x)\|^2,$$

apresentamos o seguinte algoritmo:

Algoritmo 4.3.1 Algoritmo proximal Inexato

Passo 0: *Escolha* $c_0 > 0, \delta_0 \in (0,1)$ *e* $x^0 \in \Re^n_{++}$. *Faça* $k := 0$.

Passo 1: *Se* x^k *satisfaz um critério de parada termine a iteração.*

Passo 2: *Dados* $c_k > 0$, $F^k(x) = F(x) + c_k(X^k)^{-r}(x - x^k)$, $r \geq 1$, $\delta_k \in (0,1)$ *e* $x^k \in \Re^n_{++}$, *obtenha* $x^{k+1} \in \Re^n_{++}$ *de forma que*

$$\phi^k(x^{k+1})^{\frac{1}{2}} \leq \delta_k.min\{1, \|x^{k+1} - x^k\|\} \tag{4.3}$$

Passo 3: *Escolha* $c_{k+1} > 0, \delta_{k+1} \in (0, \delta_k)$. *Faça* $k := k+1$ *e volte ao Passo 1.*

Para obtermos o iterado x^{k+1} satisfazendo o critério (4.3) no Passo 2, apresentaremos mais adiante o Método de Newton para a minimização de ϕ^k. Observe que, os conjuntos de nível de ϕ^k são limitados (Proposição 4.2.2) e como F^k é uma P-função (P_0-função), todo ponto estacionário $\overline{x}$ de ϕ^k é tal que $\phi^k(\overline{x}) = 0$ (Proposição 3.1.10 (d)). Por isso, todo algoritmo de minimização adequado ao problema produzirá uma sequência minimizante e assim o ponto x^k poderá ser determinado em um número finito de passos. Esta situação reflete o fato que os problemas perturbados são bem-postos e isto é uma das principais motivações dos métodos de regularização.

Rockafellar [28] propôs os seguintes dois critérios em 1976 para um algoritmo proximal usando operadores monótonos maximais :

Critério 1. $\|x^{k+1} - \overline{x}^k\| \leq \delta_k$, $\sum_{k=0}^{\infty} \delta_k < \infty$

Critério 2. $\|x^{k+1} - \overline{x}^k\| \leq \delta_k \|x^{k+1} - x^k\|$, $\sum_{k=0}^{\infty} \delta_k < \infty$

onde $\overline{x}^k$ é a solução do problema.

O Critério 1 garante convergência global do algoritmo proximal, enquanto o Critério 2 garante convergência superlinear.

O seguinte critério

$$\|x^{k+1} - \overline{x}^k\| \leq \delta_k min\{1, \|x^{k+1} - x^k\|\}$$

com $\delta_k \to 0$ foi utilizado por Rockafellar para provar que o algoritmo do ponto proximal para operadores monótonos maximais proposto por ele tem convergência global e superlinear, ou seja, mesclando os critérios 1 e 2.

Em 1999, Yamashita, Imai e Fukushima [30], usando este mesmo critério, apresentaram um algoritmo proximal com P_0-funções para o $PCN(F)$ utilizando a regularização $F^k(x) = F(x) + c_k(x - x^k)$ ($x^k \in \Re^n$) e provaram que o algoritmo converge globalmente e que possui taxa de convergência superlinear usando a seguinte hipótese sobre a sequência $\{c_k\}$:

Hipótese 4.3.2 . *A sequência $\{c_k\}$ satisfaz as seguintes condições:*

(A) $c_k(x^{k+1} - x^k) \to 0$ *se* $\{x^k\}$ *é limitada ;*

(B) $c_k x^k \to 0$ *se* $\{x^k\}$ *é ilimitada .*

Agora, enfatizaremos a seguinte hipótese sobre o parâmetro de regularização c_k, hipótese esta que tem importância fundamental na análise de convergência global do Algoritmo 4.3.1.

Hipótese 4.3.3 . *A sequência $\{c_k\}$ satisfaz as seguintes condições:*

(A) $c_k(X^k)^{-r}(x^{k+1} - x^k) \to 0$ *se* $\{x^k\}$ *é limitada e* $r \geq 1$*;*

(B) $c_k(X^k)^{-r} \to 0$ *se* $\{x^k\}$ *é ilimitada e* $r \geq 0$.

Observamos nessa hipótese que c_k não necessariamente necessita ir a zero para ela ser válida, como no caso proximal clássico estudado por Yamashita, Imai e Fukushima [30]. Pereira [23] mostrou, sob esta hipótese e considerando o conjunto solução do problema não-vazio e limitado, que x^k é limitada e que ϕ^k converge uniformemente a ϕ em compactos e a partir disso provou a convergência global do Algoritmo 4.3.1 utilizando o critério $\phi^k(x^{k+1})^{\frac{1}{2}} \leq \delta_k$ no Passo 2.

Porém, considerando além desta, outras hipóteses sobre c_k, analisaremos no próximo capítulo a possibilidade de o algoritmo ter taxa de convergência superlinear, utilizando o critério de Rockafellar adaptado ao nosso problema (que é o critério do Passo 2 do Algoritmo 4.3.1). Este critério engloba o critério usado por Pereira [23], pois δ_k é uma parâmetro convergindo a zero. Também veremos questões de implementação e uma versão específica deste algoritmo.

Agora veremos resultados importantes para a convergência global do Algoritmo 4.3.1.

Lema 4.3.4 *Para quaisquer a, b e c, temos*

$$|\psi(a, b+c) - \psi(a, b)| \leq 2|c|$$

Como consequência, temos que

$$\|\Theta^k(x) - \Theta(x)\| \leq 2c_k \|(X^k)^{-r}(x - x^k)\|$$

Demonstração. Visto que

$$\psi(a, b+c) - \psi(a, b) = \sqrt{a^2 + (b+c)^2} - \sqrt{a^2 + b^2} - c,$$

é suficiente mostrarmos que

$$|\sqrt{a^2 + (b+c)^2} - \sqrt{a^2 + b^2}| \leq |c|$$

Assim temos

$$\begin{aligned}
|\psi(a, b+c) - \psi(a, b)| &= |\sqrt{a^2 + (b+c)^2} - \sqrt{a^2 + b^2} - c| \\
&\leq |\sqrt{a^2 + (b+c)^2} - \sqrt{a^2 + b^2}| + |c| \\
&= |\|(a, b+c)\| - \|(a, b)\|| + |c| \\
&\leq \|(0, c)\| + |c| = 2|c|
\end{aligned}$$

Portanto, se tomarmos $a = x_i, b = F_i(x)$ e $c = c_k(x_i^k)^{-r}(x_i - x_i^k)$ temos

$$|\psi(x_i, F_i(x) + c_k(x_i^k)^{-r}(x_i - x_i^k)) - \psi(x_i, F_i(x))| \leq 2|c_k(x_i^k)^{-r}(x_i - x_i^k)|$$

Usando as definições dos operadores Θ^k e Θ, concluímos que

$$\|\Theta^k(x) - \Theta(x)\| \leq 2c_k\|(X^k)^{-r}(x - x^k)\|$$

como queríamos demonstrar. ∎

Teorema 4.3.5 *(Teorema do Passo da Montanha) Seja $f : \Re^n \to \Re$ uma função continuamente diferenciável e coerciva. Seja $S \subset \Re^n$ um conjunto compacto e não-vazio e m o mínimo de f na fronteira (compacta) de S, isto é, $m := \min_{x\in\partial S} f(x)$. Se existe $a \in S$ e $b \notin S$ tal que $f(a) < m$ e $f(b) < m$, então existe $c \in \Re^n$ tal que $\nabla f(c) = 0$ e $f(c) \geq m$.*

Prova. Veja [30]. ∎

Lema 4.3.6 *Suponha que vale o item* **B** *da Hipótese 4.3.3 e que $S \subset \Re^n$ é um conjunto compacto não-vazio. Se $\{x^k\}$ é ilimitada, então para cada $\varepsilon > 0$, existe um k_0 suficientemente grande tal que para tod $k \geq k_0$ temos que*

$$|\phi^k(x) - \phi(x)| \leq \varepsilon$$

para todo $x \in S$.

Demonstração. Pelo Lema anterior, tomando $a = x_i, b = F_i(x)$ e $c = c_k(x_i^k)^{-r}(x_i - x_i^k)$, segue pelo item **B** da Hipótese 4.3.3 que

$$|\psi(x_i, F_i^k(x) - \psi(x_i, F_i(x))| \leq 2c_k|(x_i^k)^{-r}(x_i - x_i^k)| \to 0$$

Como S é compacto e ψ, FeF^k são contínuas, temos para todo i, que

$$|\psi^2(x_i, F_i^k(x) - \psi^2(x_i, F_i(x))| =$$

$$|\psi(x_i, F_i^k(x) - \psi(x_i, F_i(x))||\psi(x_i, F_i^k(x)) + \psi(x_i, F_i(x))| \to 0$$

Assim, segue que ϕ^k converge uniformemente a ϕ em S. ∎

Agora, veremos o teorema que mostra que o algoritmo 4.3.1 converge globalmente.

Teorema 4.3.7 *Suponha que F é uma P_0-função continuamente diferenciável, vale a Hipótese 4.3.3 items* **A** *e* **B** *e que o conjunto de soluções, S^*, do $PCN(F)$ é não-vazio e limitado. Se $\delta_k \to 0$, então $\{x^k\}$ é limitada e qualquer ponto de acumulação de $\{x^k\}$ é uma solução do $PCN(F)$.*

Prova. Primeiro mostraremos que $\{x^k\}$ é limitada. Suponha que $\{x^k\}$ não é limitada, então existe uma subsequência $\{x^k\}_{k\in K}$ tal que $\|x^k\| \to \infty$ quando $k \to \infty$ com $k \in K$.

Como S^* é limitado, existe um compacto $S \subset \Re^n$ não-vazio tal que $S^* \subset int(S)$ e $x^k \notin S$ para $k \in K$, suficientemente grande. Mas, se $x^* \in S^*$, então temos que $\phi(x^*) = 0$. Assim, tomando $\alpha := \min_{x\in\partial S} \phi(x) > 0$ (para x na fronteira de S), e aplicando o Lema 4.3.6 com $\varepsilon := \frac{\alpha}{4}$ existe um k_0 tal que para todo $k \geq k_0$

$$\phi^k(x^*) \leq \frac{\alpha}{4} \quad e \quad m := \min_{x\in\partial S} \phi^k(x)(\text{para } x \text{ na fronteira de S}) \geq \frac{3\alpha}{4}$$

Do Passo 2 do algoritmo proximal 4.3.1 temos que $\phi^k(x^{k+1}) \leq \delta_k^2$, portanto, usando a hipótese que $\delta_k \to 0$, existe um k_1 tal que para todo $k \geq k_1$, $\phi^k(x^{k+1}) \leq \frac{\alpha}{4}$.

Agora, considerando um índice $k \in K$, fixo suficientemente grande, tal que $k \geq max\{k_0, k_1\}$ e, colocando $a = x^*$ e $b := x^{k+1}$ temos, pelo Teorema do Passo da Montanha que existe um $c \in \Re^n$ tal que $\nabla\phi^k(c) = 0$ e $\phi^k(c) \geq m \geq \frac{3\alpha}{4} > 0$. Assim, c é um ponto estacionário de ϕ^k, mas não é solução do $PCN(F^k)$ contradizendo a Proposição 3.1.10 (d). Portanto x^k é limitada.

Agora, mostraremos que qualquer ponto de acumulação de $\{x^k\}$ está em S^*. Como $\{x^k\}$ é limitada, temos da Hipótese 4.3.3 item A que $\|F^k(x^{k+1}) - F(x^{k+1})\| \to 0$, e de forma análoga à prova do Lema 4.3.6 temos que $\|\phi^k(x^{k+1}) - \phi(x^{k+1})\| \to 0$. Assim pelo Passo 2 do algoritmo e a hipótese que $\delta_k \to 0$, temos que $\phi^k(x^{k+1}) \to 0$. Portanto, $\phi(x^{k+1}) \to 0$, o que significa que todo ponto de acumulação da sequência é uma solução do $PCN(F)$. ∎

Capítulo 5

Análise de convergência e Escolha de parâmetros

Neste capítulo, apresentaremos o teorema de convergência superlinear usado por Yamashita, Imai e Fukushima [30] para mostrar que o algoritmo proximal apresentado por eles possui taxa de convergência superlinear, isto é, considerando a função regularizada $F^k = F(x) + c_k(x - x^k)$ ($x^k \in \Re^n$). A partir deste teorema, discutiremos esse resultado com relação ao Algoritmo 4.3.1 e analisaremos questões relacionadas à sua implementação.

5.1 Análise de Convergência

Agora, veremos uma hipótese e um lema a serem utilizados na análise da taxa de convergência .

Hipótese 5.1.1 *Suponha que o conjunto de soluções do $PCN(F)$, S^*, é não-vazio e que a sequência x^k converge para uma solução x^* do $PCN(F)$ que é BD-regular com Θ (Veja definição 2.4.10).*

Considerando F como uma P_0-função, o conjunto solução do $PCN(F)$ é conexo sempre que ele é limitado [9, Corolário 3.5]. Neste caso, a hipótese acima garante que x^* é a única solução do problema.

Lema 5.1.2 *Suponha que a Hipótese 5.1.1 está assegurada. Então existem constantes positivas K e ε tais que*

$$\|x - x^*\| \leq K\|\Theta(x)\|$$

para todo x tal que $\|\Theta(x)\| < \varepsilon$.

Demonstração. Veja [22, Proposição 3].

O teorema abaixo ([30], Teorema 3.3) foi provado para um algoritmo proximal similar ao Algoritmo 4.3.1, com o mesmo critério do Passo 2, mas usando a regularização $F^k(x) = F(x) + c_k(x - x^k)$, com $x^k \in \Re^n$. Considerando que o algoritmo converge globalmente, mostramos abaixo a prova do teorema.

Teorema 5.1.3 *Suponha que a Hipótese 5.1.1 está assegurada e que o conjunto solução do $PCN(F)$ é limitado. Suponha também que $\delta_k \to 0$ e $c_k \to 0$. Então a sequência x^k gerada pelo algoritmo converge superlinearmente para a solução x^* do $PCN(F)$.*

Prova. Como foi mencionado logo após a Hipótese 5.1.1, temos que, sob a mesma, x^* é a única solução do $PCN(F)$, e assim x^k converge para x^*, já que o algoritmo converge globalmente. Além do mais, usando a função regularizada $F^k(x) = F(x) + c_k(x - x^k)$, temos pelo Lema 4.3.4 que

$$\|\Theta^k(x^{k+1}) - \Theta(x^{k+1})\| \leq 2c_k\|x^{k+1} - x^k\|.$$

Então segue da desiguladade triangular que

$$\|\Theta(x^{k+1})\| \leq \|\Theta^k(x^{k+1})\| + 2c_k\|x^{k+1} - x^k\|.$$

Esta desigualdade, usando a definição da função de mérito ϕ (3.1), se reescreve

$$\phi(x^{k+1})^{\frac{1}{2}} \leq \phi^k(x^{k+1})^{\frac{1}{2}} + \sqrt{2}c_k\|x^{k+1} - x^k\| \tag{5.1}$$

Pelo Lema 5.1.2, existem constantes positivas K e ε tais que, para $\phi(x^{k+1})^{\frac{1}{2}} \leq \varepsilon$,

$$\|x^{k+1} - x^*\| \leq K\phi(x^{k+1})^{\frac{1}{2}} \tag{5.2}$$

Além do mais, pelo critério (4.3) temos que

$$\phi^k(x^{k+1})^{\frac{1}{2}} \leq \delta_k\|x^{k+1} - x^k\| \tag{5.3}$$

Consequentemente, por (5.2), (5.1) e (5.3), temos

$$\begin{aligned}
\|x^{k+1} - x^*\| &\leq K\phi^k(x^{k+1})^{\frac{1}{2}} + \sqrt{2}Kc_k\|x^{k+1} - x^k\| \\
&\leq K\delta_k\|x^{k+1} - x^k\| + \sqrt{2}Kc_k\|x^{k+1} - x^k\| \\
&\leq K(\delta_k + \sqrt{2}c_k)\|x^{k+1} - x^k\| \\
&\leq K(\delta_k + \sqrt{2}c_k)(\|x^k - x^*\| + \|x^{k+1} - x^*\|)
\end{aligned}$$

Rearranjando os termos nesta desigualdade, temos:

$$\|x^{k+1} - x^*\| \leq \tfrac{K(\delta_k+\sqrt{2}c_k)}{1-K(\delta_k+\sqrt{2}c_k)}\|x^k - x^*\|$$

Visto que, por hipótese, $\delta_k \to 0$ e $c_k \to 0$, temos a propriedade desejada. ■

Se usarmos esse resultado para o Algoritmo 4.3.1, isto é, usando a regularização com métrica variável $F^k(x) = F(x) + c_k(X^k)^{-r}(x - x^k)$ (com $x^k \in \Re^n_{++}$), caímos na seguinte expressão:

$$\|x^{k+1} - x^*\| \leq \tfrac{K(\delta_k+\sqrt{2}c_k\|(X^k)^{-r}\|)}{1-K(\delta_k+\sqrt{2}c_k\|(X^k)^{-r}\|)}\|x^k - x^*\|$$

Nesta expressão, encontramos o seguinte problema: Apesar de $\delta_k \to 0$, $c_k \to 0$ e de $\{x^k\}$ ser uma sequência limitada, não temos como garantir que $\|(X^k)^{-r}\|$ também é limitada, isto é, a dificuldade encontrada neste ponto é a de controlar o termo $\|(X^k)^{-r}\|$. Como pesquisa futura, seria importante provar para o Algortimo 4.3.1 um resultado de convergência análogo ao Teorema 5.1.3, de modo a fazer com que este algoritmo tenha rápida convergência assim como o algoritmo proximal apresentado por Yamashita, Imai e Fukushima.

5.2 Análise do problema regularizado

No capítulo anterior vimos o algoritmo proximal inexato usando a função F regularizada (4.1), mostramos a prova de convergência global, porém não foi apresentada nenhuma regra para escolha dos parâmetros e também nenhum procedimento para a resolução de subproblemas foi citado.

Nesta seção, discutiremos questões relacionadas a escolha destes parâmetros. Em particular, será visto como atualizar c_k e δ_k e como resolver subproblemas $PCN(F^k)$.

5.2.1 Escolha dos parâmetros

Primeiramente, analisaremos a escolha da sequência de parâmetros $\{c_k\}$. Notemos que a prova de convergência global do algoritmo 4.3.1 requer a Hipótese 4.3.3. Considerando esta hipótese , podemos determinar $\{c_k\}$ da seguinte maneira:

$$c_k = min\{\beta^k, \phi(x^k)\} \tag{5.4}$$

onde β^k converge para zero (podemos tomar $\beta^0 \in (0,1)$) . O termo $min\{\beta^k, \phi(x^k)\}$ faz c_k convergir para zero, fazendo com que o item (A) da Hipótese 4.3.3 seja satisfeito. Notemos que não foi preciso colocar nenhuma condição adicional na expressão de c_k para que o item (B) da Hipótese 4.3.3 esteje assegurado, ao contrário de Yamashita, Imai e Fukushima [30] que acrescentaram o termo $min\{1, \frac{1}{\|x^k\|^2}\}$ na expressão de c_k para que o item (B) da hipótese usada por eles (Hipótese 4.3.2) seje válido.

Com relação ao parâmetro δ_k, é suficiente tomarmos o mesmo convergindo a zero. Desse modo, simplesmente determinamos δ_k por

$$\delta_k = \beta^k.$$

5.2.2 Como resolver subproblemas

Para resolver o subproblema $PCN(F^k)$, pode ser usado um método de Newton, de forma a minimizarmos a função de mérito regularizada. De Luca, Facchinei e Kanzow [7], criaram o Método Generalizado de Newton para equações não-suaves e Yamashita, Imai e Fukushima [30] o adaptaram de maneira a solucionar o $PCN(F^k)$. Eles usaram um elemento no subdiferencial de Θ, pois já vimos que este operador não é diferenciável no ponto (0,0). Porém, como no algoritmo 4.3.1 estamos considerando $x^k \in \Re^n_{++}$, usaremos um método de Newton para equações suaves adaptado ao nosso problema.

Primeiramente, veremos uma proposição que assegura a existência de solução para o problema regularizado.

Proposição 5.2.1 *Suponha que F é uma P_0 função contínua. Então Θ^k(e ϕ^k) é coercivo.*

Prova. Seja $\{a^m\}$ uma sequência qualquer tal que $\|a^m\| \to \infty$. Por hipótese, F é uma P_0 função contínua e assim, pelo Lema 2.6.3, existe uma subsequência, que sem perda de generalidade podemos escrever como $\{a_j^m\}$, e um índice $j \in \{1...n\}$ tal que uma das seguintes condições ocorre:

(i) $a_j^m \to +\infty$ e $\{F_j(a^m)\}$ é limitada inferiormente;

(ii) $a_j^m \to -\infty$ e $\{F_j(a^m)\}$ é limitada superiormente.

Pela definição de F^k, temos que

$$F_j^k(a^m) = F_j(a^m) + c_k(x_i^k)^{-r}a_j^m - c_k(x_i^k)^{1-r}$$

Assim, se ocorrer (i) temos que

$$F_j^k(a^m) \to \infty, \text{ quando } m \to \infty \implies \|\psi(a_j^m, F_j^k(a^m))\| \to \infty$$

Mas se ocorrer (ii), segue diretamente do Lema 3.1.11 que $\|\psi(a_j^m, F_j^k(a^m))\| \to \infty$ sempre que $m \to \infty$.

Agora, pela Definição 2.5.1 (coercividade), podemos concluir que o operador regularizado Θ^k é coercivo. E isto implica diretamente que sua função de mérito correspondente também o é. ∎

Exibiremos agora um resultado clássico de Banach e Mazur que envolve o conceito de coercividade de uma função.

Teorema 5.2.2 *Se $G : \Re^n \to \Re^n$ é uma função contínua, localmente injetiva e coerciva, então G é um (sobre)homeomorfismo do $\Re^n$.*

Prova. Veja ([2], Teorema 5.1.4).

Assim, veremos agora o teorema que mostra a unicidade de solução do sistema de equações $\Theta^k(x) = 0$, ou seja, provaremos que o problema regularizado tem solução única.

Teorema 5.2.3 *Se $F : \Re^n \to \Re^n$ é uma P_0-função contínua e F^k é definida como em (4.1), então o sistema de equações $\Theta^k(x) = 0$ tem solução única.*

Prova. Visto que F é uma P_0-função, F^k é uma P-função (Proposição 4.2.1). Assim, segue da Proposição 3.1.12, que Θ^k é uma P-função. Suponhamos por contradição que existe $x, y \in \Re^n, x \neq y$ tais que $\Theta^k(x) = \Theta^k(y) = 0$. Como Θ^k é uma P-função, temos que para $x, y \in \Re^n$ existe um índice i tal que $x_i \neq y_i$ e $(x_i - y_i)(\Theta_i^k(x) - \Theta_i^k(y)) > 0$, que cai em contradição com a igualdade anterior. Logo o sistema $\Theta^k(x) = 0$ tem solução única. ∎

Assim, pelo Lema 3.1.3 e pelo Teorema anterior, temos que o $PCN(F^k)$ tem solução única.

Notemos que, obviamente, Θ^k é um homeomorfismo no $\Re^n$ pelo Teorema 5.2.2.

Assim, apresentaremos no próximo capítulo um método tipo Newton para o problema regularizado.

Capítulo 6

O Método de Newton para o problema regularizado

A seguir, veremos um método tipo Newton para equações suaves adaptado ao problema regularizado $PCN(F^k)$.

6.1 Método de Newton para o $PCN(F^k)$

Algoritmo 6.1.1 .

Passo 0 : *Faça* $z^0 := x^k > 0$ *e escolha* $\rho \in (0, \frac{1}{2})$. *Tome* $j := 0$.

Passo 1 : *Calcule* $\nabla\Theta^k$ *no ponto* z^j *e encontre uma solução* d^j *da equação linear*

$$\nabla\Theta^k(z^j)d = -\Theta^k(z^j) \tag{6.1}$$

Verifique se $z^{j+1} = z^j + d^j > 0$ *e se o critério (4.3)* $(\phi^k(x^{k+1})^{\frac{1}{2}} \leq \delta_k.min\{1, \|x^{k+1} - x^k\|\})$ *é satisfeito com* $x^{k+1} := z^{j+1}$. *Em caso afirmativo, aceite* x^{k+1} *e termine a iteração. Caso contrário, vá para o Passo 2.*

Passo 2 (Teste da razão) : *Tome* $z^{j+1} = z^j + td^j$, *onde* t *é um escalar que é dado da seguinte maneira:*

Tome um t adequado tal que $t < min\{\frac{-z_i^j}{d_i^j}\}$, para todo i tal que $d_i^j < 0$ e vá para o Passo 3.

Passo 3 (Busca de Armijo) : *Encontre o menor inteiro não-negativo l tal que*

$$\phi^k(z^j + 2^{-l}td^j) \leq \phi^k(z^j) + \rho 2^{-l}t\nabla\phi^k(z^j)^t d^j$$

Tome $z^{j+1} := z^j + 2^{-l}td^j$ e vá para o Passo 4.

Passo 4 : *Se o critério (4.3) é satisfeito com $z^{j+1} > 0$ substituído por x^{k+1}, tome $x^{k+1} := z^{j+1}$ e termine a iteração. Caso contrário, faça $j := j + 1$ e volte ao Passo 1.*

O Teste da razão no Passo 2 garante que o iterado posterior será estritamente positivo. No caso de funções semi-suaves, a equação no Passo 1 iria requerer um custo maior, mas não seria difícil, como pode ser visto em [7].

Outro ponto em que devemos prestar a atenção é que a direção obtida pela solução da equação no Passo 1 do Método de Newton é sempre uma direção de descida para a função ϕ^k, a menos que $\Theta^k(x^k)$ seja igual a zero. Esta é uma propriedade padrão da direção de Newton para a solução de um sistema suave de equações, mas não é verdadeira, em geral, quando o sistema de equações é não-suave.

Visto que F^k é uma P-função para $x^k > 0$, todos os elementos de $\partial_B\Theta^k(x^k)$ são não-singulares (veja [7]). Assim a equação (6.1) no Passo 1 é solucionável para todo j, se $x^k > 0$. Além do mais, temos o seguinte Teorema.

Teorema 6.1.2 *Para cada k, o Método de Newton encontra uma solução aproximada do $PCN(F^k)$ que satisfaz o critério (4.3) do algoritmo inexato 4.3.1.*

Prova. Pela Proposição 5.2.1, a função de mérito ϕ^k é coerciva, isto é,

$$\lim_{\|x\|\to\infty} \phi^k(x) = +\infty.$$

Assim, temos que o conjunto $L(\alpha) = \{z^j : \phi^k(z^j) \leq \alpha\}$ é compacto.

Isto garante que a sequência gerada por um método de descida para minimizar ϕ^k é limitada, isto é, que o algoritmo gera uma sequência minimizante para a função ϕ^k.

Temos que mostrar agora que todo $\{z^j\}$ no Método de Newton converge para um ponto estacionário de ϕ^k. A prova é por contradição:

Assumimos, sem perda de generalidade, que a direção é sempre dada por $\nabla\Theta^k(z^j)d = -\Theta^k(z^j)$, para $z^j > 0$.

Suponha agora, renumerando se necessário, que $\{z^j\} \to z^*$ e que $\nabla\phi^k(z^*) \neq 0$ (onde z^* é o ótimo do $PCN(F^k)$, isto é, que minimiza ϕ^k). Lembramos que como F^k é continuamente diferenciável (F continuamente diferenciável e $x^k > 0$), ϕ^k também é continuamente diferenciável pela Proposição 3.1.10 (c). Temos também que $z^* \geq 0$, já que z^* é solução do $PCN(F^k)$.

Visto que a direção d^j satisfaz (6.1), podemos escrever

$$\|\Theta^k(z^j)\| = \|\nabla\Theta^k(z^j)d^j\| \leq \|\nabla\Theta^k(z^j)\|\|d^j\| \tag{6.2}$$

e assim

$$\|d^j\| \geq \frac{\|\Theta^k(z^j)\|}{\|\nabla\Theta^k(z^j)\|} \tag{6.3}$$

(reforçando que $\|\nabla\Theta^k(z^j)\|$ não pode ser zero, senão (6.2) iria implicar que $\Theta^k(z^j) = 0$ e z^j seria um ponto estacionário e o algoritmo teria parado).

Notemos que

$$0 < m \leq \|d^j\| \leq M \tag{6.4}$$

p/ m e M positivos. De fato se, para alguma subsequência J, $\{\|d^j\|\}_J \to 0$, de (6.3) temos que $\{\|\Theta^k(z^j)\|\}_J \to 0$, porque $\nabla\Theta^k(z^j)$ é limitado sobre a sequência limitada $\{z^j\}$ (propriedade do jacobiano). Mas então, por continuidade, $\Theta^k(z^*) = 0$ e então z^* seria solução do $PCN(F^k)$, contradizendo a hipótese de que $\nabla\phi^k(z^*) \neq 0$. Por outro lado $\|d^j\|$ não pode ser ilimitado porque $\nabla\phi^k(z^j)$ é limitada e isto iria contra sua a característica de descida (já que d^j é sempre direção de descida).

Então, visto que a busca linear no Passo 3 é segura a cada iteração (pois ϕ^k é continuamente diferenciável) e que ϕ^k é limitado sobre a sequência limitada $\{x^k\}$, temos que $\{\phi^k(z^{j+1}) - \phi^k(z^j)\} \to 0$, que implica pela busca linear que

$$\{2^{-l}t\nabla\phi^k(z^j)^T d^j\} \to 0 \tag{6.5}$$

Queremos mostrar que $\{2^{-l}t\}$ é limitado longe de zero (isto é, que não convirja diretamente a zero). Suponhamos o contrário. Então, subsequenciando se necessário, temos que $\{2^{-l}t\} \to 0$ de maneira que a cada iteração o tamanho do passo é reduzido em pelo menos uma vez (em pelo menos um) e a busca linear se dá da seguinte forma:

$$\frac{\phi^k(z^j + 2^{-(l-1)}td^j) - \phi^k(z^j)}{2^{-(l-1)}t} > \rho\nabla\phi^k(z^j)^T d^j. \tag{6.6}$$

Por (6.4) podemos assumir, subsequenciando se necessário que $\{d^j\} \to \overline{d} \neq 0$ e então, tomando limite em (6.6), temos que $\nabla\phi^k(z^*)^T\overline{d} \geq \rho\nabla\phi^k(z^*)^T\overline{d}$ ($\rho < \frac{1}{2}$). Pela condição de descida de d^j temos que $\nabla\phi^k(z^*)^T\overline{d} < 0$, o que cai em contradição com a desigualdade anterior, já que $\rho < \frac{1}{2}$. Assim $\{2^{-l}t\}$ é limitado longe de zero. Mas por (6.5) e o fato de $\{d^j\}$ ser uma direção de descida ($\nabla\phi^k(z^j)^T d^j < 0$) implicam que $\{d^j\} \to 0$, contradizendo (6.4). Logo $\nabla\phi^k(z^*) = 0$.

Assim, todo ponto limite de z^j gerado pelo método é um ponto estacionário de ϕ^k. E pela Proposição 3.1.10 (d), concluímos que toda sequência gerada pelo método é uma solução do $PCN(F^k)$. Como o $PCN(F^k)$ tem solução única, toda sequência z^j gerada pelo método converge para esta única solução z^*.

Assim, temos que para cada k no algoritmo proximal encontraremos um $z^{j+1} \in \Re^n_{++}$ tal que para $x^{k+1} = z^{j+1}$ e $x^k = z^j$, $\phi^k(x^{k+1})^{\frac{1}{2}} \leq \delta_k min\{1, \|x^{k+1} - x^k\|\}$, para j suficientemente grande no Método de Newton. ■

Assim, apresentaremos uma versão específica do algoritmo inexato 4.3.1, com F sendo uma P_0-função continuamente diferenciável.

6.2 Algoritmo do Ponto Proximal II

Algoritmo 6.2.1 Passo 1:. *Escolha* $\beta^0 \in (0,1)$ *e* $x^0 \in \Re^n_{++}$. *Tome*

$$c_0 = min\{1, \phi(x^0)\}$$

e $k := 0$.

Passo 2. *Se* x^k *satisfaz um critério de parada termine a iteração.* ($\|\phi(x^k)\| < \varepsilon$).

Passo 3. *Seja* $F^k = F(x) + c_k(X^k)^{-r}(x - x^k)$. *Pelo uso do Método de Newton, obtenha uma solução aproximada do* $PCN(F^k)$, $x^{k+1} \in \Re^n_{++}$, *de forma que*

$$\phi^k(x^{k+1})^{\frac{1}{2}} \leq \beta^k min\{1, \|x^{k+1} - x^k\|\}$$

Passo 4. *Escolha* $c_{k+1} := min\{\beta^{k+1}, \phi(x^{k+1})\}$, $\beta^{k+1} \in (0, \beta^k)$, $k := k+1$ *e volte ao Passo 2.*

Sobre o cálculo do jacobiano do operador Θ^k no Método de Newton, veremos a seguir alguns resultados importantes. Ele são reformulações dos Lemas 3.3 e 4.1 da referência [30] aplicado ao caso em que Θ^k é diferenciável, já que x^k é estritamente positiva. Assim $\partial_B \Theta^k(x^k) = \nabla\Theta^k(x^k)$.

Lema 6.2.2 *Suponha satisfeita a Hipótese 5.1.1. Então existe um* $C > 0$ *tal que para todo* k *suficientemente grande*

$$\|(\nabla\Theta^k(x^k))^{-1}\| \leq C$$

Demonstração. Como $x^k > 0$ e limitada, Θ é diferenciável em x^k. Além do mais, visto que $x^k \to x^*$ e que $\partial_B\Theta(x^*)$ é não-vazio e limitado, $\nabla\Theta(x^k)$ é limitado. Pela hipótese, temos que $\partial_B\Theta(x^*)$ é não-singular. Assim $\nabla\Theta^k(x^k)$ é também limitado e não-singular, de que decorre o resultado. ∎

Lema 6.2.3 *Suponha que* F *é continuamente diferenciável e que* $x^k \to x^*$. *Então existe uma constante positiva* k_1 *tal que*

$$\|\nabla\Theta(x^k) - \nabla\Theta^k(x^k)\| \leq k_1 c_k$$

para todo k *suficientemente grande.*

Demonstração. Como $x^k > 0$ e F é continuamente diferenciável, Θ e Θ^k são diferenciáveis em x^k. Como $F^k(x^k) = F(x^k)$ segue pela Proposição 3.1.9 que

$$|\nabla\Theta_i(x_i^k) - \nabla\Theta^k(x_i^k)| = c_k| - \nabla\Big((x_i^k)^{-r}(x_i - x_i^k)\Big)\Big(\frac{F_i(x_i^k)}{\|x_i^k, F_i(x_i^k)\|} - 1\Big)| \leq k_0 c_k.$$

Logo, temos que

$$\|\nabla\Theta(x^k) - \nabla\Theta^k(x^k)\| \leq k_1 c_k.$$

■

O lema anterior mostra que $\nabla\Theta^k(x^k)$ pode servir como uma aproximação de $\nabla\Theta(x^k)$ quando k é suficientemente grande.

Não necessariamente o Algoritmo 6.2.1 é eficiente na prática, porque nada sabemos sobre custos computacionais para resolver um subproblema a cada iteração. Então, é importante estimar o número de iterações que o Método de Newton gasta a cada iteração do algoritmo. Além do mais, é particularmente interessante ver sob que condições este método gasta apenas uma simples iteração dentro do Algoritmo 6.2.1 durante o processo de convergência.

Suponhamos que o conjunto S^* é não-vazio e que $\{x^k\}$ converge para a solução x^* do $PCN(F)$ que é BD-regular com Θ. Além do mais, seja x_N^k um ponto obtido por uma simples iteração do Método de Newton para $PCN(F^k)$, que é

$$x_N^k \equiv x^k - (\nabla\Theta^k(x^k))^{-1}\Theta^k(x^k) \tag{6.7}$$

A meta é mostrar que a desigualdade

$$\phi^k(x_N^k)^{\frac{1}{2}} \leq \beta^k min\{1, \|x_N^k - x^k\|\} \tag{6.8}$$

está assegurada para todo k suficientemente grande. Agora observe que

$$\begin{aligned} \|x_N^k - x^k\| &= \|(\nabla\Theta^k(x^k))^{-1}\Theta^k(x^k)\| \\ &\leq \|(\nabla\Theta^k(x^k))^{-1}\|\|\Theta^k(x^k)\| \\ &\leq C\|\Theta^k(x^k)\| \end{aligned}$$

para alguma constante $C > 0$. A existência de C é assegurada pelo Lema 6.2.2.

Visto que $\|\Theta^k(x^k)\| \to 0$, temos que $\|x_N^k - x^k\| < 1$ para k suficientemente grande.

Assim, sabendo que $\phi^k(x) = \frac{1}{2}\|\Theta^k(x)\|^2$, a partir de (6.8), se tivermos $x_N^k > 0$ para k suficientemente grande, seria suficiente verificarmos a seguinte desigualdade:

$$\|\Theta^k(x_N^k)\| \leq \sqrt{2}\beta^k \|x_N^k - x^k\|. \tag{6.9}$$

Esse é outro ponto a ser analisado em trabalhos futuros.

Conclusões

Neste trabalho, vimos como acontece a reformulação de um problema de complementaridade não-linear em um sistema de equações usando um método de regularização proximal com métrica variável. Para isso, foi utilizada a função de Fischer-Burmeister, que possui diversas propriedades que também foram vistas. Lembrando que assumimos que a função F do problema original $PCN(F)$ é continuamente diferenciável. Também foi apresentado um algoritmo proximal inexato utilizando a mesma regularização. Foi usado um critério introduzido por Rockafellar (1976) de modo a podermos analisar questões relacionadas a taxa de convergência . Para isso, utilizamos hipóteses adequadas sobre os parâmetros $c_k's$ e consideramos regularidade da solução x^* do $PCN(F)$.

Vimos também de que maneira podemos escolher os $c_k's$, satisfazendo a Hipótese 4.3.3 e de modo a obtermos uma solução aproximada para o problema regularizado. De certa forma, estes parâmetros foram escolhidos com melhores condições do que os do método proximal usado por Yamashita, Imai e Fukushima [30].

Para a resolução do problema regularizado, apresentamos um método tipo Newton que permite que a função de mérito ϕ^k seja minimizada de maneira a aproximar a solução do $PCN(F^k)$ a solução do $PCN(F)$, isto é, que possibilita a implementação do Algoritmo 4.3.1.

Com relação aos trabalhos futuros, apesar de termos usado a função de Fischer-Burmeister em nosso problema, seria interessante tentar usar uma outra função de mérito com o objetivo de se escolher melhor os parâmetros $c_k's$. Outra questão a ser considerada é a análise do algoritmo proximal quando aplicado à funções monótonas. Também seria importante desenvolver um algoritmo específico que convirja superlinearmente até mesmo para parâmetros que não convirjam a zero. E por fim, seria de fundamental importância analisarmos o comportamento do algoritmo através de sua implementação computacional.

Referências bibliográficas

[1] Auslender, A. *Optimisation: Méthodes Numériques*, Paris, Masson, 1976.

[2] Berger, M.S. *Nonlinearity and Funtional Analysis*, New York, Academic Press, 1977.

[3] Chen, B., Chen, X., Kanzow, C. "A penalized Fischer-Burmeister NCP-function: Theoretical investigation and Numerical results", Preprint 126, *Institute of Applied Mathematics*, University of Hamburg, Hamburg, Germany, 1997.

[4] Chen, X., Qi, L., Yufei, Y. "Lagrangian globalization methods for nonlinear complementarity problems", *Journal of Optimization Theory and Applications*, v. 112, pp. 77-95, 2002.

[5] Clarke, F.H. *Optimization and Nonsmooth Analysis*, New York, John Wiley, 1983.

[6] Cottle, R.W., Pang, J.S and Stone, R.E. *The Linear complementarity Problem*, New York, Academic Press, 1988.

[7] De Luca, T., Facchinei, F. and Kanzow, C. "A semismooth equation approach to the solution of nonlinear complementarity problems", *Mathematical Programming*, v. 75, pp. 407-439, 1996.

[8] Dontchev A.L and Zolezzi T. : "Well-Posed Optimisation Problems", *Lectures Notes in Mathematics 1543*, Springer Verlag, 1993.

[9] Facchinei, F. "Structural and Stability properties of P_0 nonlinear complementarity problems", *Mathematics of Operations Research*, Vol. 23, pp. 735-745, 1998.

[10] Facchinei, F. and Kanzow, C. "A nonsmooth inexact Newton method for the solution of large-scale of nonlinear complemetarity problems", Preprint 95, *Institute of Applied Mathematics*, University of Hamburg, Hamburg, Germany, 1995.

[11] Facchinei, F. and Kanzow, C. "Beyond Monotonicity in regularization methods for nonlinear complementarity problems, *SIAM Journal on Control and Optimization*, Vol. 37, pp. 1150-1161, 1999.

[12] Facchinei, F. and Soares, J. "A new merit function for nonlinear complementarity problems and a related algorithm", *SIAM Journal of Optimization*, Vol. 7, pp. 225-247, 1997.

[13] Facchinei, F., Fischer, A. and Kanzow C. "Inexact Newton methods for semismooth equations to variational inequality problems". In: G. Di Pillo and F.Giannessi (eds), *Nonlinear Optimization and Applications*", Plenum Press, New York, 1996, pp. 155-170.

[14] Ferris, M.C. and Pang, J.S. "Engineering and economic applications of complementarity problems", *SIAM Review*, Vol. 39, pp. 669-713, 1997.

[15] Fischer, A. "A special Newton-type optimization method", *Optimization*, Vol. 24, pp. 269-284, 1992.

[16] Fischer, A. "Coercivity conditions in nonlinear complementarity problem", SIAM *Rev*, Vol. 16, pp. 1-16, 1992.

[17] Jiang, H. and QI, L.. "A new nonsmooth equations approach to nonlinear complementarity problem, *SIAM Journal on Control and Optimization*, Vol. 37, pp. 178-193, 1997.

[18] Kanzow, C. "Global convergence properties of some iterative methods for linear complementarity problems", *SIAM Journal of Optimization*, Vol. 6, pp. 326-341, 1996.

[19] Mangasarian, O.L and Solodov, M.V. "Nonlinear complementarity as unconstrained and constrained minimization", *Mathematical Programming*, v. 62, pp. 277-297, 1993.

[20] Oliveira, G.L. and Oliveira, P.R. *A new class of proximal interior-point methods for optimization positivity constraints.* Techinical Report, PESC/COPPE - Federal University of Rio de Janeiro, Rio de janeiro, RJ, Brazil, 2002.

[21] Palais, R.S and Terng, C.L. "Critical point theory and submanifold geometry", 1353 *Lecture Note in Mathematics*, 1988.

[22] Pang, J.S and Qi. L. "Nonsmooth equations: Motivations and algorithms", *SIAM Journal on Optimization*, Vol. 3, pp. 443-465, 1993.

[23] PEREIRA DA SILVA, G.J., "*Uma Nova Classe de Algoritmos proximais para o problema de complementariedade não-linear*". Tese de D.Sc, COPPE/UFRJ, RIO de Janeiro, RJ, Brasil, 2002.

[24] Qi, L. *C-differentiability, C-differential operators and generalized Newton methods.* Technical Report, *School of mathematics*, University of New South Wales, Sydney, Austrália, 1996.

[25] Qi, L. and Sun, D. "Nonsmooth Equations and Smoothing Newton Methods",

School of Mathematics, University of New South Wales, Sydney 2052, Australia, 1998.*Institute of Applied Mathematics*, University of Hamburg, Hamburg, Germany, 1995.

[26] Qi, L. and Sun, J. *A nonsmooth version's of Newton's methods.* Mathematical Programming Series A , 58, pp. 353-368, 1993.

[27] Ravindran, G. and Gowda, S. "Regularization of P_0-functions in box variational inequality problems", SIAM *Journal of Optimization*, v. 11, pp. 748-760, 2000.

[28] Rockafellar, R.T. "Monotone operators and the proximal point algorithm", SIAM *Journal on Control and Optimization*, v. 14, pp. 877-898, 1976.

[29] Tseng, P. "Growth behavior of a class of merit functions for the nonlinear complementarity problem", *Journal of Optimization Theory and Aplications*", v. 89, pp. 17-37, 1996.

[30] Yamashita, N., Imai, I. and Fukushima, M. "The proximal point algorithm for the P_0 complementarity problem". In: M.C. Ferris, O.L. Mangasarian and J.S. Pang (eds), *Complementarity: Applications, algorithms and extensions*, Boston, Kluwer Academic Publishers, pp. 361-379, 2001.

[31] Yamashita, N., Imai, I. and Fukushima, M. The proximal point algorithm with genuine superlinear convergence for the monotone complementarity problem. Technical Report 99012, *Department of Applied Mathematics and Physics*, Kyoto, Japan, 1999.

[32] Yamashita, N., Taji, K. and Fukushima, M. "Unconstrained optimization reformulations of variational inequality problems", *Journal of Optimization Theory and Applications*, v. 92, pp. 439-456, 1997.