

Estratégias de Segunda Ordem para Problemas de Complementaridade

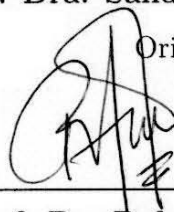
Este exemplar corresponde à redação final da tese devidamente corrigida e defendida por Wesley Vagner Inês Shirabayashi e aprovada pela comissão julgadora.

Campinas, 25 de setembro de 2009.



Profa. Dra. Sandra Augusta Santos

Orientadora



Prof. Dr. Roberto Andreani

Co-orientador

Banca Examinadora:

1. Profa. Dra. Sandra Augusta Santos
2. Prof. Dr. Lúcio Tunes dos Santos
3. Prof. Dr. Marcos Nereu Arenales
4. Profa. Dra. Susana Scheimberg de Makler
5. Profa. Dra. Elizabeth Wegner Karas

Tese apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP como requisito parcial para obtenção do título de **Doutor em Matemática Aplicada.**

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

Bibliotecária: Crislene Queiroz Custódio – CRB8 / 7966

Shirabayashi, Wesley Vagner Inês

Sh65e Estratégias de segunda ordem para problemas de complementaridade /
Wesley Vagner Inês Shirabayashi -- Campinas, [S.P. : s.n.], 2009.

Orientadores : Sandra Augusta Santos ; Roberto Andreani

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

1. Sistemas singulares. 2. Complementaridade. 3. Método preditor-
corretor. 4. Método de Newton. 5. Cones poliedrais. 6. Métodos de
pontos interiores. I. Santos, Sandra Augusta. II. Andreani, Roberto III.
Universidade Estadual de Campinas. Instituto de Matemática, Estatística
e Computação Científica. IV. Título.

Título em inglês: Second order strategies for complementarity problems.

Palavras-chave em inglês (Keywords): 1. Singular systems. 2. Complementarity. 3. Predictor-corrector methods.
4. Newton method. 5. Polyhedral cones. 6. Interior-point methods.

Área de concentração: Otimização

Titulação: Doutor em Matemática Aplicada

Banca examinadora: Profa. Dra. Sandra Augusta Santos (IMECC-UNICAMP)
Prof. Dr. Lúcio Tunes dos Santos (IMECC-UNICAMP)
Prof. Dr. Marcos Nereu Arenales (ICMC-USP)
Profa. Dra. Suzana Scheimberg de Makler (UFRJ)
Profa. Dra. Elizabeth Wegner Karas (UFPR)

Data da defesa: 25/09/2009

Programa de Pós-Graduação: Doutorado em Matemática Aplicada

Tese de Doutorado defendida em 25 de setembro de 2009 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.



Prof(a). Dr(a). SANDRA AUGUSTA SANTOS



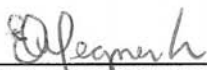
Prof(a). Dr(a). LUCIO TUNES DOS SANTOS



Prof(a). Dr(a). MARCOS NEREU ARENALES



Prof(a). Dr(a). SUSANA SCHEIMBERG DE MAKLER



Prof(a). Dr(a). ELIZABETH WEGNER KARAS

Agradecimentos

A Deus por me conceder a graça da vida e permitir a realização deste trabalho.

Aos meus pais João e Luisa.

Aos Professores Sandra Augusta Santos e Roberto Andreani, pelo companheirismo e pela excelente orientação.

Aos professores e funcionários do IMECC-Unicamp e do DMA-UEM que de um modo ou outro tiveram participação neste trabalho.

Aos sempre indispensáveis amigos. O que seria de nossa vida sem eles.

Ao CNPq pelo apoio financeiro.

Não sei se a vida é curta ou longa para nós, mas sei que nada do que vivemos tem sentido se não tocarmos o coração das pessoas.

Muitas vezes basta ser: colo que acolhe, braço que envolve, palavra que conforta, silêncio que respeita, alegria que contagia, lágrima que corre, olhar que acaricia, desejo que sacia, amor que promove.

E isso não é coisa de outro mundo, é o que dá sentido à vida.

É o que faz com que ela não seja nem curta, nem longa demais, mas que seja intensa, verdadeira, pura enquanto durar.

“Feliz aquele que transfere o que sabe e aprende o que ensina.”

(Cora Coralina)

Resumo

Neste trabalho reformulamos o problema de complementaridade não linear generalizado (GNCP) em cones poliedrais como um sistema não linear com restrição de não negatividade em algumas variáveis, e trabalhamos na resolução de tal reformulação por meio de estratégias de pontos interiores. Em particular, definimos dois algoritmos e provamos a convergência local de tais algoritmos sob hipóteses usuais. O primeiro algoritmo é baseado no método de Newton, e o segundo, no método tensorial de Chebyshev. O algoritmo baseado no método de Chebyshev pode ser visto como um método do tipo preditor-corretor. Tal algoritmo, quando aplicado a problemas em que as funções envolvidas são afins, e com escolhas adequadas dos parâmetros, torna-se o bem conhecido algoritmo preditor-corretor de Mehrotra. Também apresentamos resultados numéricos que ilustram a competitividade de ambas as propostas.

Palavras-chave: Sistemas singulares, Complementaridade, Método preditor-corretor, Método de Newton, Cones poliedrais, Métodos de pontos interiores.

Abstract

In this work we reformulate the generalized nonlinear complementarity problem (GNCP) in polyhedral cones as a nonlinear system with nonnegativity in some variables and propose the resolution of such reformulation through interior-point methods. In particular we define two algorithms and prove the local convergence of these algorithms under standard assumptions. The first algorithm is based on Newton's method and the second, on the Chebyshev's tensorial method. The algorithm based on Chebyshev's method may be considered a predictor-corrector one. Such algorithm, when applied to problems for which the functions are affine, and the parameters are properly chosen, turns into the well-known Mehrotra's predictor corrector algorithm. We also present numerical results that illustrate the competitiveness of both proposals.

Keywords: Singular systems, Complementarity, Predictor-corrector method, Newton method, polyhedral cones, interior-point methods.

Lista de Tabelas

5.1	Problemas gerados aleatoriamente no ortante positivo.	70
5.2	Resultados do problema de complementaridade implícita POZ1.	72
5.3	Resultados do problema de complementaridade implícita POZ2.	80
5.4	Resultados do problema de complementaridade implícita modificado PMIC. 81	
5.5	Resultados da Família 1 de problemas com cones poliedrais em \mathbb{R}^n (ma- trizes indefinidas e não simétricas).	82
5.6	Resultados da Família 2 de problemas com cones poliedrais em \mathbb{R}^n (ma- trizes indefinidas e simétricas).	83
5.7	Resultados da Família 3 de problemas com cones poliedrais em \mathbb{R}^n (ma- trizes indefinidas, não simétricas e singulares).	84
5.8	Resultados da Família 4 de problemas com cones poliedrais em \mathbb{R}^n (ma- trizes indefinidas, não simétricas e solução sem complementaridade estrita). 85	
5.9	Resultados da Família 1 de problemas com cones 3D (T é indefinida).	86
5.10	Resultados da Família 2 de problemas com cones 3D(T é definida positiva). 87	
5.11	Resultados da Família 4 de problemas com cones 3D (T semidefinida negativa).	88
B.1	Problemas gerados aleatoriamente no ortante positivo.	112
B.2	$\ H_k\ _1$ quando o algoritmo parou devido a $\ w_k - w_{k-1}\ _1 < \epsilon$	113
B.3	Problemas de complementaridade implícita POZ1.	113
B.4	Problemas de complementaridade implícita POZ2.	114
B.5	Problema de complementaridade implícita POZ3.	115

B.6	Família 1, matrizes indefinidas e não simétricas.	116
B.7	$\ H_k\ _1$ quando o algoritmo parou devido a $\ w_k - w_{k-1}\ _1 < \epsilon$ (família 1).	116
B.8	Família 2, matrizes indefinidas e simétricas.	117
B.9	$\ H_k\ _1$ quando o algoritmo parou devido a $\ w_k - w_{k-1}\ _1 < \epsilon$ (família 2).	117
B.10	Família 3, matrizes indefinidas, não simétricas e singulares.	118
B.11	Família 4, matrizes indefinidas, não simétricas e solução sem complementaridade estrita.	119
B.12	$\ H_k\ _1$ quando o algoritmo parou devido a $\ w_k - w_{k-1}\ _1 < \epsilon$ (família 4).	119
B.13	Família 1, T indefinida.	120
B.14	Família 2, T definida positiva.	121
B.15	Família 4, T semidefinida negativa.	122

Sumário

1	Introdução	1
2	Sistemas Não Lineares	5
2.1	Método de Newton	6
2.1.1	Métodos tipo Newton	7
2.1.2	Método de Newton em problemas não suaves	8
2.1.3	Método de Newton em problemas singulares	8
2.2	Métodos tensoriais para sistemas não lineares	15
2.2.1	Família Chebyshev-Halley	16
2.2.2	Método tensorial de Schnabel-Frank	18
3	Família Chebyshev-Halley em Sistemas Singulares	21
3.1	Resultado de Cavanagh revisitado	21
3.2	Métodos tensoriais em problemas singulares	30
3.2.1	Adaptação para o método de Chebyshev	30
3.2.2	Análise da família Chebyshev-Halley	35
4	Problemas de Complementaridade	39
4.1	Definições	40
4.1.1	Problema de complementaridade generalizado (GNCP)	41
4.1.2	Desigualdades variacionais	42
4.2	Métodos e Algoritmos	43

4.2.1	Reformulações como problemas de otimização	43
4.2.2	Reformulações como sistemas não lineares	44
4.2.3	Métodos de pontos interiores	44
5	GNCP em Cones Poliedrais	47
5.1	GNCP em cones poliedrais e um problema equivalente	48
5.2	Um algoritmo de Newton perturbado e pontos interiores	51
5.2.1	O Algoritmo 1 está bem definido	52
5.2.2	Resultados preliminares	54
5.2.3	Convergência local do Algoritmo 1	59
5.3	Um algoritmo preditor-corretor de pontos interiores	60
5.4	Experimentos Numéricos	67
5.4.1	Problemas gerados aleatoriamente no ortante positivo	68
5.4.2	Problemas de complementaridade implícita	69
5.4.3	Problemas com cones gerais em \mathbb{R}^n	74
5.4.4	Problemas em cones 3D com controle de faces geradas	76
5.4.5	Comentários gerais sobre os experimentos numéricos	78
6	Considerações Finais	89
	Bibliografia	91
A	Operações Tensoriais	109
B	Tabelas	111

Capítulo 1

Introdução

Uma ferramenta para resolver problemas de complementaridade e generalizações são os métodos de pontos interiores. A idéia consiste em reformular o problema como um sistema não linear, com não negatividade em algumas variáveis, e aplicar o método de Newton, preservando a não negatividade das variáveis. Os métodos de pontos interiores surgiram na década de 80 como alternativa ao método Simplex para resolver problemas de programação linear. Desde então este tipo de método também tem ganhado destaque na resolução de outros problemas, como programação quadrática, programação convexa, programação não linear e, como já citado, também em problemas de complementaridade.

Nesta tese trabalhamos com o problema de complementaridade não linear generalizado (GNCP) em cones poliedrais. Desenvolvemos dois algoritmos de pontos interiores baseados em métodos para resolver sistemas não lineares. O primeiro algoritmo é uma adaptação do método de Newton, e o segundo é baseado no método tensorial de Chebyshev. Mostramos a boa definição e convergência local quadrática dos métodos propostos, sob hipóteses usuais da literatura.

Uma hipótese restritiva nos problemas de complementaridade é a complementaridade estrita, que reflete na não singularidade do sistema associado. Por isso no Capítulo 2 fazemos uma revisão exaustiva dos resultados para sistemas singulares com

o método de Newton e no Capítulo 3 tratamos de sistemas singulares usando estratégias tensoriais (família de métodos Chebyshev-Halley). O estudo de métodos para resolver sistemas não lineares singulares teve maior ênfase a partir do final da década de 60, com os trabalhos de Rall [140] e Cavanagh [26]. No início da década de 80 houve um interesse de mais pesquisadores nesse tema, como Decker, Kelley, Keller e Griewank, dentre outros [36, 37, 38, 39, 40, 63, 64, 65, 66]. Em seus trabalhos, os autores focaram seus estudos no método de Newton, em particular, analisaram condições sob as quais o método gera uma sequência convergente a uma solução no caso do sistema possuir Jacobiano singular. No início da década de 90, com a advinda dos métodos de pontos interiores primais-duais, foi observado que esses métodos baseavam-se no método de Newton e conseqüentemente mostraram ter convergência quadrática em problemas não singulares, mas curiosamente vários trabalhos descreviam uma convergência super-linear em problemas singulares, onde em cada iteração o sistema a ser resolvido era não singular. Essa estrutura de não singularidade é garantida pela própria essência dos métodos de pontos interiores, por aproximarem-se da solução pelo interior do ortante não negativo, que torna-se um cone de inversibilidade para o Jacobiano do sistema não linear.

Schnabel e Frank em [151] introduzem um novo método tensorial que, segundo os autores, tem melhor desempenho em problemas singulares que os métodos do tipo Newton. Outros métodos tensoriais conhecidos e alguns bastante antigos surgiram quando pesquisadores formularam acelerações para o método de Newton, por exemplo o método de Halley que data de 1792, contemporâneo e amigo de Newton. O método de Halley foi inicialmente chamado “aceleração de Halley do método de Newton”. Posteriormente surgiram outros métodos, entre os mais conhecidos estão o método de Chebyshev e o método super-Halley. Esses métodos podem ser vistos como casos particulares de uma família de métodos chamada família de métodos Chebyshev-Halley [75], que sob hipóteses usuais da literatura têm ordem de convergência cúbica.

Durante os estudos dos vários resultados sobre o método de Newton para sistemas não lineares singulares conseguimos estender os resultados de Cavanagh [26] para

a família de métodos de Chebyshev-Halley [75]. Fizemos uma descoberta bem interessante em relação ao trabalho de Cavanagh, no qual é sugerida uma correção (teórica) para restaurar a convergência quadrática do método de Newton em problemas singulares. Percebemos que a direção corretora dos métodos Chebyshev-Halley aproxima assintoticamente a correção sugerida por Cavanagh, sendo a correção sugerida por ele “impossível” de se obter na prática, pois necessita do prévio conhecimento da solução do problema. A família Chebyshev-Halley, como descrita por Gundersen e Steihaug em [68], pode ser caracterizada como uma família do tipo preditor-corretor. A direção preditora é a direção de Newton e a corretora utiliza-se da já calculada direção preditora, e incorpora a parte tensorial no vetor dos termos independentes do sistema.

Os resultados sobre o método de Newton em problemas singulares obtidos por Reddien [141, 142] e Decker, Keller e Kelley [36] são diferentes dos obtidos por Cavanagh [26] e utilizam hipóteses diferentes. Recentemente, Oberlin e Wright em [126] obtiveram uma importante relação com a 2-regularidade introduzida por Izmailov e Solodov em [81]. Neste trabalho, Oberlin e Wright utilizaram as equivalências por eles obtidas para desenvolver novas técnicas para resolver problemas de complementaridade não linear por meio de métodos semi-suaves.

No estudo de como adaptar a família Chebyshev-Halley à filosofia de pontos interiores para resolver problemas de complementaridade não linear constatamos que o método de Chebyshev quando aplicado a um problema de complementaridade linear (LCP) nos dava um famoso método de pontos interiores: o preditor-corretor de Mehrotra [118]. Sendo mais específico, a direção corretora do método de Mehrotra é exatamente a direção corretora do método de Chebyshev. Isto nos dá uma outra explicação ao fato já conhecido e citado, em particular, por Wright em [162, p. 195], que o preditor-corretor de Mehrotra é um método de ordem superior. No trabalho de Tapia, Zhang, Saltzman e Weiser [155] os autores mostraram que o preditor-corretor de Mehrotra pode ser visto como um método de Newton composto perturbado de nível 1. Com esta observação feita por nós obtivemos um outro modo de ver o preditor-corretor de Mehrotra.

No Capítulo 2 fazemos breves comentários sobre sistemas não lineares e sobre métodos para resolvê-los. Uma ênfase maior é dada aos resultados sobre sistemas singulares. No Capítulo 3 reescrevemos o resultado de Cavanagh [26] sobre sistemas singulares e o estendemos para a família Chebyshev-Halley. No Capítulo 4 definimos o problema de complementaridade e algumas de suas variantes, e também tecemos comentários sobre alguns métodos para resolver tais problemas. No Capítulo 5 propomos uma nova reformulação do GNCP em cones poliedrais de uma maneira sob a qual possam ser aplicados métodos de pontos interiores, definimos dois algoritmos para resolver a reformulação e provamos a convergência local de tais algoritmos. O primeiro algoritmo é baseado no método de Newton e o segundo, no método de Chebyshev. Um extenso conjunto de experimentos coloca em perspectiva o desempenho dos algoritmos propostos. No Capítulo 6 apresentamos nossas considerações finais e descrevemos algumas possibilidades de trabalhos para o futuro.

Capítulo 2

Sistemas Não Lineares

O estudo de métodos para se resolver sistemas não lineares não é recente, mas o surgimento cada vez maior de problemas do cotidiano formulados sob esta ótica fez com que os estudos focassem também na eficácia prática dos métodos. Um grande exemplo disso é o desenvolvimento dos métodos quase-Newton, que não têm o poder teórico de convergência local quadrática do método de Newton, mas na prática são bastante eficientes.

Considere $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ e o problema de encontrar $x \in \mathbb{R}^n$ tal que:

$$F(x) = 0. \tag{2.1}$$

Existem vários métodos iterativos para resolver este problema, sendo o método de Newton, provavelmente, o mais conhecido. Uma característica importante deste método é sua rápida convergência local. Recentemente, tem-se tornado de grande interesse a análise de métodos para sistemas não lineares que não satisfaçam a hipótese de inversibilidade do Jacobiano na solução.

Em geral se supõe que F seja suficientemente suave. Neste caso, uma solução x_* para o problema (2.1) é dita degenerada, ou singular, se a matriz Jacobiana $F'(x_*)$ é singular, caso contrário a solução é dita não degenerada.

O problema pode ter uma única solução, muitas soluções (um número finito ou infinitas soluções) ou nenhuma solução.

O método de Newton é um dos mais conhecidos e utilizados para resolver sistemas não lineares. Apesar de ser conhecido há mais de 300 anos, ele ainda hoje é objeto de muitas pesquisas, sejam aplicadas a problemas diferenciáveis como também a problemas não diferenciáveis.

2.1 Método de Newton

Nesta seção focalizamos no método de Newton, algumas de suas variantes e modificações. Alguns trabalhos interessantes sobre o desenvolvimento do método de Newton são Yamamoto [168] e Galántai [56].

No método de Newton, dada uma aproximação x_k para uma solução x_* , a aproximação seguinte é dada por

$$x_{k+1} = x_k + d_k, \tag{2.2}$$

na qual d_k é solução do seguinte sistema linear

$$F'(x_k)d_k = -F(x_k), \tag{2.3}$$

desde que $F'(x_k)$ seja não singular.

O método de Newton resolve a cada iteração um modelo linear para F em torno do ponto x_k .

Segundo Yamamoto [168] foi Fourier, em 1818, quem provou a convergência quadrática do método de Newton para o caso real, $n = 1$, e em 1829 Cauchy provou um teorema de convergência sem assumir a existência de uma solução.

O tipo de resultado no qual se assume a existência de uma solução é conhecido como “Teorema de convergência local”, já quando não se assume a existência de uma solução, mas sim que uma aproximação inicial x_0 satisfaça algumas hipóteses, o resultado é conhecido como “Teorema de convergência semi-local”.

2.1.1 Métodos tipo Newton

Esses métodos são variações do método de Newton definidos por

$$x_{k+1} = x_k - [A(x_k)]^{-1}F(x_k), \quad (2.4)$$

em que $A(x)$ pode ser uma aproximação do Jacobiano de F ou outro tipo de matriz.

O próprio método de Newton é obtido fazendo-se $A(x) = F'(x)$. Quando $A(x) = F'(x_0)$ tem-se o método de Newton simplificado ou estacionário. Outra variação bastante comum é o método de Newton com recomeços, no qual é fixado um inteiro m , e se k é um múltiplo de m toma-se $A(x_k) = F'(x_k)$, caso contrário $A(x_k) = A(x_{k-m})$.

Quando são utilizadas diferenças finitas para se aproximar o Jacobiano de F , o método é chamado de Newton discreto. Também existem os chamados métodos de Newton inexatos, no qual cada sistema linear é resolvido por meio de um método iterativo com precisão pré-determinada. Uma classe importante dos métodos tipo Newton são os métodos secantes.

Métodos secantes

É um caso particular de método tipo Newton no qual é obtida uma aproximação para $F'(x)$ a cada iteração. No caso unidimensional $F'(x_k)$ é substituída por

$$\frac{F(x_k) - F(x_{k-1})}{x_k - x_{k-1}}.$$

Mas no caso n -dimensional existem infinitos modos de fazer esta aproximação, cada uma resultando em um método distinto. Os métodos secantes de maior valia prática são aqueles nos quais a aproximação seguinte para $F'(x_{k+1})$, digamos B_{k+1} , é obtida a partir da aproximação atual B_k por meio de “poucos” cálculos. Em geral, utiliza-se uma correção de posto 1 ou 2. Entre os métodos mais famosos e utilizados estão os métodos de Broyden; o método DFP, proposto por Davidon, Fletcher e Powell; e o método BFGS, proposto por Broyden, Fletcher, Goldfarb e Shanno. Para mais detalhes ver [42].

2.1.2 Método de Newton em problemas não suaves

Equações não suaves aparecem com frequência em vários problemas, por exemplo reformulações de problemas de complementaridade não linear (NCP's), desigualdades variacionais e otimização não linear. Clarke em [30] propôs um método de Newton generalizado usando um elemento pertencente ao Jacobiano generalizado $\partial F(x_k)$, quando F é localmente Lipschitziana mas não é diferenciável. Qi e Sun em [138] estenderam resultados de convergência do método de Newton para o caso de sistemas não suaves utilizando o Jacobiano generalizado de Clarke. Sun, Womersley e Qi em [154] propuseram um método de Newton semi-suave para resolver uma reformulação de problemas de complementaridade mista. Reformulações de problemas de complementaridade não linear foram estudadas e resolvidas por meio de métodos de Newton generalizados para equações não suaves, ver [35, 47, 74, 166, 169]. Martínez e Qi em [115] aplicaram métodos de Newton inexatos para resolver equações não suaves. Lopes e Martínez em [111] estudaram a convergência de métodos quase-Newton em problemas não suaves. Resultados de convergência superlinear foram obtidos para estes tipos de métodos.

2.1.3 Método de Newton em problemas singulares

Conforme Rall [140], Schröder em 1870 mostrou que o método de Newton para equações singulares tinha convergência linear e que a convergência quadrática podia ser restaurada. Rall [140] define um tipo de generalização do resultado de Schröder, mas no livro de Ortega e Rheinboldt [127, NR 10.2-5 p. 317], os autores contam que o resultado de Rall [140] continha uma falha, que foi posteriormente solucionada por Cavanagh em [26]. O resultado de Rall diz que a convergência quadrática pode ser restaurada aplicando-se uma determinada correção à direção de Newton. Cavanagh obteve um resultado similar, exibindo um operador projeção que serve de correção aos iterandos do método de Newton.

Nesta seção são apresentados alguns dos primeiros e principais resultados conhecidos sobre a convergência do método de Newton em problemas singulares. O

2.1. MÉTODO DE NEWTON

primeiro resultado já foi comentado no início e, segundo Rall [140], foi Schröder quem o provou.

Teorema 2.1. *Suponha que x_* seja uma solução de multiplicidade m da equação $F(x) = 0$, $F : \mathbb{R} \rightarrow \mathbb{R}$. Se a sequência $\{x_k\}$ definida por*

$$x_{k+1} = x_k - m[F'(x_k)]^{-1}F(x_k),$$

converge para x_ , então a convergência é quadrática.*

Observação 2.2. *Note que a função F no resultado anterior é uma função de apenas uma variável.*

O próximo resultado foi o obtido por Cavanagh [26] para sistemas singulares. Para prosseguir a análise estabelecemos algumas notações: Seja x_* uma solução de (2.1) tal que $F'(x_*)$ não seja inversível e suponha que $F'(x)$ seja inversível em

$$\mathcal{V}_\delta = \{x \mid 0 < \|x - x_*\|_2 < \delta\}$$

para algum $\delta > 0$. Sejam também N o núcleo de $F'(x_*)$, X tal que $\mathbb{R}^n = N \oplus X$, onde \oplus denota soma direta, P_N o projetor sobre N paralelo a X e $P_X = I - P_N$ o projetor sobre X paralelo a N . Durante todo o texto $\|\cdot\|$ denotará a norma euclidiana em \mathbb{R}^n .

Teorema 2.3. *Sejam $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, onde D é aberto, e x_* uma solução de $F(x) = 0$. Suponha a validade das seguintes hipóteses:*

1. $F \in C^2$ em D e $F''(x)$ é contínua com módulo de continuidade conhecido $w_1(r)$ (ver 2.4);
2. $F'(x)$ é não singular em uma vizinhança de x_* , $\mathcal{V}_\delta \subset D$;
3. Existe uma constante $c > 0$ tal que para todos $s \in N$ e $x \in \mathbb{R}^n$ vale a desigualdade:

$$\|F''(x_*)xs\| \geq c\|x\|\|s\|.$$

Então, o método de Newton gera uma sequência $\{x_k\}$ tal que

$$x_k \longrightarrow x_*, \quad P_X(x_k - x_*) \longrightarrow 0 \text{ superlinearmente e } \frac{\|P_N(x_k - x_*)\|}{\|P_N(x_{k-1} - x_*)\|} \longrightarrow \frac{1}{2}.$$

Observação 2.4. O módulo de continuidade é um tipo de generalização da Lipschitz continuidade e é definido por:

$$\omega_1(r) = \sup_{x \in D} \sup_{\|x-y\| \leq r} \|F(x) - F(y)\|.$$

Quando o Teorema 2.3 é citado em [127] o módulo de continuidade é substituído pela Lipschitz continuidade.

Observação 2.5. O significado de $F''(x_*)x$ é detalhado no Apêndice A.

Observação 2.6. A terceira hipótese do resultado de Cavanagh restringe a dois a multiplicidade da solução x_* .

Reddien [141] mostra um resultado um pouco mais geral que o Teorema 2.3, no sentido de que o Jacobiano de F pode ser singular em uma curva passando por x_* e não somente no ponto x_* , mas Reddien também coloca uma exigência sobre a dimensão do núcleo de N . Neste caso em [141], o autor exige que $\dim N = 1$, exigência esta que não é feita no Teorema 2.3. No que segue é considerado que $F : Y \longrightarrow Y$, onde Y é um espaço de Banach, e $Y = N \oplus X$, sendo N, X, P_N, P_X como definidos anteriormente, mas substituindo \mathbb{R}^n por Y . Antes de enunciar o resultado precisamos definir mais alguns conjuntos:

$$B_\rho(x_*) = \{x \in Y \mid \|x - x_*\| \leq \rho\}, \tag{2.5}$$

$$C_\theta(x_*) = \{x \in Y \mid \|P_X(x - x_*)\| \leq \theta \|P_N(x - x_*)\|\}. \tag{2.6}$$

Agora estamos aptos a enunciar o resultado de Reddien [141]:

Teorema 2.7. *Suponha que:*

2.1. MÉTODO DE NEWTON

1. $F \in C^3$;
2. $\dim N = 1$;
3. $F''(x_*)NN \cap X = \{0\}$, onde $F''(x_*)NN \equiv \{F''(x_*)s_1s_2 \mid s_1, s_2 \in N\}$;
4. Existe uma constante $c > 0$ tal que para todos $s \in N$ e $x \in Y$ vale a desigualdade:

$$\|F''(x_*)xs\| \geq c\|x\|\|s\|.$$

Então, existem constantes ρ , θ e c_1 positivas tais que $F'(x)$ é inversível em $W_{\rho,\theta}(x_*) \equiv B_\rho(x_*) \cap C_\theta(x_*)$, $x \neq x_*$, satisfazendo $\|(F'(x))^{-1}\| \leq c_1\|x - x_*\|^{-1}$, a aplicação $G(x) = x - (F'(x))^{-1}F(x)$ leva $W_{\rho,\theta}(x_*)$ nele próprio e $F^{-1}(0) \cap B_\rho(x_*) = x_*$. Além disso, definindo $x_k = G(x_{k-1})$ com $x_0 \in W_{\rho,\theta}(x_*)$, existe uma constante positiva c_2 tal que:

$$x_k \longrightarrow x_*, \quad \|P_X(x_k - x_*)\| \leq c_2\|x_k - x_*\|^2 \quad e \quad \frac{\|P_N(x_k - x_*)\|}{\|P_N(x_{k-1} - x_*)\|} \longrightarrow \frac{1}{2}.$$

Como no trabalho de Cavanagh, a quarta hipótese restringe a dois a multiplicidade da solução x_* . Posteriormente, Reddien em [142] estendeu o resultado do Teorema 2.7 de duas maneiras, a primeira para $\dim N > 1$ finita, e a segunda para zeros de ordem maior que dois.

Para a primeira extensão precisamos definir o seguinte conjunto:

$$T_\phi(x_*) = \{x \in X \mid \|(I - P_L)P_N(x - x_*)\| \leq \phi\|P_N(x - x_*)\|\}, \quad (2.7)$$

onde P_L é qualquer projeção limitada em um dado conjunto L .

Teorema 2.8. *Suponha que:*

1. $F \in C^3$;
2. $\dim N > 1$ e finita;
3. Seja L um subespaço unidimensional de N tal que $F''(x_*)LN \cap X = \{0\}$, onde $F''(x_*)LN \equiv \{F''(x_*)s_1s_2 \mid s_1 \in L, s_2 \in N\}$;

4. Existe uma constante $c > 0$ tal que para todos $s \in L$ e $x \in Y$ vale a desigualdade:

$$\|F''(x_*)xs\| \geq c\|x\|\|s\|.$$

Então, existem constantes ρ, θ, ϕ e c_1 positivas tais que $F'(x)$ é inversível em $W_{\rho,\theta,\phi}(x_*) \equiv B_\rho(x_*) \cap C_\theta(x_*) \cap T_\phi(x_*)$, $x \neq x_*$, satisfazendo $\|(F'(x))^{-1}\| \leq c_1\|x - x_*\|^{-1}$. Defina $G(x) = x - (F'(x))^{-1}F(x)$. Existem constantes $\rho > 0$, $\theta > 0$, ϕ e λ tais que se $x_0 \in T_\lambda(x_*)$, então a sequência $x_k = G(x_{k-1})$, $k \geq 1$, permanece em $W_{\rho,\theta,\phi}(x_*)$ e converge para x_* . Além disso, existe uma constante positiva c_2 tal que:

$$\|P_X(x_k - x_*)\| \leq c_2\|x_k - x_*\|^2 \quad e \quad \frac{\|P_N(x_k - x_*)\|}{\|P_N(x_{k-1} - x_*)\|} \longrightarrow \frac{1}{2}.$$

Para a segunda extensão seja x_* um zero de ordem maior que dois de F e suponha que F tenha derivadas contínuas até ordem $p + 2$, $p \geq 1$ e que

$$F^j(x_*) = 0, \quad j = 2, \dots, p, \quad (2.8)$$

$$\|F^{p+1}(x_*)sx^p\| \geq c\|s\|\|x^p\|, \quad s \in N, \quad x \in Y, \quad (2.9)$$

onde $F^{p+1}(x_*)mx^p = F^{p+1}(x_*)(m, x, x, \dots, x)$.

Teorema 2.9. *Suponha que:*

1. $F \in C^{p+2}$;
2. $\dim N = 1$;
3. $F''(x_*)NN \cap X = \{0\}$;
4. As condições (2.8) e (2.9) sejam satisfeitas.

Então, existem constantes ρ, θ e c_1 positivas tais que $F'(x)$ é inversível em $W_{\rho,\theta}(x_*) = B_\rho(x_*) \cap C_\theta(x_*)$, $x \neq x_*$, satisfazendo $\|(F'(x))^{-1}\| \leq c_1\|x - x_*\|^{-1}$, a aplicação $G(x) = x - (F'(x))^{-1}F(x)$ leva $W_{\rho,\theta}(x_*)$ nele próprio. Além disso, definindo $x_k = G(x_{k-1})$ com $x_0 \in W_{\rho,\theta}(x_*)$, existe uma constante positiva c_2 tal que:

$$x_k \longrightarrow x_*, \quad \|P_X(x_k - x_*)\| \leq c_2\|x_k - x_*\|^2 \quad e \quad \frac{\|P_N(x_k - x_*)\|}{\|P_N(x_{k-1} - x_*)\|} \longrightarrow \frac{p}{p+1}.$$

Decker e Kelley em [37] também obtiveram uma extensão do resultado de Reddien [141] para o caso em que a dimensão do núcleo de $F'(x_*)$ é maior que 1 e finita. O resultado de [37] baseia-se na não singularidade do seguinte operador $B(z)$ de N em N , onde $z \in Y$:

$$B(z) = -P_N F''(x_*)(z, P_N(\cdot)). \quad (2.10)$$

O resultado obtido em [37] é o seguinte:

Teorema 2.10. *As conclusões do Teorema 2.7 continuam válidas sob as seguintes hipóteses:*

1. $F \in C^3$;
2. $\dim N < \infty$;
3. para todo $z \in X$, $B(z)$ é não singular em N .

Uma observação feita por Decker e Kelley [37] é que se $\dim N = 1$, então as Hipóteses 3 e 4 do Teorema 2.7 implicam na Hipótese 3 do Teorema 2.10.

Em outro trabalho, [38], que foi uma sequência de [37], Decker e Kelley estenderam os resultados de convergência para singularidades de segunda ordem. Para obter tal resultado foram utilizados os seguintes operadores de X em X :

$$B_1(\tilde{x}) = F''(x_*)(P_N(\tilde{x}), (\cdot)), \quad (2.11)$$

$$B_2(\tilde{x}) = F'''(x_*)(P_N(\tilde{x}), P_N(\tilde{x}), (\cdot)), \quad (2.12)$$

onde $\tilde{x} = x_* - x$, com $x \in X$.

Também são utilizadas as seguintes aplicações lineares \overline{B}_j de N em N :

$$\overline{B}_j(\tilde{x}) = P_N B_j(\tilde{x}) P_N, \quad j = 1, 2. \quad (2.13)$$

\hat{F} denotará $F'(x_*)|_X$.

Para ρ, θ e K reais positivos é definido o seguinte conjunto:

$$W(\rho, \theta, K) = \{x \mid 0 < \|\tilde{x}\| < \rho, \|P_N F''(x_*)(P_X(\tilde{x}), \phi)\| \leq \theta \|P_N(\tilde{x})\|^2 \|\phi\| \\ \text{para todo } \phi \in N, \|P_X(\tilde{x})\| \leq K \|P_N(\tilde{x})\|^2\}. \quad (2.14)$$

Teorema 2.11. *Suponha a validade das seguintes hipóteses:*

1. $F \in C^4$;
2. $\bar{B}_1(\tilde{x}) = 0$;
3. Para $x \in X$ tal que $P_N(\tilde{x}) \neq 0$, $\bar{B}_2(\tilde{x})$ é uma aplicação não singular em N ;
4. Para cada $x, y, z \in X$, $P_N B_1(\tilde{z}) P_X \hat{F}^{-1} P_X B_1(\tilde{y}) P_N(x) = 0$.

Então, se

$$K_0 = 4 \sup_{\phi, \psi \in N} \left(\frac{1}{\|\phi\| \|\psi\|} \|P_X \hat{F}^{-1} P_X F''(x_*)(\phi, \psi)\| \right) + 1$$

e ρ e θ são suficientemente pequenos, $F'(x)$ é não singular para $x \in W(\rho, \theta, K_0)$, a aplicação $G(x) = x - (F'(x))^{-1} F(x)$ leva $W(\rho, \theta, K_0)$ nele próprio e existe $c_1 > 0$ tal que $\|(F'(x))^{-1}\| \leq c_1 \|\tilde{x}\|^{-2}$ para todo $x \in W(\rho, \theta, K_0)$.

Além disso, se $x_0 \in W(\rho, \theta, K_0)$ e $x_k = G(x_{k-1})$, para $k \geq 1$, a sequência $\{x_k\}$ converge para x_* e valem:

$$\text{para } k \geq 1, \|P_X(\tilde{x}_k)\| \leq \frac{1}{3} K_0 \|\tilde{x}_{k-1}\|^2, \quad \lim_{k \rightarrow \infty} \frac{\|P_N(\tilde{x}_k)\|}{\|P_N(\tilde{x}_{k-1})\|} = \frac{2}{3}.$$

e x_* é a única solução de $F(x) = 0$ no conjunto $\{\tilde{x} \mid \|\tilde{x}\| \leq \rho\}$.

O resultado obtido é basicamente o mesmo do Teorema 2.9 de Reddien [142]. Griewank e Osborne em [65] também obtiveram um resultado para o método de Newton quando a dimensão do núcleo do Jacobiano é maior que um, com condições um pouco mais fracas que as utilizadas em [37], e também levemente mais gerais que os resultados obtidos em [142]. No entanto, podemos dizer que os resultados de [37, 65, 142] são basicamente os mesmos.

O operador $B(z)$ utilizado em [65] é definido por:

$$B(z) = Q_N F''(x_*)(z, P_N(\cdot)), \quad (2.15)$$

onde P_N e Q_N são os projetores ortogonais nos núcleos de $F'(x_*)$ e $F'(x_*)^T$, respectivamente, e satisfazem,

$$Q_N F'(x_*) = F'(x_*) P_N = 0.$$

Analogamente aos resultados de [37, 142], uma hipótese básica para obter a convergência do método de Newton é a inversibilidade de $B(z)$.

Estes foram os primeiros e principais trabalhos sobre convergência do método de Newton em problemas singulares. Posteriormente houve diversos trabalhos que continuaram a tratar do assunto e também propuseram maneiras para acelerar a convergência do método de Newton, por exemplo [36, 39, 64, 66, 96, 152]. Oberlin e Wright em [126] enfraqueceram hipóteses de Griewank [63] e estenderam o resultados para resolução de sistemas singulares não suaves. Além disso, relacionaram as várias hipóteses de inversibilidade sobre os operadores $B(z)$ definidos anteriormente com a 2-regularidade introduzida por Izmailov e Solodov em [81] e estudada em [34, 82, 83].

Também surgiram alternativas ao método de Newton. Por exemplo, em [40] os autores tratam de um método de Broyden, em [59] os autores introduzem uma modificação do algoritmo de Brown para resolver sistemas singulares, em [58] os autores construíram um algoritmo ABS (introduzido por Abaffy, Broyden e Spedicato) para resolver problemas singulares, e também existe o método tensorial introduzido por Schnabel e Frank em [151], que será visto com mais detalhes na próxima seção. Outros trabalhos de interesse sobre o assunto são [53, 57, 63].

2.2 Métodos tensoriais para sistemas não lineares

Não é de hoje que métodos tensoriais para resolver sistemas, ou problemas de otimização, são estudados, [84, 85, 89, 127, 149, 151]. O método de Halley, por exemplo, data de 1694 e é apontado como um dos métodos tensoriais com convergência cúbica mais

estudados e redescobertos da história. Nesta seção vamos nos ater à discussão e resultados sobre uma classe de métodos tensoriais de terceira ordem conhecidos como família Chebyshev-Halley [69] ou classe dos métodos Halley [68], e também ao método tensorial introduzido por Schnabel & Frank em [151].

2.2.1 Família Chebyshev-Halley

Esta família de métodos foi inicialmente introduzida por Hernández e Salanova em [75] para equações escalares. Posteriormente Gutiérrez e Hernández em [69] estenderam os resultados para equações em espaços de Banach. Esta família de métodos reúne sob uma mesma relação os mais conhecidos métodos de terceira ordem para resolver sistemas não lineares:

- o método de Halley, também conhecido como modificação de Halley do método de Newton, ou ainda método das hipérbolas tangentes. Para estudos mais aprofundados sobre este método ver [1, 21, 148] para o caso em uma variável real, e [23, 45, 46, 77] em espaços de Banach;
- o método de Chebyshev, também conhecido como método das parábolas tangentes, ou ainda método de Halley melhorado. Para resultados mais detalhados sobre este método ver [24, 15, 76, 78];
- o método super-Halley, também conhecido como aceleração convexa do método de Newton. Para resultados mais detalhados sobre este método ver [27, 70, 71].

Notação: No decorrer do texto usaremos as seguintes notações:

$$F_k = F(x_k), \quad F'_k = F'(x_k), \quad F''_k = F''(x_k),$$

$$F_* = F(x_*), \quad F'_* = F'(x_*), \quad F''_* = F''(x_*).$$

Para motivar a introdução dos métodos tensoriais considere a aproximação de Taylor de segunda ordem para F :

$$F(x+h) \approx T_2(x+h) = F(x) + F'(x)h + \frac{1}{2}F''(x)hh.$$

Para obtermos $F(x + h) = 0$ fazemos $T_2(x + h) = 0$ e $h = d_N + d$, assim:

$$\underbrace{F(x) + F'(x)d_N + F'(x)d}_{=0, \text{Newton}} + \frac{1}{2}F''(x)\left[\underbrace{d_N d_N}_{(1)} + \underbrace{2d_N d}_{(2)} + \underbrace{dd}_{(3)}\right] = 0.$$

Ignorando os termos (2) e (3), obtemos o método de Chebyshev.

Ignorando o termo (2) e substituindo um d de (3) por d_N , obtemos o método de Halley.

Ignorando o termo (3), obtemos o método Super-Halley.

Conforme estabelecido em [69], os métodos da família Chebyshev-Halley são definidos pela seguinte relação:

$$x_{k+1} = x_k - \left[I + \frac{1}{2}(F'_k)^{-1}F''_k(F'_k)^{-1}F_k \left[I - \lambda(F'_k)^{-1}F''_k(F'_k)^{-1}F_k \right]^{-1} \right] (F'_k)^{-1}F_k, \quad (2.16)$$

na qual $\lambda \in [0, 1]$ é o parâmetro que define cada membro particular desta família.

Para $\lambda = 0$ temos o método de Chebyshev, para $\lambda = \frac{1}{2}$, o método de Halley e para $\lambda = 1$, o método super-Halley.

Os resultados clássicos de convergência local para tais métodos, convergência quadrática para Newton e cúbica para a família Chebyshev-Halley, são válidos somente quando $\|(F'(x))^{-1}\|$ pode ser cotada numa vizinhança de uma solução x_* . De um modo mais direto, quando $F'(x)$ é inversível numa vizinhança de x_* tem-se o resultado desejado. Quando não existe tal vizinhança não se tem garantia de boas propriedades de convergência local para tais métodos.

Gundersen e Steihaug, em [68], mostraram que esta classe de métodos pode ser definida do seguinte modo:

$$x_{k+1} = x_k + d_N + d, \quad (2.17)$$

no qual d_N e d são soluções dos seguintes sistemas lineares:

$$F'_k d_N = -F_k, \quad (F'_k + \lambda F''_k d_N) d = -\frac{1}{2} F''_k d_N d_N, \quad (2.18)$$

com $\lambda \in [0, 1]$.

Vamos utilizar os subscritos C, H e sH em d , nas equações (2.17)-(2.18), para diferenciar os métodos de Chebyshev ($\lambda = 0$), Halley ($\lambda = \frac{1}{2}$) e super-Halley ($\lambda = 1$), respectivamente.

Em [68] os autores mostraram que, para determinadas classes de problemas, a razão do custo por iteração entre um método qualquer da família Chebyshev-Halley e o método de Newton independe da dimensão do problema, sendo assintoticamente constante.

O modo de definir a família Chebyshev-Halley pelas equações (2.17) e (2.18) nos serve de inspiração para classificá-los como métodos do tipo preditor-corretor, sendo a direção preditora, a direção de Newton d_N , e a direção corretora, a direção tensorial d .

Também tem surgido variações desses métodos para resolver equações ou sistemas, por exemplo:

- Deng e Zhang em [41] estudaram um método das hipérboles tangentes melhorado (método de Chebyshev) inexato para resolver problemas de otimização irrestrita;
- modificações para acelerar a convergência de métodos de terceira ordem para equações [108, 109, 110, 139];
- variações de métodos de terceira ordem para equações sem cálculo de derivadas [29, 107].

2.2.2 Método tensorial de Schnabel-Frank

O método tensorial introduzido por Schnabel e Frank em [151], também estudado em [17, 49], é diferente dos métodos da família Chebyshev-Halley, pois nele é necessário resolver um modelo quadrático de F em cada iteração e não dois sistemas lineares como nos métodos Chebyshev-Halley.

O principal destaque desse método foi a melhora proposta para resolver problemas com Jacobianos singulares ou mal-condicionados, em comparação ao método de Newton. Apesar de ter que resolver um modelo quadrático de F em cada iteração, esse

método não utiliza as derivadas de segunda ordem de F . O termo tensorial introduzido no modelo é uma aproximação do termo correto, na qual o tensor de terceira ordem tem posto baixo. Essa característica do modelo faz com que seu cálculo e mesmo sua resolução tenham um custo não maior que o método de Newton.

Este método foi posteriormente aplicado para resolver problemas de quadrados mínimos não lineares esparsos de grande porte [20], em problemas de otimização sem restrições [19, 150], e também em problemas de otimização com restrições de igualdade [50].

Vejamos alguns detalhes do método tensorial de Schnabel-Frank:

Modelo Tensorial

O modelo tensorial de segunda ordem considerado é dado por:

$$M_T(x_c + d) = F_c + F'_c d + \frac{1}{2} T_c d d, \quad (2.19)$$

no qual o subscrito c denota o ponto corrente.

Devem ser escolhidos p pontos anteriores $\{x_{-1}, x_{-2}, \dots, x_{-p}\}$, com $p \leq \sqrt{n}$ e forma-se o seguinte sistema:

$$F_{-k} = F_c + F'_c s_k + \frac{1}{2} T_c s_k s_k, \quad (2.20)$$

$$s_k = x_{-k} - x_c, \quad k = 1, \dots, p. \quad (2.21)$$

O conjunto $\{s_k\}_{k=1}^p$ deve ser linearmente independente (LI), na prática fortemente LI, isto é, o ângulo entre um elemento s_k e o subespaço gerado pelos elementos restantes deve ser maior ou igual a um θ fixado. Na prática, os autores verificaram que θ entre 20 e 45 graus é uma escolha apropriada.

O tensor T_c deve satisfazer às seguintes relações:

$$T_c s_k s_k = z_k, \quad (2.22)$$

$$z_k = 2(F_{-k} - F_c - F'_c s_k). \quad (2.23)$$

Tem-se np equações e n^3 incógnitas, $T_c(i, j, k)$.

Escolhe-se T_c como a solução do seguinte problema de minimização:

$$\begin{aligned} \min \quad & \|T_c\|_F \\ \text{s.a.} \quad & T_c s_k s_k = z_k, \quad k = 1, \dots, p. \end{aligned}$$

A solução deste problema é dada por:

$$T_c = \sum_{k=1}^p a_k s_k s_k,$$

onde a_k é a k -ésima coluna de $A = ZM^{-1}$, e:

$$Z = [z_1, \dots, z_p] \in \mathbb{R}^{n \times p}, \quad M \in \mathbb{R}^{p \times p}, \quad \text{com } M_{ij} = (s_i^T s_j)^2.$$

De posse desses resultados sobre o método de Newton para problemas singulares e sobre métodos tensoriais no Capítulo 3 estendemos o resultado de Cavanagh para os métodos de Chebyshev-Halley.

Capítulo 3

Família Chebyshev-Halley em Sistemas Não Lineares Singulares

Neste capítulo temos por objetivo estender a demonstração do resultado de Cavanagh [26], sobre a convergência do método de Newton para problemas singulares, para alguns métodos tensoriais (métodos de Chebyshev, Halley e Super-Halley). Apresentamos também uma justificativa para a utilização de tais métodos em problemas singulares ao invés do método de Newton, seguindo o observado por Schnabel e Frank [151].

Cavanagh [26] mostrou que, sob certas condições, o método de Newton gera uma sequência linearmente convergente para uma solução x_* do problema (2.1) quando $F'(x_*)$ não é inversível. Vamos analisar o resultado de Cavanagh e estendê-lo para os métodos tensoriais citados anteriormente.

3.1 Resultado de Cavanagh revisitado

Para prosseguir a análise vamos rever algumas notações definidas no Capítulo 1: Seja x_* uma solução de (2.1) tal que $F'(x_*)$ não seja inversível e suponha que $F'(x)$ seja inversível em

$$\mathcal{V}_\delta = \{x \mid 0 < \|x - x_*\| < \delta\} \tag{3.1}$$

para algum $\delta > 0$. Sejam também N o núcleo de $F'(x_*)$, X tal que $\mathbb{R}^n = N \oplus X$, P_N o projetor em N paralelo a X e $P_X = I - P_N$ o projetor em X paralelo a N .

Observação 3.1. *Note que X não é necessariamente o complemento ortogonal de N , apenas um subespaço que completa N como soma direta para o \mathbb{R}^n .*

O método de Newton para resolver (2.1) é definido iterativamente por:

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

onde d_N é solução de:

$$F'_k d_N = -F_k.$$

De outro modo:

$$x_{k+1} = G(x_k),$$

onde

$$G(x) = \begin{cases} x - (F'(x))^{-1}F(x), & \text{se } x \neq x_* \\ x_*, & \text{se } x = x_* \end{cases} \quad (3.2)$$

é chamada função de iteração do método de Newton. Observe que no caso singular a primeira expressão para G não vale em x_* , por isso definimos $G(x_*) = x_*$, posteriormente mostraremos que desse modo a aplicação G é contínua.

O seguinte resultado refere-se a propriedades da função G e sua utilidade ficará clara mais a frente.

Proposição 3.2. *Suponha que $F \in C^2$. Para todo $x \in \mathcal{V}_\delta$, a aplicação $G(x)$ definida em (3.2) é diferenciável e $G'(x) = (F'(x))^{-1}F''(x)(F'(x))^{-1}F(x)$.*

Demonstração: Para $x \in \mathcal{V}_\delta$ temos $G(x) = x - (F'(x))^{-1}F(x)$, então

$$F'(x)G(x) = F'(x)x - F(x).$$

Derivando com relação a x , obtemos:

$$F''(x)G(x) + F'(x)G'(x) = F''(x)x + F'(x)I - F'(x) = F''(x)x.$$

Assim, substituindo a expressão (3.2) e simplificando vem:

$$F'(x)G'(x) = F''(x)x - F''(x)G(x) = F''(x)(F'(x))^{-1}F(x).$$

Portanto,

$$G'(x) = (F'(x))^{-1}F''(x)(F'(x))^{-1}F(x).$$

□

O lema seguinte será utilizado, juntamente com o Lema 3.4, para justificar a utilização de uma das hipóteses adotadas na sequência do trabalho.

Lema 3.3. *Seja $M \in \mathbb{R}^{n \times n}$ uma matriz cujo núcleo da transformação linear associada é N e seja X um subespaço de \mathbb{R}^n tal que $N \oplus X = \mathbb{R}^n$, então existe $T \in \mathbb{R}^{n \times n}$ inversível tal que $TMX = X$.*

Demonstração: Sejam $\{x_1, x_2, \dots, x_p\}$ uma base para X , $\{x_{p+1}, x_{p+2}, \dots, x_n\}$ uma base para N e defina $\bar{x}_i = Mx_i$, para $i = 1, 2, \dots, p$.

O conjunto $B = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p\}$ é linearmente independente.

De fato, considere a seguinte combinação linear:

$$\alpha_1 \bar{x}_1 + \alpha_2 \bar{x}_2 + \dots + \alpha_p \bar{x}_p = 0. \tag{3.3}$$

Mostraremos que $\alpha_i = 0$, para $i = 1, 2, \dots, p$.

Multiplicando (3.3) à esquerda por M obtemos

$$\begin{aligned} 0 &= M(\alpha_1 \bar{x}_1 + \alpha_2 \bar{x}_2 + \dots + \alpha_p \bar{x}_p) \\ &= \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_p x_p. \end{aligned}$$

Como $\{x_1, x_2, \dots, x_p\}$ é base para X , segue que $\alpha_i = 0$, para $i = 1, 2, \dots, p$, e a independência linear de B está provada.

Agora podemos completar B a uma base de \mathbb{R}^n , digamos

$$\{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}.$$

Definindo T por $T(\bar{x}_i) = x_i$, para $i = 1, 2, \dots, n$, claramente temos que T é inversível e $TMX = X$. □

O resultado a seguir, sobre a invariância dos iterandos do método de Newton por transformações não singulares, é bem conhecido e de simples demonstração, mas damos ênfase a ele pois posteriormente mostraremos a sua validade para os métodos Chebyshev-Halley. Também será útil para justificar uma das hipóteses adotadas na sequência do trabalho.

Lema 3.4. *Os iterandos gerados pelo método de Newton são invariantes por transformações não singulares.*

Demonstração: De fato, se T é uma matriz não singular, então:

$$\begin{aligned} x_{k+1} &= x_k - [(TF_k)']^{-1}(TF_k) = x_k - (TF_k')^{-1}(TF_k) \\ &= x_k - (F_k')^{-1}T^{-1}TF_k = x_k - (F_k')^{-1}F_k. \end{aligned}$$

□

As seguintes hipóteses para $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ serão adotadas nos próximos resultados:

H1: A função $F \in C^3$;

H2: O sistema $F(x) = 0$ tem uma solução x_* ;

H3: Existe $\delta > 0$ tal que $F'(x)$ é não singular na vizinhança \mathcal{V}_δ , definida em (3.1);

H4: Existe uma constante c positiva tal que para todo $s \in N$ e $x \in \mathbb{R}^n$ vale a desigualdade:

$$\|F''(x_*)sx\| \geq c\|s\| \cdot \|x\|.$$

H5: $F'(x_*)X = X$, isto é, o subespaço X é invariante pela ação do operador $F'(x_*)$.

Observação 3.5. *A Hipótese H1 é mais forte que a hipótese original de Cavanagh (ver Teorema 2.3), essa modificação é utilizada para simplificar as demonstrações. A hipótese H5 parece ser restritiva, mas observe que se $F'(x_*)X \neq X$, então pelo Lema 3.3 podemos obter uma matriz inversível T tal que $TF'(x_*)X = X$. Desse modo, pelo Lema 3.4, os iterandos gerados pelo método de Newton aplicado a F e TF serão os mesmos.*

O teorema de convergência será precedido por alguns lemas técnicos, que foram extraídos do desenvolvimento da prova de Cavanagh para simplificar a apresentação dos resultados.

O lema a seguir serve como base para a demonstração de um outro lema, enunciado posteriormente, o qual estabelece que mesmo F'_* não sendo inversível é possível obter um limitante para $(F'(x))^{-1}$ quando x está próximo de x_* .

Lema 3.6. *Suponha a validade de H1, H2, H3, H4 e H5. Então*

i) Existem $c_1 > 0$ e $\delta > 0$ tais que $\|F'(x)y\| \geq c_1\|y\|$ para todo $y \in X$ e $x \in \mathcal{V}_\delta$.

ii) Existem $c_2 > 0$ e $\delta > 0$ tais que $\|F'(x)s\| \geq c_2\|x - x_\|\|s\|$ para todo $s \in N$ e $x \in \mathcal{V}_\delta$.*

Demonstração: A aproximação de Taylor de $F'(x)$ em torno de x_* é um dos ingredientes básicos desta prova.

Para mostrar (i) considere, para $x \in \mathcal{V}_\delta$, a expansão:

$$F'(x) = F'(x_*) + O(\|x - x_*\|).$$

Assim, para $y \in X$,

$$F'(x)y = F'(x_*)y + O(\|x - x_*\|\|y\|).$$

Logo,

$$\|F'(x)y\| \geq \|F'(x_*)y\|$$

para algum $\delta > 0$.

Agora suponha que não exista $c_1 > 0$ tal que

$$\|F'(x_*)y\| \geq c_1\|y\|. \tag{3.4}$$

Tome uma sequência $\{y_n\} \subset X$ com $\|y_n\| = 1 \forall n$ e $\lim_{n \rightarrow \infty} y_n = y_*$. Sendo X fechado, então $y_* \in X$, assim, por H5, $F'(x_*)y_* \in X$. E ainda temos que $\|y_*\| = 1$. Logo, $F'(x_*)y_* \neq 0$.

Pela continuidade de $F'(x_*)$ segue que

$$F'(x_*)y_n \longrightarrow F'(x_*)y_*. \tag{3.5}$$

Pela suposição de que não existe $c_1 > 0$ tal que (3.4) seja válida, temos

$$\|F'(x_*)y_n\| \longrightarrow 0. \quad (3.6)$$

De (3.5) e (3.6) tem-se $F'(x_*)y_* = 0$, o que é uma contradição ao fato de $F'(x_*)y_* \neq 0$.

Portanto, deve existir $c_1 > 0$ tal que (3.4) seja válida para algum $\delta > 0$.

Para mostrar (ii) considere, para $x \in \mathcal{V}_\delta$, a expansão:

$$F'(x) = F'(x_*) + F''(x_*)(x - x_*) + O(\|x - x_*\|^2).$$

Se $s \in N$ então

$$\begin{aligned} F'(x)s &= F'(x_*)s + F''(x_*)(x - x_*)s + O(\|x - x_*\|^2\|s\|) \\ &= F''(x_*)(x - x_*)s + O(\|x - x_*\|^2\|s\|). \end{aligned}$$

Assim, pela hipótese H4, existe $c_2 > 0$ tal que:

$$\|F'(x)s\| \geq \|F''(x_*)(x - x_*)s\| \geq c_2\|x - x_*\|\|s\|,$$

para algum $\delta > 0$. □

Lema 3.7. *Suponha a validade de H1, H2, H3, H4 e H5. Então, existem $c_3 > 0$ e $\delta > 0$ tais que para todo $x \in \mathcal{V}_\delta$*

$$\|(F'(x))^{-1}\| \leq \frac{1}{c_3\|x - x_*\|}.$$

Demonstração: Sejam $x \in \mathcal{V}_\delta$, $R_x \equiv F'(x)(X)$ e $S_x \equiv F'(x)(N)$. Se $x \neq x_*$, então $\mathbb{R}^n = R_x \oplus S_x$.

Com efeito, como $F'(x)$ é não singular e $\mathbb{R}^n = X \oplus N$, então

$$\mathbb{R}^n = F'(x)(X) = F'(x)(X \oplus N) = F'(x)(X) \oplus F'(x)(N).$$

Do item (i) do Lema 3.6 segue que:

$$\frac{\|(F'(x))^{-1}u\|}{\|u\|} \leq \frac{1}{c_1},$$

3.1. RESULTADO DE CAVANAGH REVISITADO

para todo $u \in R_x$ e $x \in \mathcal{V}_\delta$, $u \neq 0$.

Já do item (ii) obtemos:

$$\frac{\|(F'(x))^{-1}v\|}{\|v\|} \leq \frac{1}{c_2\|x - x_*\|},$$

para todo $v \in S_x$ e $x \in \mathcal{V}_\delta$, $v \neq 0$.

Consequentemente, existem c_3 e δ positivos tais que para todo $x \in \mathcal{V}_\delta$

$$\|(F'(x))^{-1}\| \leq \frac{1}{c_3\|x - x_*\|}.$$

□

Observação 3.8. *Os Lemas 3.6 e 3.7 também foram utilizados por Reddien em [141, 142].*

O teorema a seguir é o resultado de Cavanagh [26, Teorema 2.7] com leves modificações. A hipótese H1 que estamos utilizando é mais forte que a hipótese original de [26] na qual $F \in C^2$ e $F''(x)$ possui módulo de continuidade conhecido. Essa mudança é feita apenas para simplificar as demonstrações. Também vale observar que a correção sugerida por Cavanagh para restaurar a convergência quadrática é apenas teórica e não pode ser utilizada na prática, pois necessita de prévio conhecimento da solução.

Teorema 3.9. *Suponha a validade de H1, H2, H3, H4 e H5, e escolha um $\delta > 0$ de modo que os Lemas 3.6 e 3.7 sejam válidos.*

Então, x_ é um ponto atrator para o método de Newton e a convergência é linear. E mais,*

$$P_X(x_k - x_*) \longrightarrow 0 \text{ superlinearmente, } \frac{\|P_N(x_k - x_*)\|}{\|P_N(x_{k-1} - x_*)\|} \longrightarrow \frac{1}{2} \text{ e } \frac{1}{2}P_N = G'(x_*).$$

E a convergência quadrática pode ser restaurada usando o seguinte processo iterativo:

$$x_{k+1} = x_k + d_N - \frac{1}{2}P_N(x_k - x_*).$$

Demonstração: Seja $x_0 \in \mathcal{V}_\delta$, pelo método de Newton temos:

$$x_{k+1} = x_k - (F'_k)^{-1}F_k.$$

Assim,

$$\begin{aligned} x_{k+1} - x_* &= x_k - x_* - (F'_k)^{-1}F(x_k) \\ &= x_k - x_* - (F'_k)^{-1}[F'_*(x_k - x_*) + O(\|x_k - x_*\|^2)] \\ &= x_k - x_* - (F'_k)^{-1}F'_*(P_X + P_N)(x_k - x_*) - (F'_k)^{-1}O(\|x_k - x_*\|^2) \\ &= x_k - x_* - (F'_k)^{-1}F'_*P_X(x_k - x_*) - (F'_k)^{-1}O(\|x_k - x_*\|^2). \end{aligned}$$

Logo, pelo Lema 3.6 (i), e pelo Lema 3.7, para $x_k \in \mathcal{V}_\delta$, existem $c_1, c_3, c_4, c_5 > 0$ tais que

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \|x_k - x_*\| + (F'_k)^{-1}F'_*P_X(x_k - x_*)\| + \|(F'_k)^{-1}\|c_4\|x_k - x_*\|^2 \\ &\leq \|x_k - x_*\| + \frac{1}{c_1}\|x_k - x_*\| + \frac{c_4\|x_k - x_*\|^2}{c_3\|x_k - x_*\|} \leq c_5\|x_k - x_*\|. \end{aligned}$$

Obtemos assim a convergência linear do método de Newton no caso singular, ou seja, $\lim_{k \rightarrow \infty} x_k = x_*$, e então G torna-se contínua em $\mathcal{V}_\delta \cup \{x_*\}$.

Vamos mostrar agora que $G'(x_*) = \frac{1}{2}P_N$.

De fato, usando a expansão de Taylor de segunda ordem para F e o Lema 3.6 (i), para $x_k \in \mathcal{V}_\delta$ temos:

$$\begin{aligned} \left\| G(x_k) - G(x_*) - \frac{1}{2}P_N(x_k - x_*) \right\| &= \left\| x_k - x_* - (F'_k)^{-1}F_k - \frac{1}{2}P_N(x_k - x_*) \right\| \\ &\leq \left\| \frac{1}{2}(F'_k)^{-1}F''_k(x_k - x_*) - \frac{1}{2}P_N \right\| \|x_k - x_*\| \\ &\quad + O(\|x_k - x_*\|^2). \end{aligned}$$

Como

$$(F'_k)^{-1}F''_k(x_k - x_*) = (F'_k)^{-1}[F'_k - F'_*] + O(\|x_k - x_*\|).$$

3.1. RESULTADO DE CAVANAGH REVISITADO

segue que, para $s \in N$,

$$(F'_k)^{-1}F''_k(x_k - x_*)s = s + O(\|x_k - x_*\|\|s\|),$$

e para $y \in X$,

$$\begin{aligned} (F'_k)^{-1}F''_k(x_k - x_*)y &= y + (F'_k)^{-1}F'_*y + O(\|x_k - x_*\|\|y\|) \\ &= O(\|x_k - x_*\|\|y\|). \end{aligned}$$

Consequentemente,

$$\left\| \frac{1}{2}(F'_k)^{-1}F''_k(x_k - x_*) - \frac{1}{2}P_N \right\| = O(\|x_k - x_*\|).$$

Assim,

$$\left\| G(x_k) - G(x_*) - \frac{1}{2}P_N(x_k - x_*) \right\| = O(\|x_k - x_*\|^2),$$

ou seja, $G'(x_*) = \frac{1}{2}P_N$.

Temos portanto a estimativa:

$$G(x_k) = G(x_*) + \frac{1}{2}P_N(x_k - x_*) + O(\|x_k - x_*\|^2).$$

Então para restaurar a convergência quadrática basta adotar o processo iterativo:

$$x_{k+1} = x_k + d_N - \frac{1}{2}P_N(x_k - x_*). \quad (3.7)$$

□

O seguinte resultado é consequência direta do Teorema 3.9:

Corolário 3.10. *A aplicação G como definida em (3.2) é contínua em $\mathcal{V}_\delta \cup \{x_*\}$.*

Corolário 3.11. *O operador $G'(x)$ definido por:*

$$G'(x) = \begin{cases} (F'(x))^{-1}F''(x)(F'(x))^{-1}F(x), & \text{se } x \neq x_* \\ \frac{1}{2}P_N, & \text{se } x = x_*. \end{cases}$$

é contínuo em $\mathcal{V}_\delta \cup \{x_\}$.*

Demonstração: O único ponto onde pode-se ter problemas é x_* , então vamos mostrar que $G'(x) \rightarrow \frac{1}{2}P_N$ quando $x \rightarrow x_*$. De fato,

$$\left\| \frac{1}{2}(F'(x))^{-1}F''(x)(x - x_*) - \frac{1}{2}P_N \right\| = O(\|x - x_*\|).$$

E como

$$G'(x) = (F'(x))^{-1}F''(x)(F'(x))^{-1}F(x) = \frac{1}{2}(F'(x))^{-1}F''(x)(x - x_*) + O(\|x - x_*\|),$$

segue que,

$$\left\| G'(x) - \frac{1}{2}P_N \right\| = O(\|x - x_*\|).$$

Portanto, $G'(x) \rightarrow \frac{1}{2}P_N$ quando $x \rightarrow x_*$. □

3.2 Métodos tensoriais em problemas singulares

Consideremos agora o método de Chebyshev definido pelas equações (2.17) e (2.18) com $\lambda = 0$. Mostraremos que o Teorema 3.9 vale também para (2.17)-(2.18), ou seja, o método de Chebyshev gera uma sequência convergente para uma solução do problema (2.1). A ordem de convergência também é linear, mas o passo corretor nos dá uma melhora com relação ao método de Newton. Mostraremos que a direção corretora tem uma estreita relação com a correção sugerida por Cavanagh [26]. Tal relação explica, parcialmente, a afirmação de Schnabel e Frank [151] de que os métodos tensoriais levam vantagem sobre o método de Newton nos casos singulares. As hipóteses são as mesmas do Teorema 3.9.

3.2.1 Adaptação para o método de Chebyshev

Vejam inicialmente que o resultado do Lema 3.4 também é válido para o método de Chebyshev, o que justifica a utilização da hipótese H5.

Lema 3.12. *Os iterandos gerados pelo método de Chebyshev, equações (2.17)-(2.18) com $\lambda = 0$, são invariantes por transformações não singulares.*

Demonstração: De fato, se T é uma matriz não singular, então:

$$\begin{aligned}
 x_{k+1} &= x_k - [(TF_k)']^{-1}(TF_k) - \frac{1}{2}[(TF_k)']^{-1}(TF_k'')([(TF_k)']^{-1}(TF_k))([(TF_k)']^{-1}(TF_k)) \\
 &= x_k - (TF_k')^{-1}(TF_k) - \frac{1}{2}(TF_k')^{-1}(TF_k'')((TF_k')^{-1}(TF_k))((TF_k')^{-1}(TF_k)) \\
 &= x_k - (F_k')^{-1}T^{-1}TF_k - \frac{1}{2}(F_k')^{-1}T^{-1}(TF_k'')((F_k')^{-1}T^{-1}(TF_k))((F_k')^{-1}T^{-1}(TF_k)) \\
 &= x_k - (F_k')^{-1}F_k - \frac{1}{2}(F_k')^{-1}F_k''((F_k')^{-1}F_k)((F_k')^{-1}F_k).
 \end{aligned}$$

□

Teorema 3.13. *Suponha a validade das hipóteses $H1$, $H2$, $H3$, $H4$ e $H5$ para $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Então, uma sequência gerada pelo método de Chebyshev converge para x_* e $d_C \longrightarrow \frac{1}{4}P_N(x_k - x_*)$, onde d_C é a solução de $F_k'd = -\frac{1}{2}F_k''d_Nd_N$.*

Demonstração: Da relação (2.17) e da expansão de Taylor de segunda ordem para F em torno de $x_k \in \mathcal{V}_\delta$ vem:

$$\begin{aligned}
 x_{k+1} - x_* &= x_k - x_* - (F_k')^{-1}F_k - \frac{1}{2}(F_k')^{-1}F_k''d_Nd_N \\
 &= (F_k')^{-1} \left[-F_k + F_k'(x_k - x_*) - \frac{1}{2}F_k''d_Nd_N \right] \\
 &= (F_k')^{-1} \left[\frac{1}{2}F_k''(x_k - x_*)(x_k - x_*) + O(\|x_k - x_*\|^3) - \frac{1}{2}F_k''d_Nd_N \right] \\
 &= (F_k')^{-1} \left[\frac{1}{2}F_k''(x_k - x_* + d_N)(x_k - x_* - d_N) + O(\|x_k - x_*\|^3) \right]. \quad (3.8)
 \end{aligned}$$

Veamos estimativas para os termos que aparecem na equação (3.8):

- $x_k - x_* + d_N = O(\|x_k - x_*\|)$

De fato, pela expansão de Taylor temos:

$$F_* = F_k + F_k'(x_* - x_k) + O(\|x_k - x_*\|^2).$$

Como $F_* = 0$, segue que:

$$-F_k + F_k'(x_k - x_*) = O(\|x_k - x_*\|^2).$$

Assim, do Lema 3.7 vem:

$$\begin{aligned}
 x_k - x_* + d_N &= x_k - x_* - (F'_k)^{-1}F_k \\
 &= (F'_k)^{-1}[-F_k + F'_k(x_k - x_*)] \\
 &= (F'_k)^{-1}O(\|x_k - x_*\|^2) \\
 &= O(\|x_k - x_*\|).
 \end{aligned}$$

- $x_k - x_* - d_N = O(\|x_k - x_*\|)$

Primeiro note que:

$$\begin{aligned}
 (F'_k)^{-1}F_k &= (F'_k)^{-1}[F'_*(x_k - x_*) + O(\|x_k - x_*\|^2)] \\
 &= (F'_k)^{-1}F'_*(P_X + P_N)(x_k - x_*) + (F'_k)^{-1}O(\|x - x_*\|^2)] \\
 &= (F'_k)^{-1}F'_*P_X(x_k - x_*) + (F'_k)^{-1}O(\|x - x_*\|^2)].
 \end{aligned}$$

Então, aplicando o Lema 3.7, obtemos:

$$\|d_N\| = \|(F'_k)^{-1}F_k\| \leq c_6\|x_k - x_*\|,$$

para $x_k \in \mathcal{V}_\delta$ e c_6 uma constante positiva.

Assim,

$$\|x_k - x_* - d_N\| \leq \|x_k - x_*\| + \|d_N\| \leq c_7\|x_k - x_*\|.$$

Ou seja,

$$x_k - x_* - d_N = O(\|x_k - x_*\|).$$

Retornando a (3.8),

$$\begin{aligned}
 x_{k+1} - x_* &= \frac{1}{2}(F'_k)^{-1}F''_k O(\|x_k - x_*\|)O(\|x_k - x_*\|) + (F'_k)^{-1}O(\|x_k - x_*\|^3) \\
 &= O(\|x_k - x_*\|) + O(\|x_k - x_*\|^2) = O(\|x_k - x_*\|).
 \end{aligned}$$

A última afirmação do enunciado segue de observar que,

$$d_C = -\frac{1}{2}(F'_k)^{-1}F''_k d_N d_N = \frac{1}{2}G'_k d_N.$$

□

De certo modo, o passo corretor aproxima a correção teórica sugerida por Cavanagh ($\frac{1}{2}P_N(x_k - x_*)$). Nos seguintes exemplos unidimensionais vemos a melhora obtida pelo método de Chebyshev em comparação com Newton.

Exemplo 3.14. Considere $F(x) = x^2$, $x \in \mathbb{R}$. Vejamos como se comportam as sequências geradas pelos métodos de Newton e Chebyshev para resolver $F(x) = 0$.

Temos, $F'(x) = 2x$ e $F''(x) = 2$. Todas as hipóteses dos Teoremas 3.9 e 3.13 são satisfeitas.

- Newton

$$x_{k+1} = \frac{1}{2}x_k, \text{ ou seja, } \frac{x_{k+1}}{x_k} = \frac{1}{2}.$$

- Chebyshev

$$x_{k+1} = \frac{3}{8}x_k, \text{ ou seja, } \frac{x_{k+1}}{x_k} = \frac{3}{8}.$$

Exemplo 3.15. Considere agora $F(x) = x^3 + cx^2$, $x \in \mathbb{R}$ e c uma constante real não nula. Façamos a mesma análise do exemplo anterior:

Temos, $F'(x) = 3x^2 + 2cx$ e $F''(x) = 6x + 2c$. Neste caso a equação $F(x) = 0$ tem duas raízes, a saber, $x_* = -c$ e $x_* = 0$, no entanto estamos interessados no caso em que $F'(x) = 0$, ou seja, $x_* = 0$. As hipóteses dos Teoremas 3.9 e 3.13 são satisfeitas.

- Newton

$$x_{k+1} = x_k - \frac{x_k^3 + cx_k^2}{3x_k^2 + 2cx_k}.$$

E temos que:

$$\frac{x_{k+1}}{x_k} \longrightarrow \frac{1}{2} \text{ quando } x_k \longrightarrow x_*.$$

- Chebyshev

$$x_{k+1} = x_k - \frac{x_k^3 + cx_k^2}{3x_k^2 + 2cx_k} - \frac{x_k(3x_k + c)(x_k + c)^2}{(3x_k + 2c)^3}.$$

Neste caso temos:

$$\frac{x_{k+1}}{x_k} \longrightarrow \frac{3}{8} \text{ quando } x_k \longrightarrow x_*.$$

Uma modificação que fizemos no método de Chebyshev para melhorar a correção no sentido sugerido por Cavanagh foi adotar a sobrerelaxação $d_{Cm} = 2d_C$, o que chamamos de Chebyshev modificado, no lugar de d_C , ou seja, o processo iterativo torna-se:

$$x_{k+1} = x_k + d_N + 2d_C. \quad (3.9)$$

Assim temos $d_{Cm} \longrightarrow \frac{1}{2}P_{N_1}(x_k - x_*)$, e fazendo a mesma análise do Exemplo 3.14 obtemos:

$$x_{k+1} = \frac{1}{4}x_k \text{ e } \frac{x_{k+1}}{x_k} = \frac{1}{4}.$$

Já para o Exemplo 3.15 obtemos:

$$x_{k+1} = x_k - \frac{x_k^3 + cx_k^2}{3x_k^2 + 2cx_k} - \frac{2x_k(3x_k + c)(x_k + c)^2}{(3x_k + 2c)^3}$$

e

$$\frac{x_{k+1}}{x_k} \longrightarrow \frac{1}{4} \text{ quando } x_k \longrightarrow x_*.$$

O próximo resultado é uma adaptação do Teorema 3.13 para o método definido por (3.9).

Teorema 3.16. *Suponha as mesmas hipóteses do Teorema 3.13. Então, se x_0 é tomado suficientemente próximo a x_* , o método definido por (3.9) gera uma sequência convergente a x_* .*

Demonstração: Utilizando a mesma técnica da prova do Teorema 3.13, segue de (3.9) que:

$$\begin{aligned} x_{k+1} - x_* &= x_k - x_* - (F'_k)^{-1}F_k - (F'_k)^{-1}F''_k d_N d_N \\ &= (F'_k)^{-1}[-F_k + F'_k(x_k - x_*)] - (F'_k)^{-1}F''_k d_N d_N. \end{aligned}$$

Agora usando o Lema 3.7 e os seguintes fatos, já mostrados:

$$-F_k + F'_k(x_k - x_*) = O(\|x_k - x_*\|^2), \quad d_N = O(\|x_k - x_*\|),$$

obtemos,

$$x_{k+1} - x_* = O(\|x_k - x_*\|).$$

□

3.2.2 Análise da família Chebyshev-Halley

Agora vejamos como obter resultados similares de convergência para os métodos de Halley e super-Halley. Neste caso é necessário um pouco mais de trabalho, pois devemos ter condições para que as direções de correção dos métodos de Halley e super-Halley, respectivamente, d_H e d_{sH} existam, onde:

$$d_H = -\frac{1}{2}(F'_k + \frac{1}{2}F''_k d_N)^{-1}F''_k d_N d_N, \quad d_{sH} = -\frac{1}{2}(F'_k + F''_k d_N)^{-1}F''_k d_N d_N.$$

Analisemos o caso geral, $d = -(F'_k + \lambda F''_k d_N)^{-1}F''_k d_N d_N$. Que condições devemos ter para que $(F'_k + \lambda F''_k d_N)^{-1}$ exista?

Sabemos que $(F'_k)^{-1}$ existe para $x_k \in \mathcal{V}_\delta$. Pelo Teorema 2.3.4 de [60], temos que se $r = \|(F'_k)^{-1}\lambda F''_k d_N\| < 1$ então $(F'_k + \lambda F''_k d_N)^{-1}$ existe e

$$\|(F'_k + \lambda F''_k d_N)^{-1}\| \leq \frac{\|(F'_k)^{-1}\|}{1 - r}. \quad (3.10)$$

Temos então outra questão a responder: sob quais condições teremos $r < 1$?

A resposta não é difícil. Primeiramente observe que a condição $r < 1$ é equivalente a $\lambda\|G'_k\| < 1$ (G'_k definida na Proposição 3.2), e como $G'_k \rightarrow \frac{1}{2}P_N$ quando $x_k \rightarrow x_*$, então se $\frac{\lambda}{2}\|P_N\| < 1$ existe um $\bar{k} > 0$ tal que $r < 1$ para todo $k \geq \bar{k}$.

Lembrando que $\mathbb{R}^n = N \oplus X$, os seguintes resultados são conhecidos:

- Se θ é o ângulo entre N e X , então $\|P_N\| = \frac{1}{\sin \theta}$, $\theta \in (0, \frac{\pi}{2}]$ ([119, 5.15.3 p. 450]).
- $\|P_N\| = \|P_X\|$ ([119, 5.9.9 p. 391]).
- $\|P_N\| \geq 1$ ([119, 5.9.8 p. 391]).

Podemos agora obter condições para a inversibilidade de $(F'_k + \lambda F''_k d_N)$ a partir de θ :

- Para Halley temos $\lambda = \frac{1}{2}$. Assim de $\frac{\lambda}{2}\|P_N\| < 1$, obtemos $\sin \theta > \frac{1}{4}$, ou $\theta > \arcsin(0.25)$.
- Para super-Halley temos $\lambda = 1$. Assim de $\frac{\lambda}{2}\|P_N\| < 1$, obtemos $\sin \theta > \frac{1}{2}$, ou $\theta > \frac{\pi}{6}$.

- Para o caso geral devemos ter $\theta > \frac{\lambda}{2}$.

Já temos condições para a inversibilidade de $(F'_k + \lambda F''_k d_N)$, nos falta mostrar a convergência sob estas condições, o que segue das próximas equações:

Por (3.10) e pelo fato de que $d_N = O(\|x_k - x_*\|)$, temos:

$$\frac{1}{2}(F'_k + \lambda F''_k d_N)^{-1} F''_k d_N d_N = O(\|x_k - x_*\|). \quad (3.11)$$

De (3.11) e usando que $-F'_k + F'_k(x_k - x_*) = O(\|x_k - x_*\|^2)$ vem:

$$\begin{aligned} x_{k_1} - x_* &= x_k - x_* - (F'_k)^{-1} F'_k - \frac{1}{2}(F'_k + \lambda F''_k d_N)^{-1} F''_k d_N d_N \\ &= (F'_k)^{-1} [-F'_k + F'_k(x_k - x_*)] - \frac{1}{2}(F'_k + \lambda F''_k d_N)^{-1} F''_k d_N d_N \\ &= O(\|x_k - x_*\|). \end{aligned}$$

Vejamos como ficam as sequências iterativas dos métodos de Halley e super-Halley para as funções dos Exemplos 3.14 e 3.15:

- $F(x) = x^2$

- Halley

$$x_{k+1} = \frac{1}{3}x_k, \quad \text{ou seja, } \frac{x_{k+1}}{x_k} = \frac{1}{3}.$$

- super-Halley

$$x_{k+1} = \frac{1}{4}x_k, \quad \text{ou seja, } \frac{x_{k+1}}{x_k} = \frac{1}{4}.$$

- $F(x) = x^3 + cx^2$

- Halley

$$x_{k+1} = x_k - \frac{x_k^3 + cx_k^2}{3x_k^2 + 2cx_k} - \frac{x_k(3x_k + c)(x_k + c)^2}{(6x_k^2 + 8cx_k + 3c^2)(3x_k + 2c)}$$

e

$$\frac{x_{k+1}}{x_k} \longrightarrow \frac{1}{3} \quad \text{quando } x_k \longrightarrow x^*.$$

- super-Halley

$$x_{k+1} = x_k - \frac{x_k^3 + cx_k^2}{3x_k^2 + 2cx_k} - \frac{x_k(3x_k + c)(x_k + c)^2}{(3x_k^2 + 4cx_k + 2c^2)(3x_k + 2c)}$$

e

$$\frac{x_{k+1}}{x_k} \longrightarrow \frac{1}{4} \text{ quando } x_k \longrightarrow x^*.$$

Nestes exemplos observamos que as razões de convergência dos métodos da família Chebyshev-Halley são melhores que as obtidas pelo método de Newton. Mas note que não encontramos resultados teóricos que garantam que essas razões são sempre atingidas, isso permanece como um problema em aberto.

Capítulo 4

Problemas de Complementaridade

Este capítulo será dedicado a definir e descrever alguns tipos de problemas de complementaridade, bem como alguns problemas relacionados. Também comentaremos um pouco sobre métodos de resolução de tais problemas, sem descrevermos os próprios métodos em si. Apenas daremos algumas referências sobre os métodos citados, acompanhadas de comentários pertinentes.

O problema de complementaridade não linear (NCP) foi primeiramente estudado por Cottle [31]. A importância deste tipo de problemas reside no fato de que muitas aplicações podem ser formuladas, ou reformuladas, matematicamente por meio de problemas de complementaridade, por exemplo, problemas de teoria dos jogos, problemas de equilíbrio econômico, problemas de contato em mecânica, etc. Para detalhes sobre algumas aplicações ver [4, 5, 22, 51, 73, 125]. No início, as pesquisas estavam mais voltadas à teoria sobre existência e unicidade de soluções dos NCP's, ou seja, sob quais condições um NCP possui solução e se tal solução é única, [31, 92, 94, 98].

Em [31], Cottle definiu o problema

$$\begin{aligned} \min \quad & x^T F(x) \\ \text{s.a.} \quad & x, F(x) \geq 0 \end{aligned} \tag{4.1}$$

onde $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é diferenciável.

A nomenclatura “problema de complementaridade” não foi dada por Cottle,

ela surgiu posteriormente. Nesse mesmo trabalho, Cottle redefiniu o problema (4.1) na forma que hoje é mais conhecida: encontrar $x \in \mathbb{R}^n$ tal que

$$x, F(x) \geq 0, \quad x^T F(x) = 0. \quad (4.2)$$

No caso em que F é linear, ou seja, $F(x) = Mx + q$, onde M é uma matriz $n \times n$ e $q \in \mathbb{R}^n$, o problema (4.2) já havia sido estudado anteriormente por Dorn [43].

Com a introdução deste novo tipo de problema e com os resultados obtidos surgiram também extensões e problemas relacionados, como os problemas de complementaridade mista, problemas de complementaridade generalizados, problemas de complementaridade não linear como caso particular de desigualdades variacionais, etc.

Na seção 4.1 definiremos o problema de complementaridade não linear generalizado, algumas de suas variantes e listamos alguns resultados clássicos sobre a existência de solução. Na seção 4.2 listamos alguns resultados sobre métodos e algoritmos para resolução dos problemas citados na seção 4.1.

4.1 Definições

O problema de complementaridade não linear é atualmente mais conhecido sob a formulação (4.2). A equivalência entre (4.1) e (4.2) foi feita no trabalho introdutório de Cottle [31]. Como já foi comentado, quando F tem a forma

$$F(x) = Mx + q, \quad M \in \mathbb{R}^{n \times n} \quad \text{e} \quad q \in \mathbb{R}^n, \quad (4.3)$$

o problema recebe o nome de problema de complementaridade linear (LCP).

Os primeiros resultados teóricos sobre o problema (4.1) versavam sobre condições para a existência e unicidade de solução. Para o LCP tem-se o seguinte resultado de Dorn [43].

Teorema 4.1. *Se M é definida positiva, então o problema (4.2) tem uma solução ótima x_* tal que $x_*^T F(x_*) = 0$.*

Cottle em [31] obteve um resultado similar:

Teorema 4.2. *Se M tem os menores principais positivos, então o problema definido por (4.2) tem uma solução ótima x_* tal que $x_*^T F(x_*) = 0$.*

Ingleton [79] provou que a condição do Teorema 4.2 é necessária e suficiente para a existência de uma única solução do problema (4.2).

Existem também algumas extensões do LCP:

- hLCP: problema de complementaridade linear horizontal

Consiste em encontrar $x, y \in \mathbb{R}^n$ tais que

$$Rx + Qy = q, \quad (x, y) \geq 0, \quad x^T y = 0, \quad (4.4)$$

onde $R, Q \in \mathbb{R}^{n \times n}$.

Para $R = -M$ e $Q = I$ (4.4) torna-se (4.2).

- mLCP: problema de complementaridade linear misto

Consiste em encontrar $x, y \in \mathbb{R}^n$ e $z \in \mathbb{R}^m$ tais que

$$\begin{bmatrix} y \\ 0 \end{bmatrix} = M \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}, \quad (x, y) \geq 0, \quad x^T y = 0, \quad (4.5)$$

onde

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \in \mathbb{R}^{(m+n) \times (m+n)}.$$

Para $m = 0$, $M_{11} = M$ e $q_1 = q$, (4.5) torna-se (4.2).

4.1.1 Problema de complementaridade generalizado (GNCP)

Problemas de complementaridade generalizados são tipos de extensões do NCP nos quais o ortante positivo é substituído por um cone. Antes de definir algumas extensões vejamos a definição de cone, conforme [72].

Definição 4.3. *Um subconjunto $\mathcal{K} \subset \mathbb{R}^n$ será chamado de cone se satisfaz as três seguintes condições:*

1. \mathcal{K} é fechado;
2. $ax + by \in \mathcal{K}$, para todos $x, y \in \mathcal{K}$ e a, b reais não negativos;
3. se $x \in \mathcal{K} - \{0\}$, então $-x \notin \mathcal{K}$.

Também devemos definir um cone particular, relacionado a \mathcal{K} .

Definição 4.4. *Seja \mathcal{K} um cone em \mathbb{R}^n . $\mathcal{K}^\circ = \{y \in \mathbb{R}^n \mid x^T y \geq 0, \forall x \in \mathcal{K}\}$ é chamado de cone polar de \mathcal{K} . Se $\mathcal{K} \subset \mathcal{K}^\circ$, então \mathcal{K} é chamado de cone subpolar.*

Um exemplo de cone é o ortante não negativo em \mathbb{R}^n , e mais seu cone polar é ele próprio.

O GNCP como definido em [72] consistem em encontrar $x \in \mathbb{R}^n$ tal que

$$x \in \mathcal{K}, \quad G(x) \in \mathcal{K}^\circ, \quad x^T G(x) = 0, \quad (4.6)$$

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

Se $\mathcal{K} = \mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x \geq 0\}$, então (4.6) torna-se um NCP.

Atualmente a formulação do GNCP é uma extensão de (4.6): encontrar $x \in \mathbb{R}^m$ tal que

$$F(x) \in \mathcal{K}, \quad G(x) \in \mathcal{K}^\circ, \quad F(x)^T G(x) = 0, \quad (4.7)$$

onde $F, G : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ (cf. [88, 160, 174]).

No caso em que $m = n$ e $F(x) = x$, a forma (4.7) é reduzida à forma (4.6). Quando $\mathcal{K} = \mathbb{R}_+^m$, $G(x) = x - F(x)$ e $F : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ o problema (4.7) se reduz ao problema de complementaridade implícita, como definido em [129, 130].

4.1.2 Desigualdades variacionais

Desde o início dos estudos sobre problemas de complementaridade foi observado, por Karamardian [93], a relação existente entre tais problemas e problemas variacionais. Definamos o problema variacional em dimensão finita, conforme [73]:

Definição 4.5. *Sejam $X \subset \mathbb{R}^n$ um subconjunto não vazio, convexo e fechado, e $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$. O problema de desigualdade variacional é encontrar um vetor $x_* \in X$ tal que*

$$G(x_*)^T(u - x_*) \geq 0, \quad \text{para todo } u \in X. \quad (4.8)$$

Karamardian [93] mostrou que o conjunto solução do problema (4.6) coincide com o conjunto solução do problema variacional (4.8), com $X \equiv \mathcal{K}$. Para mais detalhes sobre estas relações uma referência é o excelente trabalho de Harker e Pang [73].

4.2 Métodos e Algoritmos

Existe uma grande variedade de métodos para resolver problemas de complementaridade, o primeiro algoritmo para resolver NCP's foi dado por Cottle [31] e era uma extensão do algoritmo de pivoteamento para o caso linear, já conhecido. Atualmente a maioria dos métodos de resolução de problemas de complementaridade é baseada em reformulações do problema. Destacamos a seguir algumas dessas reformulações e para quais problemas são aplicáveis.

4.2.1 Reformulações como problemas de otimização

Uma das técnicas mais utilizadas para resolver problemas de complementaridade é reformulá-los como problemas de otimização (restrita, irrestrita, diferenciável, não diferenciável, etc.). A partir disso obtêm-se condições sob as quais as soluções do problema de otimização sejam também soluções do problema de complementaridade. Esta técnica vem das origens do problema, já que a formulação inicial de um NCP dada por Cottle [31] era a de um problema de minimização.

Nos trabalhos [2, 3, 5, 8, 9, 10, 11, 12, 55, 90] os problemas de complementaridade ou desigualdades variacionais são reformulados como problemas de minimização, alguns irrestritos e outros com restrições. Em [6, 7, 91, 157] os autores utilizam a mesma técnica mas agora em GNCP's.

Já Pang e Gabriel [131] reformulam um NCP como um problema de minimização com restrições de não negatividade nas variáveis e utilizam um método de programação quadrática sequencial não suave para resolver a reformulação.

4.2.2 Reformulações como sistemas não lineares

Outra técnica bastante estudada é a de reformular um problema de complementaridade como um sistema não linear (diferenciável ou não diferenciável). Como na técnica anterior obtêm-se condições sob as quais uma solução do sistema seja solução do problema de complementaridade.

Nesta técnica podemos citar as modificações do método de Newton para equações não suaves, quando o sistema a ser resolvido não é diferenciável, como visto no Capítulo 2. Alguns trabalhos nesta linha são [35, 47, 54, 74, 169], que tratam de reformulações de NCP's. No trabalho [166] a técnica é utilizada em desigualdades variacionais.

Nesta linha também podemos citar os métodos de continuação, também chamados de seguidores de caminho, para mais detalhes ver [101, 103, 104].

4.2.3 Métodos de pontos interiores

Os métodos de pontos interiores surgiram na década de 80, com o trabalho de Karmarkar [95]. Neste trabalho o autor tratava de um método para resolver problemas de programação linear (PL) que tinha complexidade polinomial, o que causou uma grande expectativa na área, pois pouco tempo antes disso foi mostrado que existiam problemas para os quais a complexidade do método Simplex era exponencial. A partir daí houve uma intensa busca por melhores métodos de pontos interiores com boas propriedades teóricas, complexidade polinomial, e que também tivessem um bom desempenho prático. O final da década de 80 e início da década de 90 foi muito frutífero em trabalhos nesse sentido, tendo sido desenvolvida a base de grande parte dos métodos de pontos interiores hoje estudados e utilizados na prática, ver

[61, 102, 113, 116, 117, 120, 134, 144, 146, 172, 176, 177, 178].

Com o sucesso dos métodos de pontos interiores em PL começou a surgir o interesse de aplicar tal técnica a outros tipos de problemas, como programação quadrática [25, 158, 159], semidefinida [86, 87, 136], convexa, não linear [44, 170] e também a problemas de complementaridade [62, 67, 80, 99, 100, 105, 112, 118, 121, 122, 123, 135, 137, 153, 163, 164, 171, 179, 180, 181].

Em [67, 100, 112, 137, 163] os autores desenvolvem métodos de pontos interiores para NCP's monótonos, mais precisamente, Kojima, Noma e Yoshise em [100] definem uma vasta classe de métodos de pontos interiores e mostram a convergência, em [112] os autores modificam um método preditor-corretor de Mizuno-Todd-Ye, em [137] os autores sugerem um algoritmo de redução potencial, e em [163] os autores utilizam o referencial do trabalho [100] e definem um novo algoritmo.

Iusem em [80] propõe um método de pontos interiores para resolver NCP's com operadores paramonótonos, isto é, nos quais $\langle F(x) - F(y), x - y \rangle = 0$ implica que $F(x) = F(y)$. Monteiro, Pang e Wang em [123] combinam idéias de métodos de pontos interiores com idéias do algoritmo NE/SQP de Pang e Gabriel [131], para NCP's não monótonos.

Em [62, 99, 105, 118, 164, 171] os autores desenvolvem métodos de pontos interiores para LCP's monótonos. Em particular, Gonzaga em [62], Mehrotra e Stubbs em [118], e Ye e Anstreicher em [171] trabalham com métodos do tipo preditor-corretor. Simantiraki e Shanno em [153] desenvolvem um método de pontos interiores para LCP's não monótonos. Já em [135, 179, 180, 181] os autores desenvolvem métodos de pontos interiores para hLCP's.

Capítulo 5

GNCP em Cones Poliedrais

Muitos problemas em matemática, física, economia e engenharia podem ser formulados como desigualdades variacionais, problemas de complementaridade e problemas relacionados com (4.7), ver p.ex. [4, 5, 22, 51, 52, 97, 125, 133]. Apesar de (4.7) não ser um caso particular de *problema de desigualdade variacional* (VIP), ele é um caso especial de *problema de desigualdade variacional geral* (GVIP), como formulado por [124].

No caso em que $\mathcal{K} = \mathbb{R}_+^n$, a solução do GNCP por meio de reformulações foi considerada por Peng e Yuan [132] e Jiang et al. [88], baseadas na função de Fischer-Burmeister (F-B). A primeira reformulação para problema (4.7) em sua forma geral foi introduzida em [6], por meio do sistema KKT associado ao problema. Desde então esta reformulação tem sido adotada [160, 161, 174, 175], mas usando a função F-B aplicada às variáveis de folga e não a $F(x)$ ou a $G(x)$, o que impede a possibilidade de tirar proveito das propriedades teóricas da função F-B. As reformulações obtidas em [6, 7, 160, 161] foram feitas para métodos de primeira ordem. Em [174, 175] foram introduzidos métodos de segunda ordem não diferenciáveis. Também foram propostos e analisados algoritmos de projeção para problemas relacionados. Em [165] os autores desenvolveram uma estratégia de redes neurais para resolver desigualdades variacionais monótonas, enquanto em [145] foi apresentado um esquema de projeção para aplicação ao GVIP.

Apesar do método de Newton já ter sido estudado na resolução do NCP, seja por estratégias não diferenciáveis (ver [47, 74, 169]) ou por métodos de pontos interiores [100, 137, 163], ele ainda não havia sido estudado na resolução do GNCP sobre uma perspectiva que combina idéias diferenciáveis e pontos interiores.

Considerando a classe dos métodos de pontos interiores, variantes do algoritmo preditor-corretor original de Mehrotra [116] estão entre os algoritmos mais utilizados em implementações comerciais de Métodos de Pontos Interiores (ver [33, 167]). Além disso, o algoritmo preditor-corretor de Mehrotra é uma técnica de segunda ordem que tem sido objeto de recentes pesquisas ([146, 155, 180, 181]). Portanto, nada mais natural que considerar estes dois elementos: um problema importante com algoritmos proeminentes.

Este capítulo é organizado do seguinte modo. Na Seção 5.1 estabelecemos o GNCP em cones poliedrais e introduzimos um sistema não linear equivalente com algumas variáveis não negativas. Na Seção 5.2 é apresentado um algoritmo de pontos interiores baseado no método de Newton para resolver o problema equivalente definido na Seção 5.1, e também é feita a análise de convergência local de tal algoritmo. Na Seção 5.3 estabelecemos um algoritmo do tipo preditor-corretor para resolver a reformulação do GNCP e analisamos a sua convergência. Na Seção 5.4 são apresentados experimentos numéricos para avaliar o desempenho dos algoritmos introduzidos nas seções anteriores. O material deste capítulo foi sistematizado no artigo [13].

5.1 GNCP em cones poliedrais e um problema equivalente

No que segue, consideraremos o caso do problema (4.7) com $m = n$, $F, G \in C^1$ e \mathcal{K} um cone poliedral em \mathbb{R}^n , isto é,

$$\mathcal{K} = \{v \in \mathbb{R}^n \mid Av \geq 0, Bv = 0\}, \tag{5.1}$$

onde $A \in \mathbb{R}^{q \times n}$ e $B \in \mathbb{R}^{s \times n}$. Assim, o cone polar é dado por

$$\mathcal{K}^\circ = \{u \in \mathbb{R}^n \mid u = A^T \lambda^1 + B^T \lambda^2, \lambda^1 \geq 0\}. \quad (5.2)$$

Seja x_* uma solução do problema (4.7) com as escolhas acima para \mathcal{K} e \mathcal{K}° e considere o seguinte problema de programação convexa

$$\begin{aligned} \min \quad & y^T G(x_*) \\ \text{s.a.} \quad & y \in \mathcal{K}, \end{aligned}$$

que pode ser reescrito como

$$\begin{aligned} \min \quad & y^T G(x_*) \\ \text{s.a.} \quad & By = 0 \\ & Ay \geq 0. \end{aligned} \quad (5.3)$$

Note que devido à linearidade das restrições, o problema (5.3) naturalmente satisfaz uma condição de qualificação (*constraint qualification*). Identificando $y \equiv F(x_*)$, o sistema KKT associado ao problema (5.3) é dado por

$$\begin{aligned} G(x_*) - A^T \lambda^1 - B^T \lambda^2 &= 0 \\ BF(x_*) &= 0 \\ AF(x_*) - z &= 0 \\ \lambda^1 \geq 0, z \geq 0, (\lambda^1)_i z_i &= 0, \forall i. \end{aligned}$$

Este raciocínio motiva a introdução do seguinte sistema não linear com restrição de não negatividade em algumas variáveis:

$$H(x, z, \lambda^1, \lambda^2) = \begin{pmatrix} AF(x) - z \\ BF(x) \\ G(x) - A^T \lambda^1 - B^T \lambda^2 \\ Z \lambda^1 \end{pmatrix} = 0 \quad (5.4)$$

com $z, \lambda^1 \geq 0$, e $Z = \text{diag}(z)$.

O seguinte resultado estabelece uma relação entre as soluções dos problemas (4.7) e (5.4):

Teorema 5.1. *Se x_* é solução do problema (4.7), então existem $z_*, \lambda_*^1 \in \mathbb{R}^q$ e $\lambda_*^2 \in \mathbb{R}^s$, tais que $w_* = (x_*, z_*, \lambda_*^1, \lambda_*^2)$ é solução do problema (5.4). Reciprocamente, se w_* é solução do problema (5.4), então x_* é solução do problema (4.7).*

Demonstração: Seja x_* solução do GNCP. Então, de $F(x_*) \in \mathcal{K}$ obtemos

$$AF(x_*) \geq 0 \text{ e } BF(x_*) = 0.$$

Assim, deve existir

$$z_* \in \mathbb{R}^q, z_* \geq 0, \text{ tal que } AF(x_*) - z_* = 0.$$

Agora, pelo fato de $G(x_*) \in \mathcal{K}^\circ$ segue que existem $\lambda_*^1 \in \mathbb{R}^q, \lambda_*^1 \geq 0$, e $\lambda_*^2 \in \mathbb{R}^s$ tais que

$$G(x_*) = A^T \lambda_*^1 + B^T \lambda_*^2.$$

Só resta mostrar que $Z_* \lambda_*^1 = 0$, ou equivalentemente, $z_*^T \lambda_*^1 = 0$.

$$\begin{aligned} z_*^T \lambda_*^1 &= [AF(x_*)]^T \lambda_*^1 \\ &= F(x_*)^T A^T \lambda_*^1 \\ &= F(x_*)^T (G(x_*) - B^T \lambda_*^2) \\ &= F(x_*)^T G(x_*) - F(x_*)^T B^T \lambda_*^2 \\ &= F(x_*)^T G(x_*) - [BF(x_*)]^T \lambda_*^2 \\ &= F(x_*)^T G(x_*) = 0. \end{aligned}$$

Suponhamos agora que w_* seja solução do problema (5.4). Pelas duas primeiras equações de (5.4), e pelo fato de $z_* \geq 0$, vem que $F(x_*) \in \mathcal{K}$. Da terceira equação de (5.4) e de $\lambda_*^1 \geq 0$ obtemos $G(x_*) \in \mathcal{K}^\circ$, e como $F(x_*)^T G(x_*) = z_*^T \lambda_*^1$, da quarta equação de (5.4) vem que $F(x_*)^T G(x_*) = 0$. Portanto, x_* é solução do problema (4.7). □

5.2 Um algoritmo de Newton perturbado e pontos interiores

A principal idéia deste capítulo é utilizar métodos de pontos interiores para encontrar uma solução do sistema (5.4), obtendo assim uma solução de (4.7) pelo Teorema 5.1.

Um protótipo de algoritmo de pontos interiores de Newton para resolver (5.4) será dado a seguir. A função que define o sistema não linear, calculada na sequência gerada, é denotada por $H_k \equiv H(w_k)$, enquanto que sua matriz Jacobiana será representada por $H'_k \equiv H'(w_k)$. Explicitamente, esta matriz é dada por

$$H'(w) = \begin{bmatrix} AF'(x) & -I & 0 & 0 \\ BF'(x) & 0 & 0 & 0 \\ G'(x) & 0 & -A^T & -B^T \\ 0 & \Lambda^1 & Z & 0 \end{bmatrix} \quad (5.5)$$

na qual $Z = \text{diag}(z)$, $\Lambda^1 = \text{diag}(\lambda^1)$. Para completar a notação utilizada, a dimensão total é $n + 2q + s \equiv n_1$, e os vetores \hat{e} e d são dados por $\hat{e} = (0, 0, 0, e)^T \in \mathbb{R}^{n_1}$, com $e = (1, 1, \dots, 1)^T \in \mathbb{R}^s$ e $d = (d_x, d_z, d_{\lambda^1}, d_{\lambda^2})^T \in \mathbb{R}^{n_1}$.

Algoritmo 1. Seja w_0 com $z_0, \lambda_0^1 > 0$ dados. Para $k = 0, 1, 2, \dots$ faça:

1. Escolha $\mu_k \in \mathbb{R}$ tal que $\mu_k > 0$.
2. Calcule d , solução do sistema linear $H'_k d = -H_k + \mu_k \hat{e}$.
3. Escolha $\gamma_k \in (0, 1)$ e encontre

$$\alpha_k = \min \left\{ 1, \gamma_k \min_i \left\{ -\frac{\lambda_i^1}{(d_{\lambda^1})_i} \mid (d_{\lambda^1})_i < 0 \right\}, \gamma_k \min_i \left\{ -\frac{z_i}{(d_z)_i} \mid (d_z)_i < 0 \right\} \right\}.$$

4. Faça $w_{k+1} = w_k + \alpha_k d$.

No que diz respeito às escolhas inerentes aos Passos 1 e 3, condições teóricas serão estabelecidas na análise de convergência e expressões particulares utilizadas na implementação serão oportunamente apresentadas.

Para provar que o Algoritmo 1 está bem definido e analisar sua convergência local, as seguintes hipóteses são necessárias.

Hipóteses

- A1.** O sistema (5.4) tem uma solução $w_* \in \mathbb{R}^{n_1}$.
- A2.** As funções $F, G \in C^2$, F'' e G'' são Lipschitz contínuas, $F'(x)$ é inversível e $F'(x)^T G'(x)$ é definida positiva no núcleo de $BF'(x)$ para cada $x \in \mathbb{R}^n$.
- A3.** A matriz $B \in \mathbb{R}^{s \times n}$ tem posto completo.

Verificar que a matriz $F'(x)^T G'(x)$ é definida positiva no núcleo de $BF'(x)$ não é fácil de ser feito, na prática. Uma condição mais forte e no entanto mais simples de ser verificada é $F'(x)^T G'(x) > 0$, e nesta condição já fica implícita a inversibilidade de F' . Porém, em nossa hipótese, a não singularidade de G' não é obrigatória. Vale a pena mencionar que os autores de [160, 174], que utilizam abordagem não diferenciável, assumem que ambos os Jacobianos são não singulares na solução (denotados por F'_* e G'_*), e além disso B tem posto completo e $A(G'_*)^{-T}(F'_*)^T A^T$ é uma P -matriz, isto é, todos os seus menores principais são positivos. Infelizmente, tais hipóteses restringem a q o número de faces do cone poliedral, que não pode ser maior que a dimensão do espaço n .

5.2.1 O Algoritmo 1 está bem definido

O resultado abaixo garante que a matriz Jacobiana de (5.4) admite uma inversa sempre que for calculada na sequência de pontos gerados pelo Algoritmo 1.

Lema 5.2. *Se $w = (x, z, \lambda^1, \lambda^2)$ é tal que $z, \lambda^1 > 0$ e as hipóteses A2 e A3 são satisfeitas, então a matriz $H'(w)$ é não singular.*

Demonstração: Para simplificar a apresentação, no que segue omitiremos a dependên-

cia em w e x . Considere o sistema

$$H' \begin{bmatrix} u \\ v \\ y \\ t \end{bmatrix} = 0,$$

que pode ser reescrito como

$$AF'u - v = 0 \tag{5.6}$$

$$BF'u = 0 \tag{5.7}$$

$$G'u - A^T y - B^T t = 0 \tag{5.8}$$

$$\Lambda^1 v + Zy = 0. \tag{5.9}$$

Pré-multiplicando a equação (5.8) por $u^T(F')^T$ e usando (5.6) e (5.7), temos

$$u^T(F')^T G'u - v^T y = 0. \tag{5.10}$$

Lembrando que $z, \lambda^1 > 0$, então da equação (5.9) vem que

$$y = -Z^{-1}\Lambda^1 v. \tag{5.11}$$

Substituindo (5.11) em (5.10) obtemos

$$u^T(F')^T G'u + v^T Z^{-1}\Lambda^1 v = 0. \tag{5.12}$$

Como $Z^{-1}\Lambda^1$ é definida positiva, assumindo que $F'(x)$ é inversível e $F'(x)^T G'(x)$ é definida positiva no núcleo de $BF'(x)$ teremos, da equação (5.12), $v = 0$ e $u = 0$. E por (5.11), $y = 0$. A equação (5.8) torna-se $B^T t = 0$, mas B tem posto completo, logo $t = 0$. \square

Um caso particular de interesse surge quando no cone \mathcal{K} não há a matriz B , ou seja,

$$\mathcal{K} = \{v \in \mathbb{R}^n \mid Av \geq 0\}.$$

Em tal caso, as condições do Lema 5.2 são simplificadas como segue:

Corolário 5.3. *Se $w = (x, z, \lambda^1)$ é tal que $z, \lambda^1 > 0$, $G'(x)$ é inversível e $F'(x)^T G'(x)$ é semidefinida positiva, então a matriz $H'(w)$ é não singular.*

5.2.2 Resultados preliminares

Na sequência vamos introduzir alguns lemas técnicos que auxiliarão o estabelecimento dos resultados principais deste capítulo.

Lema 5.4. *Sejam $D_1 \subset \mathbb{R}^n, D_2, D_3 \subset \mathbb{R}^q, D_4 \subset \mathbb{R}^s$ vizinhanças abertas de x_*, z_*, λ_*^1 e λ_*^2 , respectivamente. Então, existe $L > 0$ tal que:*

$$\|G'(x) - G'(x')\| \leq L\|x - x'\| \quad \text{para todos } x, x' \in D_1; (5.13)$$

$$\|F'(x) - F'(x')\| \leq L\|x - x'\| \quad \text{para todos } x, x' \in D_1; (5.14)$$

$$\|H(w) - H(w')\| \leq L\|w - w'\| \quad \text{para todos } w, w' \in D = D_1 \times D_2 \times D_3 \times D_4; (5.15)$$

$$\|H'(w) - H'(w')\| \leq L\|w - w'\| \quad \text{para todos } w, w' \in D; (5.16)$$

$$\|H''(w) - H''(w')\| \leq L\|w - w'\| \quad \text{para todos } w, w' \in D; (5.17)$$

$$\|H(w) - H(w_*) - H'(w)(w - w_*)\| \leq \frac{1}{2}L\|w - w_*\|^2 \quad \text{para todo } w \in D. (5.18)$$

Demonstração: As desigualdades (5.13)-(5.17) seguem da hipótese A2. Para obter (5.18) note que

$$\begin{aligned} H(w_*) - H(w) - H'(w)(w_* - w) &= \int_0^1 H'(w + t(w_* - w))(w_* - w)dt + H'(w)(w_* - w) \\ &= \int_0^1 [H'(w + t(w_* - w)) - H'(w)](w_* - w)dt, \end{aligned}$$

e portanto

$$\begin{aligned} \|H(w) - H(w_*) - H'(w)(w - w_*)\| &\leq \|w - w_*\| \int_0^1 \|H'(w + t(w_* - w)) - H'(w)\| dt \\ &\leq L\|w - w_*\| \int_0^1 \|w + t(w_* - w) - w\| dt \\ &\leq L\|w - w_*\|^2 \int_0^1 t dt = \frac{1}{2}L\|w - w_*\|^2. \end{aligned}$$

□

O próximo lema é clássico em análise matricial, e $\|\cdot\|_p$ denota uma p -norma geral.

Lema 5.5. *Se $A \in \mathbb{R}^{n \times n}$ é não singular e $r = \|A^{-1}E\|_p < 1$, então $(A + E)$ é não singular e valem as seguintes desigualdades*

$$\|(A + E)^{-1}\|_p \leq \frac{\|A^{-1}\|_p}{1 - r}, \quad \|(A + E)^{-1} - A^{-1}\|_p \leq \frac{\|E\|_p \|A^{-1}\|_p^2}{1 - r}.$$

Demonstração: Ver [60, Teorema 2.3.4] □

Os próximos dois lemas foram adaptados de Yamashita e Yabe [170], que trabalharam com métodos de pontos interiores para otimização com restrições. Os resultados foram reescritos em termos das variáveis do GNCP, no contexto do Algoritmo 1.

Para a demonstração dos próximos resultados vamos precisar da seguinte hipótese:

A4. A complementaridade estrita se cumpre na solução w_* , isto é, $z_* + \lambda_*^1 > 0$.

Lema 5.6. *Suponha que a hipótese A4 seja satisfeita e defina*

$$\kappa = 2 \max \left\{ \left\{ \frac{1}{(z_*)_i}; (z_*)_i > 0, i = 1, 2, \dots, q \right\}, \left\{ \frac{1}{(\lambda_*^1)_i}; (\lambda_*^1)_i > 0, i = 1, 2, \dots, q \right\} \right\}.$$

Existe $\epsilon_0 > 0$ tal que, se $\|w_k - w_\| \leq \epsilon_0$ e d é definido no Passo 2 do Algoritmo 1 então, para cada i tal que $(\lambda_*^1)_i = 0$,*

$$\frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} = -1 + \frac{\mu_k}{(z_k)_i (\lambda_k^1)_i} + (p_k)_i, \quad |(p_k)_i| \leq \kappa \|d\|, \quad \left| \frac{(d_z)_i}{(z_k)_i} \right| \leq \kappa \|d\|,$$

e para cada i tal que $(\lambda_^1)_i > 0$ ($(z_*)_i = 0$) temos*

$$\frac{(d_z)_i}{(z_k)_i} = -1 + \frac{\mu_k}{(z_k)_i (\lambda_k^1)_i} + (q_k)_i, \quad |(q_k)_i| \leq \kappa \|d\|, \quad \left| \frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} \right| \leq \kappa \|d\|.$$

Demonstração: Se $(\lambda_*^1)_i = 0$ então, pela hipótese A4, $(z_*)_i > 0$. Da definição de κ sabemos que $\frac{1}{(z_*)_i} < \kappa$. Deste modo, existe ϵ_0 suficientemente pequeno tal que, se $\|w_k - w_*\| \leq \epsilon_0$, então $\frac{1}{(z_k)_i} \leq \kappa$. Logo

$$\left| \frac{(d_z)_i}{(z_k)_i} \right| \leq \frac{\|d\|}{(z_k)_i} \leq \kappa \|d\|. \tag{5.19}$$

Agora, do Passo 2 do Algoritmo 1, temos

$$\Lambda_k^1 d_z + Z_k d_{\lambda^1} = \mu_k e - Z_k \lambda_k^1,$$

que pode ser reescrita como

$$Z_k d_z + \Lambda_k^1 d_{\lambda^1} = \mu_k (Z_k \Lambda_k^1)^{-1} e - e,$$

ou ainda

$$\frac{(d_z)_i}{(z_k)_i} + \frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} = \frac{\mu_k}{(z_k)_i (\lambda_k^1)_i} - 1, \quad i = 1, 2, \dots, q. \quad (5.20)$$

Por (5.19) e (5.20) obtemos que

$$\frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} = -1 + \frac{\mu_k}{(z_k)_i (\lambda_k^1)_i} + (p_k)_i, \quad \text{onde } |(p_k)_i| \leq \kappa \|d\|.$$

Supondo agora que $(\lambda_*^1)_i = 0$ segue, da hipótese A4, que $(z_*)_i = 0$. De modo similar podemos obter que

$$\left| \frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} \right| \leq \kappa \|d\|,$$

e conseqüentemente

$$\frac{(d_z)_i}{(z_k)_i} = -1 + \frac{\mu_k}{(z_k)_i (\lambda_k^1)_i} + (q_k)_i, \quad \text{onde } |(q_k)_i| \leq \kappa \|d\|.$$

□

O Lema 5.6 é usado para provar o seguinte resultado.

Lema 5.7. *Sob as hipóteses do Lema 5.6, se $\kappa \|d_k\| \leq \gamma_k$, então*

$$1 \geq \alpha_k \geq \gamma_k - \kappa \|d_k\|.$$

Demonstração: Se $(d_z)_i, (d_{\lambda^1})_i \geq 0$ para todo i , então $(z_k + d_z, \lambda_k^1 + d_{\lambda^1}) > 0$ e o resultado é válido com $\alpha_k \equiv 1$.

Suponha que existe um índice i tal que $(d_{\lambda^1})_i < 0$. Se $(\lambda_*^1)_i > 0$, então do Lema 5.6 e de $\kappa \|d_k\| \leq \gamma_k$ temos

$$-\gamma_k \frac{(\lambda_k^1)_i}{(d_{\lambda^1})_i} \geq \frac{\gamma_k}{\kappa \|d_k\|} \geq 1.$$

Se $(\lambda_*^1)_i = 0$, então do Lema 5.6 obtemos

$$-\frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} = 1 - \frac{\mu_k}{(z_k)_i(\lambda_k^1)_i} - (p_k)_i \leq 1 + \kappa \|d_k\|,$$

e assim,

$$-\gamma_k \frac{(\lambda_k^1)_i}{(d_{\lambda^1})_i} \geq \frac{\gamma_k}{1 + \kappa \|d_k\|} \geq \gamma_k(1 - \kappa \|d_k\|).$$

Utilizando o mesmos argumentos para d_z obtemos

$$\alpha_k \geq \min\{1, \gamma_k(1 - \kappa \|d_k\|)\} = \gamma_k(1 - \kappa \|d_k\|) \geq \gamma_k - \kappa \|d_k\|.$$

Portanto,

$$1 \geq \alpha_k \geq \gamma_k - \kappa \|d_k\|.$$

□

O próximo resultado fornece a inversibilidade da matriz Jacobiana na solução w_* de (5.4), e é utilizado no lema que vem depois dele.

Lema 5.8. *Suponha a validade das hipóteses A1, A2, A3 e A4. Então, a matriz $H'(w_*)$ é não singular.*

Demonstração: Considere o sistema linear:

$$H'(w_*) \begin{bmatrix} u \\ v \\ y \\ t \end{bmatrix} = 0,$$

que pode ser reescrito como

$$AF_*'u - v = 0, \tag{5.21}$$

$$BF_*'u = 0, \tag{5.22}$$

$$G_*'u - A^T y - B^T t = 0, \tag{5.23}$$

$$\Lambda_*^1 v + Z_* y = 0. \tag{5.24}$$

Da equação (5.24) e da hipótese A3 segue que $v^T y = 0$.

Utilizando a mesma idéia da demonstração do Lema 5.2, $u^T (F'_*)^T G'_* u - v^T y = 0$, e assim, $u^T (F'_*)^T G'_* u = 0$.

Deste modo, pela hipótese A2, devemos ter $u = 0$. Portanto, $v = 0$, $y = 0$ e $t = 0$. □

Note que, apesar das equações (5.9) e (5.24) serem parecidas, a hipótese $z, \lambda^1 > 0$ do Lema 5.2 é mais forte que a hipótese A3 requerida pelo Lema 5.8. De qualquer modo, a complementaridade estrita na solução $z_* + \lambda_*^1 > 0$ é suficiente para garantir que a matriz Jacobiana calculada nesta solução seja inversível.

Lema 5.9. *Suponha a validade das hipóteses A1, A2, A3 e A4 e seja $\{w_k\}$ uma seqüência gerada pelo Algoritmo 1. Sejam $\epsilon, \delta_1, \delta_2 > 0$ tais que*

$$\|w_k - w_*\| < \epsilon, \quad \|F'_k - F'_*\| < \delta_1, \quad \|G'_k - G'_*\| < \delta_2,$$

então existem $\zeta, \tau > 0$ tais que

$$\|H'_k - H'_*\| \leq \zeta, \quad \|(H'_k)^{-1}\| < \tau,$$

e

$$(1 - \alpha_k) \leq (1 - \gamma_k) + O(\|H_k\|) + O(\mu_k).$$

Demonstração: Como

$$H'_k - H'_* = \begin{bmatrix} A(F'_k - F'_*) & 0 & 0 & 0 \\ B(F'_k - F'_*) & 0 & 0 & 0 \\ G'_k - G'_* & 0 & 0 & 0 \\ 0 & \Lambda_k^1 - \Lambda_*^1 & Z_k - Z_* & 0 \end{bmatrix}.$$

temos

$$\begin{aligned} \|H'_k - H'_*\| &\leq \|H'_k - H'_*\|_F \leq \|A\|_F \|F'_k - F'_*\|_F + \|B\|_F \|F'_k - F'_*\|_F + \|G'_k - G'_*\|_F \\ &\quad + \|\Lambda_k^1 - \Lambda_*^1\|_F + \|Z_k - Z_*\|_F \\ &\leq (k_1 + k_2 + 1)L\epsilon + 2k_3\epsilon \equiv k_4\epsilon \equiv \zeta, \end{aligned}$$

onde $\|\cdot\|_F$ é a norma de Frobenius.

Tomando $A = H'_*$, $E = H'_k - H'_*$ e escolhendo $\epsilon > 0$ suficientemente pequeno, tal que $\zeta \epsilon \|(H'_*)^{-1}\| < 1$ segue, da desigualdade anterior e do Lema 5.5, que a matriz Jacobiana H'_k é inversível e

$$\|(H'_k)^{-1}\| \leq \frac{\|(H'_*)^{-1}\|}{1 - \zeta \epsilon \|(H'_*)^{-1}\|} \equiv \tau.$$

Agora,

$$\begin{aligned} \|d_k\| &= \|(H'_k)^{-1}(-H_k + \mu_k \hat{e})\| \leq \|(H'_k)^{-1}\|(\|H_k\| + \mu_k \|\hat{e}\|) \\ &\leq \tau(\|H_k\| + \mu_k \|\hat{e}\|). \end{aligned}$$

Deste modo, pelo Lema 5.7 temos

$$0 \leq (1 - \alpha_k) \leq (1 - \gamma_k) + k\|d_k\| \leq (1 - \gamma_k) + k(O(\|H_k\|) + O(\mu_k)),$$

e portanto,

$$(1 - \alpha_k) \leq (1 - \gamma_k) + O(\|H_k\|) + O(\mu_k).$$

□

5.2.3 Convergência local do Algoritmo 1

Sob as hipóteses usuais feitas em métodos de pontos interiores para os parâmetros γ_k e μ_k do Algoritmo 1 [44, 170], a convergência local quadrática é estabelecida.

Teorema 5.10. *Suponha que sejam válidas A1, A2, A3 e A4 e que a sequência $\{w_k\}$ é gerada pelo Algoritmo 1. Se $1 - \gamma_k = O(\|H_k\|)$ e $\mu_k = O(\|H_k\|^2)$, então existe $\epsilon > 0$ tal que, para $\|w_0 - w_*\| < \epsilon$, a sequência $\{w_k\}$ está bem definida, $w_k \rightarrow w_*$ e*

$$\|w_{k+1} - w_*\| \leq \nu \|w_k - w_*\|^2,$$

para alguma constante $\nu > 0$.

Demonstração: Suponha $\|w_k - w_*\| < \epsilon$. Como H'_k é não singular e $\|(H'_k)^{-1}\| \leq \tau$, temos

$$\begin{aligned} w_{k+1} - w_* &= w_k - w_* + \alpha_k d_k \\ &= w_k - w_* + \alpha_k (H'_k)^{-1} (-H_k + \mu_k \hat{e}) + \alpha_k (w_k - w_*) - \alpha_k (w_k - w_*) \\ &= (1 - \alpha_k)(w_k - w_*) + \alpha_k (H'_k)^{-1} (-H_k + H_* + H'_k (w_k - w_*)) \\ &\quad + \alpha_k (H'_k)^{-1} \mu_k \hat{e} \end{aligned}$$

Agora, pelos Lemas 5.4 e 5.9:

$$\begin{aligned} \|w_{k+1} - w_*\| &\leq |1 - \alpha_k| \|w_k - w_*\| + \alpha_k \|(H'_k)^{-1}\| \|H_k - H_* - H'_k (w_k - w_*)\| \\ &\quad + \alpha_k \|(H'_k)^{-1}\| \|\mu_k \hat{e}\| \\ &\leq [(1 - \gamma_k) + O(\|H_k\|) + O(\mu_k)] \|w_k - w_*\| + O(\|w_k - w_*\|^2) + O(\mu_k) \end{aligned}$$

Como, por hipótese, $1 - \gamma_k = O(\|H_k\|)$ e $\mu_k = O(\|H_k\|^2)$, obtemos:

$$\|w_{k+1} - w_*\| = O(\|w_k - w_*\|^2).$$

□

5.3 Um algoritmo preditor-corretor de pontos interiores

Nesta seção vamos nos utilizar das técnicas e resultados das seções anteriores para estabelecer a convergência de um método preditor-corretor para resolver o GNCP na forma (5.4). Nosso método é baseado no método de Chebyshev para sistemas não lineares [24, 68, 69], que será adaptado à filosofia dos métodos de pontos interiores. Vale a pena mencionar que o bem conhecido método preditor-corretor de Mehrotra, [155], pode ser visto com um caso particular do algoritmo aqui proposto.

Seguindo a idéia de Mehrotra, nosso método também utiliza a mesma matriz Jacobiana no cálculo das duas direções, preditora e corretora, como será detalhado a seguir. Em alguns casos vamos omitir o índice k para simplificar a notação.

Direção preditora. A direção preditora é a direção de Newton, solução do seguinte sistema linear

$$H'_k d_N = -H_k, \quad (5.25)$$

no qual H'_k e H_k são como definidas no Algoritmo 1.

Do último conjunto de equações do sistema (5.25) temos

$$\Lambda_k^1 d_{N_z} + Z_k d_{N_{\lambda^1}} = -Z_k \Lambda_k^1 e, \quad (5.26)$$

no qual usamos a notação $d_N = (d_{N_x}, d_{N_z}, d_{N_{\lambda^1}}, d_{N_{\lambda^2}}) \in \mathbb{R}^{n_1}$.

Direção corretora. A direção corretora do método de Chebyshev é a solução do seguinte sistema linear, no qual a perturbação já está incluída

$$H'_k d_C = r_k + \mu_k \hat{e}, \quad (5.27)$$

no qual $D_{N_z} = \text{diag}(d_{N_z})$, $D_{N_{\lambda^1}} = \text{diag}(d_{N_{\lambda^1}})$ e

$$r_k = -\frac{1}{2} \begin{bmatrix} AF''_k d_{N_x} d_{N_x} \\ BF''_k d_{N_x} d_{N_x} \\ G''_k d_{N_x} d_{N_x} \\ 2D_{N_z} D_{N_{\lambda^1}} e \end{bmatrix} = -\frac{1}{2} H''_k d_N d_N. \quad (5.28)$$

Note que $H''(w) \in \mathbb{R}^{n_1 \times n_1 \times n_1}$ é um tensor de terceira ordem, com elementos

$$[H''(w)]_{lij} = \frac{\partial^2 H_\ell}{\partial w_i \partial w_j}(w),$$

e sua ação em vetores $u, v \in \mathbb{R}^{n_1}$ é dada no Apêndice A.

Do último conjunto de equações do sistema (5.27) temos:

$$\Lambda_k^1 d_{C_z} + Z_k d_{C_{\lambda^1}} = -D_{N_z} D_{N_{\lambda^1}} e + \mu_k e, \quad (5.29)$$

e denotamos $d_c = (d_{C_x}, d_{C_z}, d_{C_{\lambda^1}}, d_{C_{\lambda^2}}) \in \mathbb{R}^{n+2q+s}$.

A direção final $d = (d_x, d_z, d_{\lambda^1}, d_{\lambda^2})$ é dada pela soma das duas direções d_N e d_C , isto é, $d = d_N + d_C$.

Das equações (5.26) e (5.29) obtemos:

$$\Lambda_k^1 d_z + Z_k d_{\lambda^1} = -Z_k \Lambda_k^1 e - D_{Nz} D_{N\lambda^1} e + \mu_k e. \quad (5.30)$$

Visto como calcular as direções podemos definir o algoritmo preditor-corretor.

Algoritmo 2: Sejam $w_0 = (x_0, z_0, \lambda_0^1, \lambda_0^2)$ com $z_0, \lambda_0^1 > 0$, para $k = 0, 1, 2, \dots$, faça:

1. Encontre d_N solução do sistema linear de Newton: $H'_k d_N = -H_k$.
2. Escolha $\mu_k > 0$.
3. Encontre d_C solução do sistema linear de Chebyshev: $H'_k d_C = r_k + \mu_k \hat{e}$.
4. Faça $d = d_N + d_C$.
5. Escolha $\gamma_k \in (0, 1)$ e encontre

$$\alpha_k = \min \left\{ 1, \gamma_k \min_i \left\{ -\frac{(\lambda_k^1)_i}{(d_{\lambda^1})_i} \mid (d_{\lambda^1})_i < 0 \right\}, \gamma_k \min_i \left\{ -\frac{(z_k)_i}{(d_z)_i} \mid (d_z)_i < 0 \right\} \right\}.$$

6. Faça $w_{k+1} = w_k + \alpha_k d$.

No que se refere às escolhas dos parâmetros μ_k e γ_k , como observado após o Algoritmo 1, condições teóricas e práticas serão dadas no momento oportuno.

Como a matriz H'_k é utilizada no cálculo das duas direções, sob as Hipóteses A1, A2, A3 e A4, segue pelo Lema 5.2 que H'_k é não singular. Consequentemente, o Algoritmo 2 está bem definido.

Para poder chegar a um resultado de convergência para o Algoritmo 2, similar ao Teorema 5.10, os Lemas 5.6, 5.7 e 5.9 precisam ser adaptados, o que será feito a seguir.

Lema 5.11. *Suponha que a hipótese A3 seja satisfeita, seja $\{w_k\}$ uma sequência gerada pelo Algoritmo 2, e defina*

$$\kappa_1 = 2 \max \left\{ \left\{ \frac{1}{(z_*)_i} \mid (z_*)_i > 0, i = 1, 2, \dots, q \right\}, \left\{ \frac{1}{(\lambda_*^1)_i} \mid (\lambda_*^1)_i > 0, i = 1, 2, \dots, q \right\} \right\}.$$

Existe $\epsilon_0 > 0$ tal que, se $\|w_k - w_*\| \leq \epsilon_0$ e considerando d como definida no Passo 4 do Algoritmo 2, então para cada i tal que $(\lambda_*^1)_i = 0$,

$$\frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} = -1 + \frac{\mu_k}{(z_k)_i(\lambda_k^1)_i} + (p_k)_i, \quad |(p_k)_i| = O(\|H_k\|) + O(\mu_k), \quad (5.31)$$

$$\left| \frac{(d_z)_i}{(z_k)_i} \right| \leq \kappa_1 \|d\| = O(\|H_k\|) + O(\mu_k), \quad (5.32)$$

e para cada i tal que $(\lambda_*^1)_i > 0$ ($(z_*)_i = 0$) temos

$$\frac{(d_z)_i}{(z_k)_i} = -1 + \frac{\mu_k}{(z_k)_i(\lambda_k^1)_i} + (q_k)_i, \quad |(q_k)_i| = O(\|H_k\|) + O(\mu_k), \quad (5.33)$$

$$\left| \frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} \right| \leq \kappa_1 \|d\| = O(\|H_k\|) + O(\mu_k). \quad (5.34)$$

Demonstração: As estimativas

$$\|d_N\| = \|(H'_k)^{-1}(-H_k)\| = O(\|H_k\|) \quad (5.35)$$

e

$$\|r_k\| = O(\|d_N\|^2) = O(\|H_k\|^2), \quad (5.36)$$

são essenciais para obter os resultados desejados e elas são consequências da definição (5.28) e do Lema 5.4.

De (5.35) e (5.36) temos

$$\|d\| = \|(H'_k)^{-1}(-H_k + r_k + \mu_k \hat{e})\| = O(\|H_k\|) + O(\mu_k). \quad (5.37)$$

No caso em que $(\lambda_*^1)_i = 0$, segue de A3 que $(z_*)_i > 0$.

Da definição de c_1 , existe $\epsilon_1 > 0$ suficientemente pequeno tal que, para $\|w_k - w_*\| \leq \epsilon_1$,

$$\left| \frac{(d_{Nz})_i}{(z_k)_i} \right| \leq c_1 \|d_N\|. \quad (5.38)$$

De (5.26) temos

$$\frac{(d_{N\lambda^1})_i}{(\lambda_k^1)_i} = -1 - \frac{(d_{Nz})_i}{(z_k)_i}. \quad (5.39)$$

Utilizando (5.38), (5.39) e (5.35) obtemos

$$\left| \frac{(d_{Nz})_i (d_{N\lambda^1})_i}{(z_k)_i (\lambda_k^1)_i} \right| = O(\|d_N\|) = O(\|H_k\|). \quad (5.40)$$

Agora, da definição de κ_1 e de (5.37), existe $\epsilon_2 > 0$ suficientemente pequeno tal que, para $\|w_k - w_*\| \leq \epsilon_2$,

$$\left| \frac{(d_z)_i}{(z_k)_i} \right| \leq \kappa_1 \|d\| = O(\|H_k\|) + O(\mu_k). \quad (5.41)$$

Utilizando (5.30), (5.40) e (5.41) segue que, para $\epsilon_0 = \min\{\epsilon_1, \epsilon_2\}$, se $\|w_k - w_*\| \leq \epsilon_0$, então

$$\frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} = -1 + \frac{\mu_k}{(z_k)_i (\lambda_k^1)_i} + (p_k)_i, \quad |(p_k)_i| = O(\|H_k\|) + O(\mu_k). \quad (5.42)$$

Suponha agora que $(\lambda_*^1)_i > 0$. Consequentemente $(z_*)_i = 0$ e, para $\epsilon_3 > 0$ suficientemente pequeno, se $\|w_k - w_*\| \leq \epsilon_3$ então

$$\left| \frac{(d_{N\lambda^1})_i}{(\lambda_k^1)_i} \right| \leq \kappa_1 \|d_N\|. \quad (5.43)$$

De (5.26) vem

$$\frac{(d_{Nz})_i}{(z_k)_i} = -1 - \frac{(d_{N\lambda^1})_i}{(\lambda_k^1)_i}. \quad (5.44)$$

Por (5.43), (5.44) e (5.35) obtemos

$$\left| \frac{(d_{Nz})_i (d_{N\lambda^1})_i}{(z_k)_i (\lambda_k^1)_i} \right| = O(\|d_N\|) = O(\|H_k\|). \quad (5.45)$$

Segue da definição de κ_1 e de (5.37) que existe $\epsilon_4 > 0$ suficientemente pequeno tal que, se $\|w_k - w_*\| \leq \epsilon_4$ então

$$\left| \frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} \right| \leq \kappa_1 \|d\| = O(\|H_k\|) + O(\mu_k). \quad (5.46)$$

De (5.30), (5.45) e (5.46), e definindo $\epsilon_0 = \min\{\epsilon_3, \epsilon_4\}$, se $\|w_k - w_*\| \leq \epsilon_0$ então

$$\frac{(d_z)_i}{(z_k)_i} = -1 + \frac{\mu_k}{(z_k)_i (\lambda_k^1)_i} + (q_k)_i, \quad |(q_k)_i| = O(\|H_k\|) + O(\mu_k). \quad (5.47)$$

□

Lema 5.12. *Sob as mesmas hipóteses do Lema 5.11, se $\gamma_k \geq O(\|H_k\|) + O(\mu_k)$, então*

$$1 \geq \alpha_k \geq \gamma_k - [O(\|H_k\|) + O(\mu_k)]. \quad (5.48)$$

Demonstração: Se $(d_z)_i \geq 0$ e $(d_{\lambda^1})_i \geq 0$ para todo i , então $(z_k + d_z, \lambda_k^1 + d_{\lambda^1}) > 0$ e o resultado é válido com $\alpha_k \equiv 1$.

Suponha que existe um índice i tal que $(d_{\lambda^1})_i < 0$. Se $(\lambda_*^1)_i > 0$, então de (5.34) e de $\gamma_k \geq O(\|H_k\|) + O(\mu_k)$ temos

$$-\gamma_k \frac{(\lambda_k^1)_i}{(d_{\lambda^1})_i} \geq \frac{\gamma_k}{O(\|H_k\|) + O(\mu_k)} \geq 1.$$

Se $(\lambda_*^1)_i = 0$, então de (5.31) obtemos

$$-\frac{(d_{\lambda^1})_i}{(\lambda_k^1)_i} = 1 - \frac{\mu_k}{(z_k)_i (\lambda_k^1)_i} - (p_k)_i \leq 1 + O(\|H_k\|) + O(\mu_k),$$

e assim,

$$-\gamma_k \frac{(\lambda_k^1)_i}{(d_{\lambda^1})_i} \geq \frac{\gamma_k}{1 + O(\|H_k\|) + O(\mu_k)} \geq \gamma_k (1 - [O(\|H_k\|) + O(\mu_k)]).$$

Utilizando o mesmos argumentos para d_z temos

$$\alpha_k \geq \min\{1, \gamma_k(1 - [O(\|H_k\|) + O(\mu_k)])\} = \gamma_k(1 - [O(\|H_k\|) + O(\mu_k)]) \geq \gamma_k - [O(\|H_k\|) + O(\mu_k)].$$

Portanto,

$$1 \geq \alpha_k \geq \gamma_k - [O(\|H_k\|) + O(\mu_k)].$$

□

Lema 5.13. *O Lema 5.9 é válido para uma sequência $\{w_k\}$ gerada pelo Algoritmo 2.*

Demonstração: É uma consequência direta do Lema 5.12. □

Teorema 5.14. *Suponha as mesmas hipóteses do Teorema 5.10 com $\{w_k\}$ uma sequência gerada pelo Algoritmo 2. Então, existe $\epsilon > 0$ tal que, para $\|w_0 - w_*\| < \epsilon$ a sequência $\{w_k\}$ está bem definida, $w_k \rightarrow w_*$ e*

$$\|w_{k+1} - w_*\| \leq \nu \|w_k - w_*\|^2,$$

para cada $k \geq 0$, onde ν é uma constante positiva.

Demonstração: Supondo que $\|w_k - w_*\| < \epsilon$. Lembrando que H'_k é não singular, então $\|(H'_k)^{-1}\| \leq \tau$, e assim

$$\begin{aligned}
 w_{k+1} - w_* &= w_k - w_* + \alpha_k d_k \\
 &= w_k - w_* + \alpha_k (H'_k)^{-1} (-H_k + r_k + \mu_k \hat{e}) + \alpha_k (w_k - w_*) - \alpha_k (w_k - w_*) \\
 &= (1 - \alpha_k)(w_k - w_*) + \alpha_k (H'_k)^{-1} (-H_k + H_* + H'_k(w_k - w_*)) \\
 &\quad + \alpha_k (H'_k)^{-1} r_k + \alpha_k (H'_k)^{-1} \mu_k \hat{e}.
 \end{aligned}$$

Pelos Lemas 5.4 e 5.13 e utilizando (5.36):

$$\begin{aligned}
 \|w_{k+1} - w_*\| &\leq [(1 - \gamma_k) + O(\|H_k\|) + O(\mu_k)] \|w_k - w_*\| \\
 &\quad + O(\|w_k - w_*\|^2) + O(\|H_k\|^2) + O(\mu_k).
 \end{aligned}$$

Como $O(\|H_k\|^2) = O(\|w_k - w_*\|^2)$ e das hipóteses, obtemos

$$\|w_{k+1} - w_*\| = O(\|w_k - w_*\|^2).$$

□

No caso em que as funções F e G que definem o GNCP são afins, o método preditor-corretor do Algoritmo 2 converte-se no método preditor-corretor de Mehrotra (como apresentado em [155]), com as devidas escolhas dos parâmetros. De fato, para F e G afins, a expressão (5.28) torna-se

$$\tilde{r}_k = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -D_{N_z} D_{N_{\lambda_1}} e \end{bmatrix}. \tag{5.49}$$

Para o caso não linear, vale a pena notar que definindo o lado direito do sistema (5.27) do Passo 3 do Algoritmo 2 com o vetor (5.49) ao invés de (5.28), os resultados e demonstrações continuam válidos, pois a relação $\|\tilde{r}_k\| = O(\|d_N\|^2)$ ainda é válida. Tal adaptação gera uma versão de menor custo do algoritmo para problemas não lineares, que nós chamamos de Algoritmo 2 econômico, que evita o cálculo

das derivadas de segunda ordem e dos produtos tensoriais. O desempenho deste algoritmo econômico é analisado comparativamente no segundo conjunto de experimentos da próxima seção.

5.4 Experimentos Numéricos

Os Algoritmos 1 e 2 foram implementados em Matlab 7.0.1 R14, e foi utilizado um computador com um processador Phenom X4 9550 2.2 GHz, 4 GB de memória RAM e sistema operacional Windows Vista Business 64.

Os critérios de parada dos algoritmos foram a obtenção de uma solução do sistema (5.4), isto é $\|H(w_k)\|_1 \leq 10^{-10}$ (Flag 1), dois iterandos consecutivos muito próximos, $\|w_k - w_{k-1}\|_1 \leq 10^{-10}$ (Flag 2), ou realização do número máximo de iterações, 150 para o Algoritmo 1 e 100 para o Algoritmo 2, indicado por Flag 3 quando necessário.

Observação 5.15. *Nos critérios de parada foi utilizado erro absoluto ao invés de erro relativo pois os dados dos experimentos foram gerados com distribuição uniforme. A diferença entre os valores máximos de iteração para os algoritmos se deve ao fato de que, em geral, algoritmos do tipo preditor-corretor necessitam de um número menor de iterações para encontrar uma solução do referido problema, com precisão dada.*

Para calcular o parâmetro associado ao tamanho do passo a ser obtido utilizamos uma expressão similar a dada em [44], a saber, $\gamma_k = 1 - \min\{0.005, 100z_k^T \lambda_k^1\}$.

Lembrando que q é a dimensão dos vetores z e λ^1 , o parâmetro μ_k que fornece a perturbação nos sistemas de ambos os algoritmos foi calculado como:

$$\begin{aligned} \text{Newton Perturbado: } \mu_k &= \frac{z_k^T \lambda_k^1}{q^\beta}, \\ \text{Preditor-corretor: } \mu_k &= \begin{cases} \left(\frac{z_k^T \lambda_k^1}{q^\beta} \right)^2, & \text{if } z_k^T \lambda_k^1 < 1, \\ \left(\frac{\bar{z}_k^T \bar{\lambda}_k^1}{z_k^T \lambda_k^1} \right)^3 \left(\frac{z_k^T \lambda_k^1}{q^\beta} \right), & \text{caso contrário} \end{cases} \end{aligned}$$

onde

$$\bar{z}_k = z_k + \bar{\alpha}d_{N_z} \quad \text{e} \quad \bar{\lambda}_k^1 = \lambda_k^1 + \bar{\alpha}d_{N_{\lambda^1}}.$$

As expressões “standard” μ_k e “modificado” μ_k correspondem às escolhas $\beta = 1$ e $\beta = 2$, respectivamente, iremos deixar clara a escolha utilizada quando apresentarmos os resultados numéricos. As expressões acima para μ_k têm semelhança com as adotadas para resolver problemas de programação linear (ver [116]).

Os quatro grupos de problemas que compõem nosso conjunto de testes também foram considerados em [6, 7].

Observação 5.16. *A notação $\text{rnd}(a, b)$ significa que os dados foram gerados aleatoriamente entre a e b com distribuição uniforme.*

5.4.1 Problemas gerados aleatoriamente no ortante positivo

No primeiro conjunto de experimentos consideramos o problema de encontrar $x \in \mathbb{R}^n$ tal que $Mx + c \geq 0$, $Px + d \geq 0$ e $(Mx + c)^T(Px + d) = 0$, onde as matrizes $M, P \in \mathbb{R}^{n \times n}$ e os vetores $c, d \in \mathbb{R}^n$ são dados. Com a notação \mathcal{K} dada por (5.1), temos $A = I_n$ e $B = 0$. Este problema é levemente mais geral que um LCP e mesmo para operadores singulares ou não monótonos M e/ou P , os exemplos gerados têm solução.

Iniciamos com uma aproximação inicial gerada aleatoriamente do seguinte modo: $(x_0)_i = \text{rnd}(-10, 10)$, $(z_0)_i = \text{rnd}(1, 10)$ $(\lambda_0)_i = \text{rnd}(1, 10)$, $i = 1, \dots, n$.

No cálculo do parâmetro μ_k , utilizamos a opção “modificado” para o método de Newton (a opção “standard” não foi suficiente para obter solução dos problemas) e a escolha “standard” para o algoritmo preditor-corretor.

De acordo com as características das matrizes M e P , dividimos o conjunto de testes em nove diferentes famílias, numeradas como em [6, 7]: M e P podem ser iguais (1, 2, 3, 4 e 6) ou não (7, 10, 11 e 14); M e P podem ser simétricas (1, 2, 3, 7 e 10) ou não (4, 6, 11 e 14) e as matrizes M e P podem ser inversíveis (1, 2, 4, 7 e 11) ou singulares (3, 6, 10 e 14).

Para cada família foram utilizados dois valores para a dimensão n (5 e 500). Para cada dimensão, três problemas foram resolvidos, com diferentes pontos iniciais (a, b e c conforme Apêndice B), com ambos os métodos: Newton (N) e o preditor-corretor (PC). Os resultados estão compilados na Tabela 5.1, onde temos as triplas, com os valores mínimo, médio e máximo, do número de iterações (ITER) e do tempo de CPU, em segundos, (CPU) requeridos por cada método. A partir dessas triplas podemos avaliar precisamente a variabilidade dos resultados obtidos. Em geral, neste conjunto de problemas teste, o custo de uma iteração do preditor-corretor, em termos de tempo de CPU, é o dobro do custo de uma iteração de Newton, para qualquer critério de parada.

Para as cinco famílias geradas com matrizes inversíveis (respec. 1, 2, 4, 7 e 11), as hipóteses teóricas são satisfeitas e ambos os métodos obtêm sucesso nas suas resoluções, com o critério de parada indicado por **Flag 1**, exceto para um caso da família 11 e dimensão $n = 500$ que parou com **Flag 3**. Vale observar que para a família 2, as matrizes $M = P$ são indefinidas mas inversíveis, e assim $M^T P = M^T M = M^2$ é definida positiva.

Para a família 10, com matrizes semidefinidas e singulares, os casos com dimensão menor foram resolvidos com sucesso (**Flag 1**), enquanto os maiores pararam devido a falta de progresso (**Flag 2**), para ambos os métodos, com $\|H_k\|_1$ pelo menos da ordem 10^{-7} (para os valores exatos ver Tabelas B.1–B.2 do Apêndice B). Para as outras famílias com matrizes singulares, a saber 3, 6 e 14, geralmente a parada ocorreu pelo número máximo de iterações (**Flag 3**).

5.4.2 Problemas de complementaridade implícita

No segundo conjunto de experimentos resolvemos problemas de complementaridade implícita, como definidos em (ver [128]), da forma: encontrar $x \in \mathbb{R}^n$ tal que

$$x - m(x) \geq 0, \quad F(x) \geq 0 \quad \text{e} \quad F(x)^T(x - m(x)) = 0, \quad (5.50)$$

Características do problema	n	Método	ITER			CPU		
			min.	med.	max.	min.	med.	max.
Família 1 $M = P$ $M = M^T$ $M > 0$	5	N	19	19.0	19	0.003	0.008	0.017
		PC	7	7.0	7	0.002	0.007	0.016
	500	N	14	14.0	14	2.544	2.693	2.768
		PC	8	8.0	8	2.757	2.911	3.013
Família 2 $M = P$ $M = M^T$ M indefinida	5	N	19	19.0	19	0.003	0.008	0.017
		PC	7	7.0	7	0.001	0.001	0.002
	500	N	14	14.0	14	2.621	2.686	2.763
		PC	8	8.0	8	2.967	2.987	3.005
Família 3 $M = P$ $M = M^T$ $M \geq 0$	5	N	150	150.0	150	1.446	1.544	1.639
		PC	100	100.0	100	1.620	1.952	2.162
	500	N	150	150.0	150	28.88	30.41	33.28
		PC	100	100.0	100	38.84	39.14	39.40
Família 4 $M = P$ $M \neq M^T$ $M > 0$	5	N	19	19.0	19	0.001	0.002	0.002
		PC	7	7.0	7	0.001	0.002	0.003
	500	N	14	14.0	14	1.545	1.670	1.738
		PC	19	19.0	19	1.841	1.898	1.993
Família 6 $M = P$ $M \neq M^T$ $M \geq 0$	5	N	150	150.0	150	1.087	1.436	1.805
		PC	100	100.0	100	1.661	1.735	1.842
	500	N	150	150.0	150	17.40	18.07	19.36
		PC	100	100.0	100	22.98	23.13	23.43
Família 7 $M \neq P$ $M = M^T > 0$ $P = P^T > 0$	5	N	19	19.0	19	0.002	0.002	0.002
		PC	7	7.0	7	0.001	0.001	0.002
	500	N	14	15.3	16	1.646	1.716	1.769
		PC	8	8.3	9	1.814	1.940	2.127
Família 10 $M \neq P$ $M = M^T \geq 0$ $P = P^T \geq 0$	5	N	19	19.3	20	0.002	0.003	0.004
		PC	7	7.7	8	0.001	0.002	0.003
	500	N	25	34.0	39	3.054	4.179	4.758
		PC	11	14.0	16	2.815	3.367	3.934
Família 11 $M \neq P$ $M \neq M^T > 0$ $P \neq P^T > 0$	5	N	20	20.3	21	0.002	0.003	0.004
		PC	11	12.7	15	0.002	0.003	0.005
	500	N	27	71.0	150	3.128	8.776	18.76
		PC	22	74.0	100	5.223	17.26	23.59
Família 14 $M \neq P$ $M \neq M^T \geq 0$ $P \neq P^T \geq 0$	5	N	21	64.0	150	0.003	0.008	0.017
		PC	11	41.0	100	0.004	0.013	0.031
	500	N	150	150.0	150	17.47	18.35	18.81
		PC	100	100.0	100	23.13	23.75	24.29

Tabela 5.1: Problemas gerados aleatoriamente no ortante positivo.

onde $m_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, n$,

$$F(x) = Mx + c = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & \ddots & 0 \\ 0 & \ddots & \ddots & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (5.51)$$

e $m(x) = \varphi(Mx + c)$, com $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe C^2 .

Observação 5.17. *Note que o problema (5.50) é mais geral que o definido em [129, 130], onde $G(x) = x - F(x)$.*

Como nos exemplos 4.3 e 4.4 de [128], as seguintes escolhas para as funções φ definem nossos problemas teste: POZ1: $\varphi_i(\lambda) = -0.5 - \lambda_i, i = 1, \dots, n$ e POZ2: $\varphi_i(\lambda) = -0.5 - 1.5\lambda_i + 0.25\lambda_i^2, i = 1, \dots, n$. Como no primeiro conjunto de experimentos, com a notação \mathcal{K} de (5.1), temos $A = I_n$ e $B = 0$.

Para cada problema, três pontos iniciais diferentes foram utilizados, a saber, $(x_0)_i = 0.0, (x_0)_i = -1.0$ e $(x_0)_i = -0.5, i = 1, \dots, n$, indicados respectivamente por a,b,c nas Tabelas B.3–B.4 do Apêndice B. Para estas três escolhas, definimos $(z_0)_i = (\lambda_0^1)_i = \max\{(F(x_0))_i, 1\}, i = 1, \dots, n$.

Devido à estrutura esparsa dos operadores que definem os problemas, exemplos com dimensão maior que as consideradas em [6, 7] foram resolvidos. Como no primeiro conjunto de experimentos, utilizamos a opção “modificado” para atualizar μ_k no método de Newton e “standard” para o preditor-corretor.

Para o problema POZ1, comparamos o método de Newton novamente com o preditor-corretor definido pelo Algoritmo 2. Para o problema POZ2, a leve não linearidade da função φ proporcionou condições para avaliar o desempenho do Algoritmo 2 econômico, onde definimos o lado direito do sistema (5.27) do Passo 3 do Algoritmo 2 com o vetor (5.49) ao invés de (5.28). Os resultados estão relatados nas Tabelas 5.2 e 5.3, que seguem o mesmo esquema da Tabela 5.1. Na Tabela 5.3, PCM denota o preditor-corretor de Mehrotra (Algoritmo 2 econômico). Para os resultados obtidos com cada um dos pontos iniciais ver Tabelas B.3–B.4 do Apêndice B.

n	Método	ITER			CPU		
		min.	med.	max.	min.	med.	max.
4	N	19	19.0	19	0.006	0.007	0.008
	PC	6	6.0	6	0.003	0.004	0.006
40	N	9	9.0	9	0.004	0.007	0.009
	PC	5	5.0	5	0.003	0.006	0.007
400	N	7	7.0	7	0.190	0.269	0.314
	PC	6	6.0	6	0.172	0.218	0.257
4000	N	7	7.0	7	24.18	24.58	25.24
	PC	6	6.0	6	20.48	20.96	21.89
8000	N	7	7.0	7	131.0	134.7	139.0
	PC	6	6.0	6	114.0	119.3	123.0
16000	N	7	7.0	7	579.0	579.7	580.0
	PC	6	6.0	6	485.0	499.3	508.0

Tabela 5.2: Resultados do problema de complementaridade implícita POZ1.

Da Tabela 5.2, observamos que o algoritmo preditor-corretor é competitivo com Newton para os casos envolvendo o problema POZ1. O tempo de CPU gasto por iteração é semelhante para ambos os métodos, e o preditor-corretor mostra uma leve vantagem em termos do número de iterações. Todos os problemas pararam com sucesso, isto é **Flag 1**.

Com respeito ao problema de complementaridade implícita POZ2, observamos pela Tabela 5.2 que o **Algoritmo 2** e o **Algoritmo 2 econômico** têm comportamento similar. Parece que a não linearidade presente em POZ2 não é forte o suficiente para influenciar nos resultados. Com a notação do problema (4.7), a função G é quadrática e F é afim. Também neste caso, ambos os algoritmos são competitivos com Newton. A razão do tempo de CPU pelo número de iterações é bem similar para os três métodos, com uma leve vantagem para o preditor-corretor em termos do número

de iterações. Todos os problemas foram resolvidos com sucesso, parada com **Flag 1**.

Para investigar melhor a influência da não linearidade na classe dos problemas de complementaridade implícita, e seu impacto quando tais problemas são resolvidos pelos **Algoritmos 2 e 2 econômico**, foi criado um terceiro conjunto de problemas-teste, a saber PCIM: $\varphi_i(\lambda) = -0.5 - 1.5\lambda_i + 0.25\lambda_i^2 + \lambda_i^3, i = 1, \dots, n$.

Em testes preliminares com a formulação (5.50) observamos que o mal-condicionamento da matriz M , que define a função F em (5.51), tinha grande influência no desempenho dos **Algoritmos 1, 2 e 2 econômico** em outras escolhas não lineares para a função φ , especialmente quando a dimensão n aumentava. Portanto, fizemos uma leve modificação na função $G(x) = x - m(x)$ da formulação (5.50), adotando $G(x) = 100x - m(x)$ para contornar o mal-condicionamento dos operadores envolvidos e de algum modo evitar essa influência nos resultados. Nos testes preliminares também observamos que para prevenir o excesso de paradas prematuras devido a dois iterandos consecutivos estarem muito próximos, $\|w_k - w_{k-1}\|_1 \leq \varepsilon$ (**Flag 2**), a tolerância ε foi alterada de $\varepsilon = 10^{-10}$ para $\varepsilon = 10^{-12}$, mas somente para esta instância. Os resultados estão relatados na Tabela 5.4, que segue o mesmo esquema da Tabela 5.3. Com a modificação da tolerância, mencionada acima, todas os problemas pararam com **Flag 1**. Apesar do problema gerado com a função φ no PCIM ser mais não linear que a função do POZ2, da Tabela 5.4 podemos observar que não houve diferença significativa nos resultados dos **Algoritmos 2 e 2 econômico**. Além disso, para esta instância é evidente que o desempenho dos métodos preditores-corretores é bastante competitivo com o do método de Newton perturbado.

Nossos resultados para os problemas POZ2 e PCIM são similares aos obtidos por Gundersen e Steihaug [68] para minimização irrestrita da função de Rosembrock, na qual a Hessiana é tridiagonal. Eles obtiveram o limitante teórico de 5 para a razão entre o custo por iteração do método de Chebyshev e do método de Newton, independente da dimensão, para uma grande classe de problemas esparsos. Na prática, eles atingiram um limitante bem menor, devido à especialização do código para a estrutura esparsa do problema, como ocorreu em nossos testes.

5.4.3 Problemas com cones gerais em \mathbb{R}^n

No terceiro conjunto de experimentos os problemas são: encontrar $x \in \mathbb{R}^n$ tal que $Mx + c \in \mathcal{K}$, $Px + d \in \mathcal{K}^\circ$ e $(Mx + c)^T(Px + d) = 0$, onde os conjuntos \mathcal{K} , \mathcal{K}° são definidos por (5.1), (5.2), com $A \in \mathbb{R}^{q \times n}$, $B \in \mathbb{R}^{s \times n}$ matrizes dadas. As matrizes $M, P \in \mathbb{R}^{n \times n}$ e os vetores $c, d \in \mathbb{R}^n$ também são dados.

Os problemas foram gerados aleatoriamente, de um modo bem semelhante ao primeiro conjunto de experimentos. Para maiores detalhes, ver [7]. É importante salientar que estes problemas englobam matrizes não nulas B na definição dos cones \mathcal{K} e \mathcal{K}° , dados (5.1) e (5.2), respectivamente.

As aproximações iniciais foram geradas aleatoriamente do seguinte modo: $(x_0)_i = \text{rnd}(-10, 10)$, $i = 1, \dots, n$, $(z_0)_i = \text{rnd}(1, 10)$ e $(\lambda_0^1)_i = \text{rnd}(1, 10)$, $i = 1, \dots, p$ e $(\lambda_0^2)_i = 0$, $i = 1, \dots, s$.

De acordo com as características das matrizes M e P , dividimos este conjunto de testes em três famílias: 1) $M = P$, indefinida e não simétrica; 2) $M \neq P$, indefinida e simétrica; 3) $M \neq P$, indefinida, não simétrica e singular.

Para estas três famílias, a solução gerada satisfaz a complementaridade estrita, hipótese A4. Para analisar a influência de tal hipótese no desempenho dos métodos, uma quarta família de experimentos foi introduzida, com as mesmas características da primeira família, mas neste caso a solução gerada não satisfaz a complementaridade estrita.

Note que a hipótese teórica A2 de monotonia e regularidade é válida para as famílias 1) e 4): visto que $M = P$, segue que $M^T P = M^T M$ é semidefinida positiva, e como também é não singular, a referida matriz é definida positiva. A hipótese A2 também é verificada pela família 2), pois $M = P = M^T$ é indefinida e não singular, logo $M^T P = M^2$ é definida positiva.

Para cada família, seis combinações de dimensões (n, q, s) foram consideradas: (10, 5, 1); (10, 15, 1); (100, 50, 5); (100, 100, 5); (100, 150, 5) e (500, 250, 25). Para cada uma delas, três problemas foram resolvidos, com diferentes pontos

iniciais (indicados por a,b,c nas Tabelas B.6–B.12 do Apêndice B). Os resultados estão relatados nas Tabelas 5.5–5.8, que seguem o mesmo padrão das tabelas anteriores.

Como no primeiro conjunto de experimentos, utilizamos a opção “modificada” para atualizar μ_k no método de Newton e a escolha “standard” para o preditor-corretor. Para permitir uma perspectiva comparativa, incluímos os resultados com a opção “modificada” na escolha de μ_k no método preditor-corretor, denotando, nas tabelas, os métodos preditor-corretor por PCstd e PCmod, respectivamente.

Analisando as Tabelas 5.5 e 5.6 comparativamente, notamos que a simetria não produz impacto significativo nos resultados. Newton e as variantes preditor-corretor obtiveram sucesso, como esperado, pois as Famílias 1 e 2 satisfazem as hipóteses teóricas. Dos 54 casos (6 conjuntos de dimensões \times 3 pontos iniciais \times 3 métodos) que geraram os resultados em cada tabela, a parada com **Flag 1** aconteceu em 40 problemas na Tabela 5.5 e também na Tabela 5.6. Para os problemas com dimensões maiores, ambos tiveram várias paradas com **Flag 2**, e a ordem dos valores de $\|H_k\|_1$ na última iteração variou entre 10^{-6} e 10^{-10} (para os valores exatos ver Tabelas B.6,B.9 no Apêndice B).

Quando se trata da Família 3, cujos resultados então relatados na Tabela 5.7, observamos que a singularidade causa um grande inconveniente para todos os métodos, de modo que para a grande maioria dos problemas os algoritmos pararam com **Flag 3**. Para os problemas com menor dimensão, os três métodos encontraram uma solução do problema.

No que diz respeito à Família 4, para a qual a solução gerada não satisfaz a complementaridade estrita, hipótese A4, dos resultados da Tabela 5.8 observamos que tal hipótese parece não influenciar na convergência de qualquer um dos métodos. De fato, o desempenho dos métodos é bastante similar ao apresentado para a Família 1. As mesmas observações feitas anteriormente valem neste caso: 42 dos casos pararam com **Flag 1**, e 12 pararam com **Flag 2** (a ordem dos valores de $\|H_k\|_1$ na última iteração variou entre 10^{-6} e 10^{-10}), no total de 54 casos. Para mais detalhes ver também as Tabelas B.11–B.12 no Apêndice B.

Na comparação de “standard” *versus* “modificado” para atualizar μ_k no método preditor-corretor, não observamos evidências significativas que apontem vantagens de uma escolha sobre a outra, (ver Tabelas 5.5–5.8). Considerando este conjunto de experimentos, vale também mencionar que, apesar do número de iterações das variantes preditor-corretor serem menores que do método de Newton, o método de Newton exibe uma leve vantagem em termos de tempo de CPU por iteração.

5.4.4 Problemas em cones 3D com controle de faces geradas

No quarto conjunto de experimentos abordamos o problema de encontrar $x \in \mathcal{K} = \{v \in \mathbb{R}^n \mid Av \geq 0\}$ tal que $Tx + c \in \mathcal{K}^\circ = \{v \in \mathbb{R}^n \mid A^T \lambda = v, \lambda \geq 0\}$. Geramos os cones poliedrais \mathcal{K} com q faces, tais que suas arestas são as retas

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos\left(\frac{2\pi}{q}k\right) \\ r \sin\left(\frac{2\pi}{q}k\right) \\ 1 \end{pmatrix} t, \quad t \in \mathbb{R}, \quad k = 1, \dots, q.$$

Portanto, \mathcal{K} foi definido computando as linhas da matriz A como sendo os vetores normais ao plano suporte das faces do cone. Em outras palavras, o vetor que define a i -ésima linha da matriz A ($i = 1, \dots, q$) é dado pelo produto vetorial:

$$\begin{pmatrix} \cos\left(\frac{2\pi}{q}(i-1)\right) \\ \sin\left(\frac{2\pi}{q}(i-1)\right) \\ \frac{1}{r} \end{pmatrix} \times \begin{pmatrix} r \cos\left(\frac{2\pi}{q}i\right) \\ r \sin\left(\frac{2\pi}{q}i\right) \\ 1 \end{pmatrix} = \begin{pmatrix} \sin\left(\frac{2\pi}{q}i\right)\left(\cos\frac{2\pi}{q} - 1\right) - \cos\left(\frac{2\pi}{q}i\right)\sin\frac{2\pi}{q} \\ \cos\left(\frac{2\pi}{q}i\right)\left(1 - \cos\frac{2\pi}{q}\right) - \sin\left(\frac{2\pi}{q}i\right)\sin\frac{2\pi}{q} \\ r \sin\frac{2\pi}{q} \end{pmatrix}$$

Os problemas foram gerados do seguinte modo. Uma vez definido o valor do raio r e da dimensão q (número de faces do cone \mathcal{K}), construímos a matriz A e definimos dois tipos de x_* , um na borda e outro no interior de \mathcal{K} , respectivamente. Depois foi gerada uma matriz T mantendo a simetria, e então definimos três famílias de problemas, numeradas de acordo com [6, 7], a saber: 1) T indefinida; 2) T definida positiva, e 4) T semidefinida negativa e singular. Para mais detalhes, ver [7]. Observe que os Algoritmos 1 e 2 estão bem definidos somente para a Família 2.

Os testes foram construídos variando $r \in \{0.1, 10\}$, $q \in \{3, 12, 20, 40\}$, as três famílias de matrizes T e os dois tipos de solução gerada x_* . Como estes casos são de dimensão pequena, e geralmente resolvidos rapidamente, adotamos dimensões q maiores que as utilizadas em [6, 7], assim as medidas de tempo de CPU aumentaram e puderam ser detectadas com mais precisão.

Foram gerados três pontos iniciais diferentes, utilizando $(x_0)_i = \text{rnd}(-10, 10)$, $i = 1, \dots, n$, $(z_0)_i = \text{rnd}(1, 10)$ e $(\lambda_0^1)_i = \text{rnd}(1, 10)$, $i = 1, \dots, q$, para cada escolha de r, q, T e x_* . Os problemas foram resolvidos pelo método de Newton com a opção “modificada” para atualizar μ_k , e para o método preditor-corretor foi utilizada a escolha “standard” para μ_k .

Quanto às potenciais causas de resultados distintos, a saber, raio r pequeno ou grande, solução gerada no interior do cone \mathcal{K} , ou na sua borda, e o tamanho dos problemas, dependendo do número de faces q , organizamos a apresentação dos dados de acordo com a estrutura da matriz T . As Tabelas 5.9–5.11 contém o mínimo, médio e máximo dos resultados obtidos para os diferentes pontos iniciais gerados aleatoriamente, para cada família, seguindo o mesmo padrão das tabelas anteriores. Cada tabela contém 96 casos (4 escolhas para o número de faces $q \times 2$ tipos de soluções geradas $\times 2$ escolhas para o raio $r \times 2$ métodos $\times 3$ pontos iniciais).

Analisando os resultados da Tabela 5.9, observamos que apesar da matriz T não ser monótona para os problemas da Família 1, a regularidade é provavelmente decisiva para a obtenção de sucesso dos algoritmos, neste caso. De fato, considerando os casos com solução gerada no interior do cone, somente 4 dos 48 casos pararam com **Flag 3**, e para os casos com solução na borda do cone, somente 1 dos 48 parou com **Flag 3**; todos os demais pararam com **Flag 1**. Vale mencionar que para dois casos com solução gerada no interior do cone, o método de Newton atingiu o número máximo de iterações, enquanto o preditor-corretor encontrou uma solução, com 12 e 14 iterações, respectivamente (ver Tabela B.13 no Apêndice B).

Observamos que para as dimensões $q = 3$ e $q = 12$, os problemas gerados com raio $r = 0.1$ foram levemente mais difíceis. Quando aumenta-se q , o tamanho do raio parece não ter uma influência significativa nos resultados. Em geral, a localização da solução gerada também não gera diferenças significativas nos resultados.

Sem dúvida o cumprimento das hipóteses do Corolário 5.3 pela Família 2 é determinante para atingir o resultado de 100% de sucesso para ambos os métodos, como pode ser

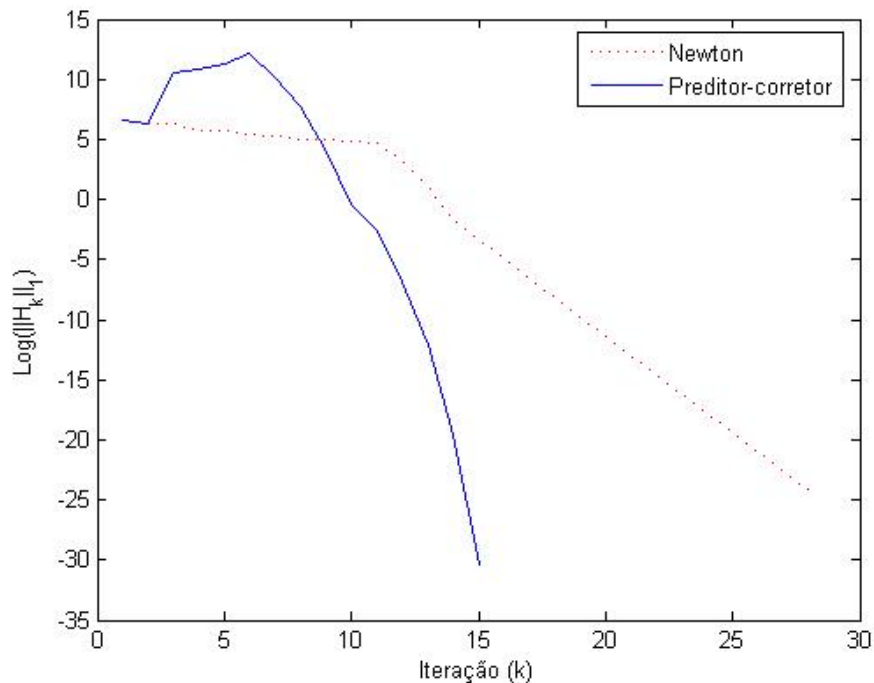
verificado na Tabela 5.10. Todos os problemas pararam com **Flag 1**. O algoritmo preditor-corretor sempre obteve uma solução com menos iterações que Newton. O custo, em termos de tempo de CPU por iteração, variou e em 18% dos casos (17 de 96), o preditor-corretor exigiu menos tempo que Newton por iteração. Para os problemas com menor número de faces, a diferença entre o desempenho de ambos os métodos foi mais significativa. Com o aumento do número de faces, essa diferença diminuiu. Por outro lado, no que diz respeito à localização da solução gerada, os problemas com solução gerada na borda do cone são levemente mais difíceis que aqueles cujas soluções estão no interior do cone, especialmente quando se aumenta o número de faces.

Em se tratando da Família 4, devido à singularidade da matriz T , houve vários casos de insucesso, parada com **Flag 3**, a saber 27% (26 de 96). Destes 26 casos, 22 correspondem a 11 casos nos quais ambos os métodos falharam. Nos 4 casos restantes, em 1 Newton falhou e o preditor-corretor obteve sucesso, com 8 iterações, enquanto nos outros 3, ocorreu o contrário. Geralmente o preditor-corretor demanda menos iterações que Newton.

5.4.5 Comentários gerais sobre os experimentos numéricos

Para os 4 diferentes grupos de problemas foram gerados 108 problemas teste, variando dimensão e propriedades das matrizes Jacobianas. Para cada um destes 108 problemas foram utilizados 3 pontos iniciais distintos, gerados aleatoriamente conforme especificado nas descrições dos problemas, resultando em um total de 324 instâncias. Quanto à robustez, o método de Newton (**Algoritmo 1**) resolveu 271 problemas (84%), enquanto o método preditor-corretor (**Algoritmo 2**) resolveu 268 problemas (83%). Todas as hipóteses foram satisfeitas por 168 dos 324 problemas, destes o método de Newton resolveu 167 e o preditor-corretor resolveu 166. Com relação aos 156 problemas restantes, nos quais pelo menos uma hipótese falhou, Newton resolveu 92 (59%) e o preditor-corretor resolveu 90 (58%).

Quanto ao tempo de CPU, de um modo geral os métodos obtiveram resultados similares para o segundo (problemas de complementaridade implícita) e quarto (cones em \mathbb{R}^3) grupos de experimentos, o método de Newton obteve vantagem sobre o preditor-corretor no primeiro grupo (problemas gerados no ortante positivo) de experimentos e a vantagem se inverteu no terceiro grupo (cones gerais em \mathbb{R}^n) de experimentos.

Figura 5.1: Família 4, $(n, q, s) = (10, 15, 1)$ e ponto inicial **a**

Um observação interessante foi que todos os problemas sem complementaridade estrita na solução foram resolvidos por ambos os métodos. Esses problemas são todos do terceiro grupo de problemas teste. Neste caso, houve uma clara vantagem do preditor-corretor sobre Newton, em termos de número de iterações e também em termos de tempo de CPU. Analisando $\|H_k\|_1$ em cada iteração pudemos observar o fenômeno ilustrado na Figura 5.1, onde temos os gráficos de $\text{log}(\|H_k\|_1)$ para uma instância do terceiro grupo de problemas teste, sem complementaridade estrita na solução. Este fenômeno ocorreu em geral para os problemas da Família 4.

n	Método	ITER			CPU		
		min.	med.	max.	min.	med.	max.
4	N	19	19.0	19	0.003	0.004	0.006
	PC	6	6.0	6	0.002	0.003	0.004
	PCM	6	6.0	6	0.002	0.003	0.003
40	N	9	9.0	9	0.005	0.007	0.009
	PC	5	5.0	5	0.004	0.005	0.007
	PCM	5	5.0	5	0.003	0.004	0.007
400	N	7	7.0	7	0.230	0.215	0.225
	PC	6	6.0	6	0.171	0.187	0.203
	PCM	6	6.0	6	0.172	0.183	0.191
4000	N	7	7.0	7	23.50	23.83	24.10
	PC	6	6.0	6	19.52	21.10	22.39
	PCM	6	6.0	6	20.71	21.32	22.15
8000	N	7	7.0	7	131.0	132.0	133.0
	PC	6	6.0	6	108.0	109.3	111.0
	PCM	6	6.0	6	115.0	115.3	116.0
16000	N	7	7.0	7	562.0	566.0	572.0
	PC	6	6.0	6	494.0	512.3	524.0
	PCM	6	6.0	6	492.0	500.7	510.0

Tabela 5.3: Resultados do problema de complementaridade implícita POZ2.

5.4. EXPERIMENTOS NUMÉRICOS

n	Método	ITER			CPU		
		min.	med.	max.	min.	med.	max.
4	N	18	18.0	18	0.003	0.005	0.006
	PC	4	4.7	5	0.003	0.003	0.003
	PCM	4	4.7	5	0.001	0.002	0.003
40	N	8	8.0	8	0.017	0.019	0.020
	PC	5	5.0	5	0.015	0.015	0.016
	PCM	5	5.0	5	0.015	0.015	0.016
400	N	6	6.3	7	0.187	0.224	0.281
	PC	5	5.3	6	0.156	0.187	0.234
	PCM	5	5.7	6	0.156	0.174	0.196
4000	N	6	6.7	7	26.00	27.67	29.03
	PC	6	6.0	6	24.93	25.23	25.79
	PCM	6	6.0	6	24.82	25.04	25.27
8000	N	6	6.7	7	129.9	144.4	152.7
	PC	6	6.0	6	130.9	131.6	132.7
	PCM	6	6.0	6	131.5	132.0	132.5
16000	N	6	6.7	7	529.0	593.6	631.5
	PC	6	6.0	6	534.3	540.2	546.8
	PCM	6	6.0	6	513.7	531.4	550.8

Tabela 5.4: Resultados do problema de complementaridade implícita modificado PMIC.

(n, q, s)	Método	ITER			CPU		
		min.	med.	max.	min.	med.	max.
(10,5,1)	N	25	25.0	25	0.014	0.015	0.015
	PCstd	11	13.7	18	0.011	0.013	0.015
	PCmod	10	12.3	17	0.010	0.013	0.015
(10,15,1)	N	29	29.3	30	0.047	0.057	0.063
	PCstd	11	12.3	14	0.015	0.026	0.031
	PCmod	12	12.3	13	0.031	0.031	0.032
(100,50,5)	N	53	59.0	64	0.702	0.884	1.061
	PCstd	16	18.0	20	0.515	0.598	0.702
	PCmod	20	22.0	24	0.562	0.660	0.780
(100,100,5)	N	55	62.3	67	2.464	2.704	2.871
	PCstd	16	16.3	17	0.874	0.952	1.076
	PCmod	20	21.7	23	1.123	1.217	1.311
(100,150,5)	N	90	93.0	97	4.970	4.996	5.026
	PCstd	23	25.0	28	2.037	2.167	2.370
	PCmod	27	31.3	37	2.481	2.798	3.198
(500,250,25)	N	111	126.0	135	273.0	292.3	311.0
	PCstd	24	28.3	34	80.00	86.73	91.00
	PCmod	36	41.7	47	125.0	142.3	163.0

Tabela 5.5: Resultados da Família 1 de problemas com cones poliedrais em \mathbb{R}^n (matrizes indefinidas e não simétricas).

5.4. EXPERIMENTOS NUMÉRICOS

(n, q, s)	Método	ITER			CPU		
		min.	med.	max.	min.	med.	max.
(10,5,1)	N	21	21.0	21	0.012	0.014	0.016
	PCstd	8	10.3	14	0.006	0.009	0.012
	PCmod	8	9.0	11	0.010	0.012	0.013
(10,15,1)	N	25	26.3	29	0.046	0.052	0.062
	PCstd	12	12.7	14	0.031	0.036	0.046
	PCmod	14	17.7	20	0.015	0.138	0.251
(100,50,5)	N	60	63.0	66	0.811	0.920	0.998
	PCstd	16	17.3	19	0.531	0.598	0.640
	PCmod	19	21.7	24	0.484	0.603	0.702
(100,100,5)	N	80	82.0	84	3.369	3.635	3.869
	PCstd	19	19.7	21	1.108	1.181	1.217
	PCmod	29	53.3	100	1.701	3.068	5.663
(100,150,5)	N	77	78.0	79	4.012	4.189	4.335
	PCstd	24	31.7	41	2.087	2.720	3.436
	PCmod	24	26.0	27	2.215	2.433	2.636
(500,250,25)	N	118	132.0	150	266.0	287.3	317.0
	PCstd	24	27.0	30	74.00	84.00	91.00
	PCmod	32	36.0	41	110.0	123.7	138.0

Tabela 5.6: Resultados da Família 2 de problemas com cones poliedrais em \mathbb{R}^n (matrizes indefinidas e simétricas).

(n, q, s)	Método	ITER			CPU		
		min.	med.	max.	min.	med.	max.
(10,5,1)	N	18	18.7	19	0.013	0.015	0.016
	PCstd	7	7.7	9	0.008	0.009	0.010
	PCmod	6	7.3	9	0.007	0.008	0.010
(10,15,1)	N	150	150.0	150	0.109	0.129	0.147
	PCstd	100	100.0	100	0.109	0.120	0.125
	PCmod	100	100.0	100	0.062	0.099	0.125
(100,50,5)	N	150	150.0	150	1.709	1.808	1.884
	PCstd	100	100.0	100	2.464	2.631	2.730
	PCmod	100	100.0	100	2.528	2.663	2.855
(100,100,5)	N	150	150.0	150	5.514	5.638	5.838
	PCstd	100	100.0	100	5.803	6.151	6.411
	PCmod	100	100.0	100	5.740	5.829	5.882
(100,150,5)	N	150	150.0	150	8.796	9.009	9.264
	PCstd	100	100.0	100	10.20	10.55	10.80
	PCmod	100	100.0	100	9.906	10.01	10.21
(500,250,25)	N	150	150.0	150	284.0	284.7	285.0
	PCstd	100	100.0	100	271.0	271.7	273.0
	PCmod	100	100.0	100	297.0	314.3	327.0

Tabela 5.7: Resultados da Família 3 de problemas com cones poliedrais em \mathbb{R}^n (matrizes indefinidas, não simétricas e singulares).

(n, q, s)	Método	ITER			CPU		
		min.	med.	max.	min.	med.	max.
(10, 5, 1)	N	21	21.3	22	0.016	0.022	0.031
	PCstd	13	13.3	14	0.012	0.014	0.016
	PCmod	13	13.0	13	0.012	0.014	0.017
(10, 15, 1)	N	12	12.7	13	0.015	0.021	0.031
	PCstd	6	6.0	6	0.010	0.012	0.015
	PCmod	6	6.0	6	0.010	0.012	0.017
(100, 50, 5)	N	30	32.7	36	0.374	0.421	0.453
	PCstd	14	16.3	20	0.390	0.437	0.514
	PCmod	15	15.0	15	0.327	0.363	0.419
(100, 100, 5)	N	44	47.3	53	2.309	3.000	3.744
	PCstd	19	19.0	19	1.061	1.071	1.077
	PCmod	21	22.3	24	1.090	1.259	1.482
(100, 150, 5)	N	46	48.3	50	3.167	3.659	3.978
	PCstd	19	21.0	23	1.823	2.126	2.325
	PCmod	21	23.0	25	1.904	2.177	2.463
(500, 250, 25)	N	121	136.0	144	242.0	296.3	324.0
	PCstd	22	23.3	26	68.00	76.30	85.00
	PCmod	23	25.3	30	77.00	82.30	92.00

Tabela 5.8: Resultados da Família 4 de problemas com cones poliedrais em \mathbb{R}^n (matrizes indefinidas, não simétricas e solução sem complementaridade estrita).

q	Solução gerada	r	Método	ITER			CPU		
				min.	med.	max.	min.	med.	max.
3	interior	0.1	N	26	26.3	27	0.009	0.012	0.018
			PC	10	12.7	17	0.005	0.011	0.019
		10	N	24	66.3	150	0.004	0.016	0.039
			PC	9	39.7	100	0.002	0.015	0.040
	borda	0.1	N	26	26.0	26	0.004	0.004	0.004
			PC	7	7.3	8	0.001	0.001	0.001
		10	N	29	29.7	31	0.004	0.005	0.005
			PC	14	14.7	15	0.003	0.014	0.016
12	interior	0.1	N	25	108.3	150	0.015	0.060	0.089
			PC	12	12.7	14	0.009	0.009	0.010
		10	N	15	15.0	15	0.004	0.006	0.009
			PC	8	8.0	8	0.003	0.003	0.003
	borda	0.1	N	15	17.7	20	0.004	0.006	0.008
			PC	10	47.3	100	0.004	0.028	0.069
		10	N	15	16.3	17	0.005	0.005	0.005
			PC	11	14.7	18	0.007	0.008	0.009
20	interior	0.1	N	11	14.0	16	0.008	0.011	0.015
			PC	7	8.3	10	0.008	0.009	0.010
		10	N	14	15.0	16	0.007	0.007	0.008
			PC	10	10.3	11	0.006	0.010	0.013
	borda	0.1	N	15	15.0	15	0.008	0.012	0.014
			PC	11	11.7	13	0.008	0.012	0.015
		10	N	14	14.7	15	0.007	0.009	0.014
			PC	9	9.7	10	0.006	0.008	0.013
40	interior	0.1	N	11	14.3	16	0.019	0.020	0.022
			PC	7	8.3	10	0.009	0.011	0.014
		10	N	11	11.3	12	0.014	0.017	0.022
			PC	7	7.0	7	0.010	0.012	0.014
	borda	0.1	N	9	9.7	11	0.011	0.015	0.018
			PC	7	8.7	10	0.008	0.018	0.023
		10	N	14	14.0	14	0.017	0.027	0.037
			PC	9	25.3	51	0.024	0.060	0.129

Tabela 5.9: Resultados da Família 1 de problemas com cones 3D (T é indefinida).

5.4. EXPERIMENTOS NUMÉRICOS

q	Solução gerada	r	Método	ITER			CPU		
				min.	med.	max.	min.	med.	max.
3	interior	0.1	N	25	25.7	26	0.008	0.037	0.093
			PC	7	7.7	8	0.003	0.013	0.032
		10	N	24	25.0	26	0.004	0.011	0.020
			PC	6	6.0	6	0.001	0.002	0.003
	borda	0.1	N	26	26.0	26	0.004	0.004	0.004
			PC	7	7.0	7	0.001	0.001	0.001
		10	N	25	26.0	27	0.004	0.007	0.009
			PC	7	8.0	9	0.001	0.002	0.003
12	interior	0.1	N	14	14.3	15	0.008	0.009	0.010
			PC	8	8.7	10	0.006	0.008	0.010
		10	N	14	14.3	15	0.005	0.006	0.007
			PC	7	7.0	7	0.003	0.004	0.005
	borda	0.1	N	13	13.7	14	0.004	0.012	0.026
			PC	7	8.3	10	0.003	0.004	0.006
		10	N	13	14.3	15	0.005	0.008	0.010
			PC	7	7.7	8	0.006	0.006	0.006
20	interior	0.1	N	13	13.3	14	0.006	0.006	0.007
			PC	8	8.0	8	0.004	0.005	0.005
		10	N	12	12.3	13	0.009	0.011	0.013
			PC	7	7.3	8	0.007	0.008	0.009
	borda	0.1	N	15	15.7	17	0.007	0.013	0.017
			PC	9	9.7	10	0.007	0.009	0.012
		10	N	14	14.0	14	0.006	0.007	0.009
			PC	8	8.7	9	0.005	0.007	0.010
40	interior	0.1	N	11	11.3	12	0.013	0.014	0.014
			PC	7	7.0	7	0.011	0.018	0.021
		10	N	11	11.7	12	0.015	0.022	0.031
			PC	7	7.3	8	0.011	0.015	0.022
	borda	0.1	N	15	16.7	18	0.028	0.038	0.044
			PC	10	10.3	11	0.018	0.026	0.032
		10	N	14	15.0	16	0.017	0.027	0.039
			PC	10	10.0	10	0.026	0.029	0.030

Tabela 5.10: Resultados da Família 2 de problemas com cones 3D(T é definida positiva).

q	Solução gerada	r	Método	ITER			CPU		
				min.	med.	max.	min.	med.	max.
3	interior	0.1	N	22	24.0	25	0.007	0.031	0.080
			PC	8	8.3	9	0.003	0.003	0.004
		10	N	20	20.7	21	0.012	0.013	0.015
			PC	7	41.3	100	0.009	0.014	0.024
	borda	0.1	N	24	24.7	25	0.004	0.007	0.008
			PC	8	8.0	8	0.002	0.003	0.004
		10	N	24	108.0	150	0.004	0.022	0.038
			PC	7	69.0	100	0.003	0.032	0.053
12	interior	0.1	N	150	150.0	150	0.046	0.062	0.083
			PC	100	100.0	100	0.037	0.049	0.072
		10	N	13	13.3	14	0.004	0.009	0.015
			PC	9	11.7	15	0.003	0.010	0.015
	borda	0.1	N	12	14.3	18	0.004	0.006	0.007
			PC	8	8.7	9	0.003	0.005	0.007
		10	N	14	14.7	15	0.006	0.009	0.012
			PC	8	10.3	12	0.003	0.005	0.009
20	interior	0.1	N	11	12.3	14	0.010	0.011	0.012
			PC	9	9.7	10	0.010	0.012	0.013
		10	N	13	104.3	150	0.006	0.047	0.076
			PC	8	69.3	100	0.005	0.055	0.096
	borda	0.1	N	150	150.0	150	0.080	0.085	0.088
			PC	100	100.0	100	0.053	0.055	0.057
		10	N	16	16.7	18	0.010	0.012	0.016
			PC	16	72.0	100	0.011	0.053	0.080
40	interior	0.1	N	11	14.3	16	0.026	0.032	0.040
			PC	8	12.0	17	0.024	0.036	0.050
		10	N	10	10.3	11	0.014	0.021	0.032
			PC	7	7.0	7	0.012	0.020	0.032
	borda	0.1	N	14	104.7	150	0.014	0.131	0.225
			PC	12	70.7	100	0.030	0.121	0.192
		10	N	14	15.7	17	0.022	0.023	0.024
			PC	11	12.0	14	0.019	0.020	0.022

Tabela 5.11: Resultados da Família 4 de problemas com cones 3D (T semidefinida negativa).

Capítulo 6

Considerações Finais

A principal contribuição deste trabalho foi analisar as propriedades de convergência local, implementar e ilustrar o desempenho de algoritmos de pontos interiores de segunda ordem adaptados para resolver o GNCP em cones poliedrais. Primeiramente, introduzimos um sistema não linear equivalente, com restrições de não negatividade em algumas variáveis, de modo a ser abordado por algoritmos de pontos interiores, tal que as estruturas intrínsecas das funções envolvidas F e G do problema (4.7) pudessem ser totalmente exploradas.

Dois algoritmos foram propostos: O **Algoritmo 1** é baseado em um método de Newton perturbado. Foi provado que tal método está bem definido, sob hipóteses usuais, e que ele tem convergência local quadrática. O **Algoritmo 2** é um método do tipo preditor-corretor baseado no método de Chebyshev, e também provamos que tal método está bem definido e é localmente convergente, com taxa quadrática. No caso em que as funções que definem o GNCP são afins, o método preditor-corretor do **Algoritmo 2** torna-se o bem conhecido método preditor-corretor de Mehrotra (como apresentado por [155]), desde que os parâmetros sejam adequadamente escolhidos. Observando isto tivemos a idéia de criar uma versão econômica do **Algoritmo 2** que chamamos **Algoritmo 2 econômico**, onde no cálculo da direção corretora não utiliza as derivadas parciais de segunda ordem.

Experimentos numéricos, baseados em problemas já abordados em [6, 7] ilustram o desempenho de ambos os algoritmos. Em geral, os algoritmos têm desempenhos bastante semelhantes, com clara vantagem para o **Algoritmo 2** em termos do número de iterações, mas uma leve desvantagem em termos de tempo de CPU por iteração.

O Algoritmo 2 e sua versão econômica, Algoritmo 2 econômico, tiveram um comportamento bem semelhante nos problemas resolvidos neste trabalho. Para os problemas de complementaridade implícita e para alguns cones poliedrais gerais em \mathbb{R}^n , o custo-benefício de computar os passos preditores-corretores vale a pena na redução do esforço total do algoritmo quando comparado ao método de Newton perturbado.

Apesar dos algoritmos propostos serem de natureza local, nos experimentos numéricos, variando as opções de pontos iniciais, uma solução do correspondente GNCP foi obtida na maioria dos testes. De fato, os casos de parada com Flag 3 (atingiu o número máximo de iterações) foram em sua maioria associados ao não cumprimento de alguma hipótese teórica.

Uma profunda investigação das escolhas dos parâmetros γ_k e μ_k e sua consistência com as hipóteses teóricas que garantam a convergência local quadrática (Teoremas 5.10 e 5.14) são objetos de pesquisa futura. Um estudo numérico comparando várias escolhas dos parâmetros em estratégias de pontos interiores aplicadas ao problema de programação não linear é feito em [14], mas até onde sabemos não existem referências relacionadas a escolhas práticas na implementação de algoritmos de pontos interiores aplicados a problemas de complementaridade e suas generalizações. Outra linha de investigação futura é a adaptação dos métodos aqui apresentados para resolver extensões de problemas de complementaridade. Também sugerimos como proposta de pesquisa futura a possibilidade de enfraquecimento das hipóteses, em especial a exigência da complementaridade estrita na solução. Nesse sentido, as técnicas vistas nos Capítulos 2 e 3 deste trabalho podem ser de grande valia, aliás esse foi o principal motivo que nos levou à investigação daqueles resultados.

Bibliografia

- [1] G. Alefeld. On the Convergence of Halley's Method. *Amer. Math. Monthly*, 88, p. 530 - 536, 1981.
- [2] R. Andreani; A. Friedlander. Bound constrained smooth optimization for solving variational inequalities and related problems. Operations research and systems (CLAIO 2000), Part I (Mexico City). *Ann. Oper. Res.*, 116, p. 179-198, 2002.
- [3] R. Andreani; A. Friedlander; J. M. Martínez. Solution of finite-dimensional variational inequalities using smooth optimization with simple bounds. *J. of Opt. Theory and Appl.*, 94, p. 635-657, 1997.
- [4] R. Andreani; A. Friedlander; M. P. Mello; S. A. Santos. Mixed nonlinear complementarity problems via nonlinear optimization: numerical results on multi-rigid-body contact problems with Coulomb friction. *Intern. J. for Comp. Methods in Eng. Science and Mechanics*, 6, p. 85-94, 2005.
- [5] R. Andreani; A. Friedlander; M. P. Mello; S. A. Santos. Box-constrained minimization reformulations of complementarity problems in second-order cones. *J. Global Optim.*, 40, p. 505-527, 2008.
- [6] R. Andreani; A. Friedlander; S. A. Santos. On the resolution of the generalized nonlinear complementarity problem. *SIAM J. Optim.*, 12, p. 303-321, 2001.
- [7] R. Andreani; A. Friedlander; S. A. Santos. Solving generalized nonlinear complementarity problems: Numerical experiments on polyhedral cones. Technical Report, IMECC, State University of Campinas, Campinas, Brazil, 2001. Disponível em http://www1.ime.unicamp.br/rel_pesq/2001/rp17-01.html.

- [8] R. Andreani; J. M. Martínez. On the solution of the extended linear complementarity problem. *Linear Algebra Appl.*, 281, p. 247-257, 1998.
- [9] R. Andreani; J. M. Martínez. Solving complementarity problems by means of a new smooth constrained nonlinear solver. In: *Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing methods* Edited by M. Fukushima and L. Qi. Kluwer, p. 1-24, 1999.
- [10] R. Andreani; J. M. Martínez. On the reformulation of nonlinear complementarity problems using the Fischer-Burmeister function. *Appl. Math. Letters*, 12, p. 7-12, 1999.
- [11] R. Andreani; J. M. Martínez. Reformulation of variational inequalities on a simplex and compactification of complementarity problems. *SIAM J. Optim.*, 10, p. 878-895, 2000.
- [12] R. Andreani; J. M. Martínez. On the solution of bounded and unbounded mixed complementarity problems. *Optimization*, p. 265-278, 2001.
- [13] R. Andreani; S. A. Santos; W. V. I. Shirabayashi. Second-order interior-point methods for solving generalized complementarity problems in polyhedral cones. Submetido para publicação.
- [14] M. Argáez; R. Tapia; L. Velázquez. Numerical comparisons of path-following strategies for a primal-dual interior-point method for nonlinear programming. *J. Optim. Theory Appl.*, 114, no. 2, 255-272, 2002.
- [15] I. Argyros; D. Chen. Results on the Chebyshev method in Banach spaces. *Proyecciones*, 12, p. 119-128, 1993.
- [16] H. Bateman. Halley's Methods for Solving Equations. *Amer. Math. Monthly*, 45, p. 11-17, 1938.
- [17] B. Bader; R. B. Schnabel. Curvilinear linesearch for tensor methods. *SIAM J. Sci. Comput.*, 25, p. 604-622, 2003.
- [18] J. F. Bonnans; C. C. Gonzaga. Convergence of interior point algorithms for the monotone linear complementarity problem. *Math. Oper. Res.*, 21, p. 1-25, 1996.

-
- [19] A. Bouaricha. Tensor methods for large, sparse unconstrained optimization. *SIAM J. Optim.*, 7, p. 732-756, 1997.
- [20] A. Bouaricha; R. B. Schnabel. Tensor methods for large, sparse nonlinear least squares problems. *SIAM J. Sci. Comput.*, 21, p. 1199-1221, 1999.
- [21] G. H. Brown Jr.. On Halley's variation of Newton's method. *Amer. Math. Monthly*, 84, p. 726-728, 1977.
- [22] F. E. Buffo; M. C. Maciel. A mixed nonlinear complementarity technique for solving the dynamics of a dexterous manipulation system. *Comput. Appl. Math.*, 25, p. 79-95, 2006.
- [23] V. Candela; A. Marquina. Recurrence relations for rational cubic methods I: the Halley method. *Computing*, 44, p. 169-184, 1990.
- [24] V. Candela; A. Marquina. Recurrence relations for rational cubic methods II: the Chebyshev method. *Computing*, 45, p. 355-367, 1990.
- [25] T. J. Carpenter; I. J. Lustig; J. M. Mulvey; D. F. Shanno. Higher-order predictor-corrector interior point methods with application to quadratic objectives. *SIAM J. Optim.*, 3, p. 696-725, 1993.
- [26] R. C. Cavanagh, *Difference equations and iterative process*, Thesis, Computer Sci. Dept., Univ. of Maryland, College Park, 1970.
- [27] D. Chen; I.K. Argyros; Q. Qian. A local convergence theorem for the super-Halley method in a Banach space. *Appl. Math. Lett.*, 7, p. 49-52, 1994.
- [28] J.-S. Chen; P. Tseng. An unconstrained smooth minimization reformulation of the second-order cone complementarity problem. *Math. Program.*, 104, p. 293-327, 2005.
- [29] C. Chun. Some second-derivative-free variants of Chebyshev-Halley methods. *Appl. Math. and Comp.*, 191, p. 410-414, 2007.
- [30] F. H. Clarke. *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
- [31] R. W. Cottle. Nonlinear programs with positively bounded Jacobians. *SIAM J. on Appl. Math.*, 14, p. 147-158, 1966.

- [32] R. W. Cottle; J.-S. Pang; R. E. Stone. *The linear complementarity problem*, Academic Press, Boston, MA, 1992.
- [33] CPLEX: ILOG Optimization. <http://www.ilog.com>.
- [34] A. N. Daryina; A. F. Izmailov; M. V. Solodov. A class of active-set Newton methods for mixed complementarity problems. *SIAM J. Optim.*, 15, p. 409-429, 2004/05.
- [35] T. De Luca; F. Facchinei; C. Kanzow. A semismooth equation approach to the solution of nonlinear complementarity problems. *Math. Program.*, 75, p. 407-439, 1996.
- [36] D. W. Decker; H. B. Keller; C. T. Kelley. Convergence rates for Newton's method at singular points. *SIAM J. of Numer. Anal.*, 20, p. 296-314, 1983.
- [37] D. W. Decker; C. T. Kelley. Newton's method at singular points, I. *SIAM J. of Numer. Anal.*, 17, p. 66-70, 1980.
- [38] D. W. Decker; C. T. Kelley. Newton's method at singular points, II. *SIAM J. of Numer. Anal.*, 17, p. 465-471, 1980.
- [39] D. W. Decker; C. T. Kelley. Convergence acceleration for Newton's method at singular points. *SIAM J. of Numer. Anal.*, 19, p. 219-229, 1981.
- [40] D. W. Decker; C. T. Kelley. Broyden's method for a class of problems having singular jacobian at the root. *SIAM J. of Numer. Anal.*, 22, p. 566-574, 1985.
- [41] N. Deng; H. Zhang. Theoretical efficiency of a new inexact method of tangent hiperbolas. *Optim. Methods Softw.*, 19, p. 247-265, 2004.
- [42] J.E. Dennis; R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*, Englewood Cliffs, Prentice Hall, 1983.
- [43] W. S. Dorn. Self-dual quadratic programs. *J. of the Soc. Ind. Appl. Math.*, 9, p. 51-54, 1961.
- [44] A. S. El-Bakry; R. A. Tapia; T. Tsuchiya; Y. Zhang. On the formulation and theory of the Newton interior-point method for nonlinear programming. *J. of Optim. Theory and Appl.*, 89, p. 507-541, 1996.

-
- [45] J. A. Ezquerro; M. A. Hernández. On the R -order of the Halley method., *J. Math. Anal. Appl.*, 303, p. 591-601, 2005.
- [46] J. A. Ezquerro; M. A. Hernández. Halley's method for operators with unbounded second derivative. *Appl. Numer. Math.*, 57, p. 354-360, 2007.
- [47] F. Facchinei; C. Kanzow. A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems. *Math. Program.*, 76, p. 493-512, 1997.
- [48] F. Facchinei; J. Soares. A new merit function for nonlinear complementarity problems and a related algorithm. *SIAM J. Optim.*, 7, p. 225-247, 1997.
- [49] D. Feng; P. D. Frank; R. B. Schnabel. Local convergence analysis of tensor methods for nonlinear equations. *Math. Program.*, 62, p. 427-459, 1993.
- [50] D. Feng; R. B. Schnabel. Tensor methods for equality constrained optimization. *SIAM J. Optim.*, 6, p. 653-673, 1996.
- [51] M. C. Ferris; J.-S. Pang. Engineering and economic applications for nonlinearly complementarity problems. *SIAM Review*, 39, p. 669-713, 1997.
- [52] M. C. Ferris and J.-S. Pang. *Complementarity and Variational Problems: State of the Art*, Philadelphia, PA: SIAM, 1997.
- [53] J.P. Fink; W. Rheinboldt. A geometric framework for the numerical study of singular points. *SIAM J. of Numer. Anal.*, 24, p. 618-633, 1987.
- [54] A. Fischer. Solution of monotone complementarity problems with locally Lipschitzian functions. *Math. Program.*, 76, p. 513-532, 1997.
- [55] A. Friedlander; J. M. Martínez; S. A. Santos. Solution of linear complementarity problems using minimization with simple bounds. *J. Global Optim.*, 6, p. 253-267, 1995.
- [56] A. Galántai. The theory of Newton's method. *J. of Comp. and Appl. Math.*, 124, p. 25-44, 2000.
- [57] D. Gay. Modifying singular values: existence of solutions to systems of nonlinear equations having a possibly singular jacobian matrix. *Math. of Comp.*, 31, p. 962-973, 1977.

- [58] R.-D. Ge; Z.-Q. Xia. An ABS algorithm for solving singular nonlinear systems with rank defects. *J. Appl. Math. & Comp.*, 12, p. 1-20, 2003.
- [59] R.-D. Ge; Z.-Q. Xia; J.-G Liu. A modified Brown algorithm for solving singular nonlinear systems with rank defects. *J. of Comp. and Appl. Math.*, 181, p. 252-265, 2005.
- [60] G. H. Golub; C. F. Van Loan. *Matrix Computations, Third Edition*, Baltimore, Maryland: The Johns Hopkins University Press, 1996.
- [61] C. C. Gonzaga. Path-following methods for linear programming. *SIAM Review*, 34, p. 167-224, 1992.
- [62] C. C. Gonzaga. Complexity of predictor-corrector algorithms for LCP based on a large neighborhood of the central path. *SIAM J. Optim.*, 10, p. 183-194, 1999.
- [63] A. Griewank. Starlike domains of convergence for Newton's method at singularities. *Numer. Math.*, 35, p. 95-111, 1980.
- [64] A. Griewank. On solving nonlinear equations with simple singularities or nearly singular solutions. *SIAM Review*, 27, p. 537-563, 1985.
- [65] A. Griewank; M. R. Osborne. Newton's method for singular problems when the dimension of the null space is > 1 . *SIAM J. of Numer. Anal.*, 18, p. 145-149, 1981.
- [66] A. Griewank; M. R. Osborne. Analysis and modification of Newton's method at singularities. *SIAM J. of Numer. Anal.*, 18, p. 145-149, 1981.
- [67] O. Güler. Existence of interior points and interior paths in nonlinear monotone complementarity problems. *Math. Oper. Res.*, 18, p. 128-147, 1993.
- [68] G. Gundersen; T. Steihaug. On large scale unconstrained optimization problems and higher order methods. *Optimization Online 2007*. A aparecer em *Optimization Methods and Software* (2009).
- [69] J. M. Gutiérrez; M. A. Hernández. A family of Chebyshev-Halley type methods in Banach spaces. *Bull. Austral. Math. Soc.*, 55, p. 113-130, 1997.

-
- [70] J. M. Gutiérrez; M. A. Hernández. Recurrence relations for the super-Halley method. *Comput. Math. Appl.*, 36, p. 1-8, 1998.
- [71] J. M. Gutiérrez; M. A. Hernández. An acceleration of Newton's method: Super-Halley method. *Appl. Math. and Comp.*, 117, p. 223-239, 2001.
- [72] G. J. Hatebler; A. L. Price. Existence theory for generalized nonlinear complementarity problems. *J. of Opt. Theory and Appl.*, 7, p. 223-239, 1971.
- [73] P. T. Harker; J. S. Pang. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Math. Program.*, 48, p. 161-220, 1990.
- [74] P. T. Harker; B. Xiao. Newton's method for the nonlinear complementarity problem: A B-differentiable equation approach. *Math. Program.*, 48, p. 339-357, 1990.
- [75] M. A. Hernández; M. A. Salanova. A family of Chebyshev-Halley type methods. *Intern. J. of Computer Math.*, 47, p. 59-63, 1993.
- [76] M. A. Hernández; M. A. Salanova. Chebyshev method and convexity. *Appl. Math. and Comp.*, 95, p. 51-62, 1998.
- [77] M. A. Hernández; M. A. Salanova. Indices of convexity and concavity. Application to Halley method. *Appl. Math. and Comp.*, 103, p. 27-49, 1999.
- [78] M. A. Hernández; M. A. Salanova. Modification of the Kantorovich assumptions for semilocal convergence of the Chebyshev method. *Appl. Math. and Comp.*, 126, p. 131-143, 2000.
- [79] A. W. Ingleton. A problem in linear inequalities. *Proc. London Math. Soc.*, 16, p. 519-536, 1966.
- [80] A. N. Iusem. An interior point method for the nonlinear complementarity problem. *Appl. Numer. Math.*, 24, p. 469-482, 1997.
- [81] A. F. Izmailov; M. V. Solodov. Error bounds for 2-regular mappings with Lipschitzian derivatives and their applications. *Math. Program.*, 89, p. 413-435, 2001.

- [82] A. F. Izmailov; M. V. Solodov. Superlinearly convergent algorithms for solving singular equations and smooth reformulations of complementarity problems. *SIAM J. Optim.*, 13, p. 386-405, 2002.
- [83] A. F. Izmailov; M. V. Solodov. The theory of 2-regularity for mappings with Lipschitzian derivatives and its applications to optimality conditions. *Math. Oper. Res.*, 27, p. 614-635, 2002.
- [84] R. H. F. Jackson; G. P. McCormick. The polyadic structure of factorable function tensors with applications to high-order minimization techniques. *J. of Opt. Theory and Appl.*, 51, p. 63-94, 1986.
- [85] R. H. F. Jackson; G. P. McCormick. Second-order sensitivity analysis in factorable programming: Theory and applications. *Math. Program.*, 41, p. 1-27, 1988.
- [86] J. Ji; F. A. Potra; R. Sheng. On the local convergence of a predictor-corrector method for semidefinite programming. *SIAM J. Optim.*, 10, p. 195-210, 1999.
- [87] J. Ji; F. A. Potra; R. Sheng. On a general class of interior-point algorithms for semidefinite programming with polynomial complexity and superlinear convergence. *Methods Appl. Anal.*, 6, p. 549-569, 1999.
- [88] H. Jiang; M. Fukushima; L. Qi; D. Sun. A trust region method for solving generalized complementarity problems. *SIAM J. Optim.*, 8, p. 140-157, 1998.
- [89] R. Kalaba; A. Tishler. A generalized Newton algorithm using higher-order derivatives. *J. of Opt. Theory and Appl.*, 39, p. 1-17, 1983.
- [90] C. Kanzow. Nonlinear complementarity as unconstrained optimization. *J. of Optim. Theory and Appl.*, 88, p. 139-155, 1996.
- [91] C. Kanzow; M. Fukushima. Equivalence of the generalized complementarity problem to differentiable unconstrained minimization. *J. of Opt. Theory and Appl.*, 90, p. 581-603, 1996.
- [92] S. Karamardian. The nonlinear complementarity problem with applications, part 1. *J. of Opt. Theory and Appl.*, 4, p. 87-98, 1969.

-
- [93] S. Karamardian. Generalized complementarity problem. *J. of Opt. Theory and Appl.*, 8, p. 161-168, 1971.
- [94] S. Karamardian. The complementarity problem. *Math. Program.*, 2, p. 107-129, 1972.
- [95] N. K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4, p. 373-395, 1984.
- [96] C. T. Kelley; R. Suresh. A new acceleration method for Newton's method at singular points. *SIAM J. of Numer. Anal.*, 20, p. 1001-1009, 1983.
- [97] M. Kocvara and J. V. Outrata. On the solution of optimum design problems with variational inequalities. In: it Recent Advances in Nonsmooth Optimization (D.-Z. Du, L. Qi and R.L. Womersley, eds.), p. 171-191. World Scientific Publishers, 1995.
- [98] M. Kojima. A unification of the existence theorems of the nonlinear complementarity problem. *Math. Program.*, 9, p. 257-277, 1975.
- [99] M. Kojima; Y. Kurita; S. Mizuno. Large-step interior point algorithms for linear complementarity problems. *SIAM J. Optim.*, 3, p. 398-412, 1993.
- [100] M. Kojima; T. Noma; A. Yoshise. Global convergence in infeasible-interior-point algorithms. *Math. Program.*, 65, p. 43-72, 1994.
- [101] M. Kojima; N. Meggido; S. Mizuno. A general framework of continuation methods for complementarity problems. *Math. Op. Res.*, 18, p. 945-963, 1993.
- [102] M. Kojima; N. Meggido; S. Mizuno. A primal-dual infeasible-interior point algorithm for linear programming. *Math. Program.*, 61, p. 263-280, 1993.
- [103] M. Kojima; N. Meggido; T. Noma. Homotopy continuation method for nonlinear complementarity problems. *Math. Op. Res.*, 16, p. 754-774, 1991.
- [104] M. Kojima; S. Mizuno; T. Noma. A new continuation method for complementarity problems with uniform P-functions. *Math. Program.*, 43, p. 107-113, 1989.
- [105] M. Kojima; S. Mizuno; A. Yoshise. A polynomial-time algorithm for a class of linear complementarity problems. *Math. Program.*, 44, p. 1-26, 1989.

- [106] T. G. Kolda. Multilinear operators for higher-order decompositions. Technical Report SAND2006-2081, Sandia National Laboratories, Livermore, California, 2006.
- [107] J. Kou; Y. Lin; X. Wang. Modified Halley's method free from second derivative. *Appl. Math. and Comp.*, 183, p. 704-708, 2006.
- [108] J. Kou; Y. Lin; X. Wang. A variant of super-Halley method with accelerated fourth-order convergence. *Appl. Math. and Comp.*, 186, p. 535-539, 2007.
- [109] J. Kou; Y. Lin; X. Wang. A family of fifth-order iterations composed of Newton and third-order methods. *Appl. Math. and Comp.*, 186, p. 1258-1262, 2007.
- [110] J. Kou; X. Wang; Six-order variants of Chebyshev-Halley methods for solving non-linear equations. *Appl. Math. and Comp.*, 190, p. 1839-1843, 2007.
- [111] V. L. R. Lopes; J. M. Martínez. On the convergence of quasi-Newton methods for nonsmooth problems. *Num. Func. Anal. and Optim.*, 16, p. 1193-1209, 1995.
- [112] Z. Luo; S. Wu; Y. Ye. Predictor-corrector method for nonlinear complementarity problem. *Acta Math. Appl. Sinica*, 13, p. 321-328, 1997.
- [113] I. J. Lustig; R. E. Marsten; D. F. Shanno. On implementing Mehrotra's predictor-corrector interior-point method for linear programming. *SIAM J. Optim.*, 2, p. 435-449, 1992.
- [114] J. M. Martínez. Quasi-Newton methods with derivatives. *Calcolo*, 32, p. 105-135, 1995.
- [115] J. M. Martínez; L. Qi. Inexact Newton methods for solving nonsmooth equations. *J. of Comp. and Appl. Math.*, 60, p. 127-145, 1995.
- [116] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM J. Optim.*, 2, p. 575-601, 1992.
- [117] S. Mehrotra. Asymptotic convergence in a generalized predictor-corrector method. *Math. Program.*, 74, p. 11-28, 1996.
- [118] S. Mehrotra; R. A. Stubbs. Predictor-corrector methods for a class of linear complementarity problems. *SIAM J. Optim.*, 4, p. 441-453, 1994.

-
- [119] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*, Philadelphia, PA: SIAM, 2000.
- [120] S. Mizuno; M. J. Todd; Y. Ye. On adaptive-step primal-dual interior-point algorithms for linear programming. *Math. Oper. Res.*, 18, p. 964-981, 1993.
- [121] R. D. C. Monteiro; J. S. Pang. Properties of an interior-point mapping for mixed complementarity problems. *Math. Oper. Res.*, 21, p. 629-654, 1996.
- [122] R. D. C. Monteiro; J. S. Pang. On two interior-point mappings for nonlinear semidefinite complementarity problems. *Math. Oper. Res.*, 23, p. 39-60, 1998.
- [123] R. D. C. Monteiro; J. S. Pang; T. Wang. A positive algorithm for the nonlinear complementarity problem. *SIAM J. Optim.*, 5, p. 129-148, 1995.
- [124] M. A. Noor. General variational inequalities. *Applied Mathematics Letters*, 1, p. 119-122, 1988.
- [125] M. A. Noor. Variational inequalities in physical oceanography, in: M. Rahman (Ed.) *Ocean Waves Engineering, Advances in Fluid Mechanics*, vol. 2, WIT Press, p. 201-266, 1994.
- [126] C. Oberlin; S. J. Wright. An accelerated Newton method for equations with semismooth Jacobians and nonlinear complementarity problems. *Math. Program.*, 117, p. 355-386, 2009.
- [127] J. M. Ortega; W.C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York , 1970.
- [128] J. V. Outrata; J. Zowe. A Newton method for a class of quasi-variational inequalities. *Comput. Optim. Appl.*, 4, p. 5-21, 1995.
- [129] J.-S. Pang. The implicit complementarity problem. *Nonlinear Programming 4*, Editors: O. Mangasarian, M. Robinson & R. Meyer, Academic Press, New York, p. 487-518, 1981.
- [130] J.-S. Pang. Complementarity Problem, *Handbook of Global Optimization*, R. Horst and P. Pardalos (Eds.), Kluwer Academic Publishers, Boston, p. 271-338, 1995.

- [131] J.-S. Pang; S. A. Gabriel. NE/SQP: A robust algorithm for the nonlinear complementarity problem. *Math. Program.*, 60, p. 295-337, 1993.
- [132] J. Peng; Y. Yuan. Unconstrained methods for generalized complementarity problems, *J. of Comp. Mathematics*, 15, p. 253-264, 1997.
- [133] A. Pinto da Costa; A. Seeger. Numerical resolution of cone-constrained eigenvalue problems. *Comput. Appl. Math.*, 28, p. 37-61, 2009.
- [134] F. A. Potra. A quadratically convergent predictor-corrector method for solving linear programs from infeasible starting points. *Math. Program.*, 67, p. 383-406, 1994.
- [135] F. A. Potra. The Mizuno-Todd-Ye algorithm in a larger neighborhood of the central path. *European J. Oper. Res.*, 143, p. 257-267, 2002.
- [136] F. A. Potra; R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. *SIAM J. Optim.*, 8, p. 1007-1028, 1998.
- [137] F. A. Potra; Y. Ye. Interior-point methods for nonlinear complementarity problems. *J. Optim. Theory Appl.*, 88, p. 617-642, 1996.
- [138] L. Qi; J. Sun. A nonsmooth version of Newton's method. *Math. Program.*, 58, p. 353-367, 1993.
- [139] A. Rafiq; M. Awais; F. Zafar. Modified efficient variant of super-Halley method. *Appl. Math. and Comp.*, 189, p. 2004-2010, 2007.
- [140] L. B. Rall. Convergence of the Newton process to multiple solutions. *Numer. Math.*, 9, p. 23-37, 1966.
- [141] G. W. Reddien. On Newton's method for singular problems. *SIAM J. of Numer. Anal.*, 15, p. 993-996, 1978.
- [142] G. W. Reddien. Newton's method and high order singularities. *Comput. Math. with Appl.*, 5, p. 79-86, 1979.
- [143] C. Roos; T. Terlaky; J.-Ph. Vial. *Interior Point Methods for Linear Optimization*, Second edition, New York: Springer, 2006.

-
- [144] M. Salahi; J. Peng; T. Terlaky. On Mehrotra-type predictor-corrector algorithms. *SIAM J. Optim.*, 18, p. 1377-1397, 2007.
- [145] P. S. M. Santos; S. Scheimberg. A projection algorithm for general variational inequalities with perturbed constraint sets. *Appl. Math. and Comp.*, 181, p. 649-661, 2006.
- [146] M. Salahi; T. Terlaky. Mehrotra-type predictor-corrector algorithm revisited. *Optim. Methods Soft.*, 23, p. 259-273, 2008.
- [147] L. A. Santaló. *Vectores y Tensores con sus aplicaciones*, Buenos Aires, Ed. Univ. de Buenos Aires, 1961.
- [148] T. R. Scavo; J. B. Thoo. On the Geometry of Halley's Method. *Amer. Math. Monthly*, 102, p. 417-426, 1995.
- [149] R. B. Schnabel. Conic methods for unconstrained minimization and Tensor methods for nonlinear equations. *Math. Programming: the state of the art*, Springer - Berlin, p. 413-438, 1983.
- [150] R. B. Schnabel; T. Chow. Tensor methods for unconstrained optimization using second derivatives. *SIAM J. Optim.*, 1, p. 293-315, 1991.
- [151] R. B. Schnabel; P. D. Frank. Tensor methods for nonlinear equations. *SIAM J. of Numer. Anal.*, 21, p. 815-843, 1984.
- [152] Y.-Q. Shen; T. J. Ypma. Newton's method for singular nonlinear equations using approximate left and right null spaces of the jacobian. *Appl. Numer. Math.*, 54, p. 256-265, 2005.
- [153] E. M. Simantiraki; D. F. Shanno. An infeasible-interior-point method for linear complementarity problems. *SIAM J. Optim.*, 7, p. 620-640, 1997.
- [154] D. Sun; R.S. Womersley; H. Qi. A feasible semismooth asymptotically Newton method for mixed complementarity problems. *Math. Program.*, 94, p. 167-187, 2002.
- [155] R. Tapia; Y. Zhang; M. Saltzman; A. Weiser. The Mehrotra predictor-corrector interior-point method as a perturbed composite Newton method. *SIAM J. Optim.*, 6, p. 47-56, 1996.

- [156] T. Terlaky (Ed.) *Interior Point Methods of Mathematical Programming*, Applied Optimization Series, Vol. 5, Dordrecht: Kluwer Academic Publishers, 1996.
- [157] P. Tseng; N. Yamashita; M. Fukushima. Equivalence of complementarity problems to differentiable minimization: a unified approach. *SIAM J. Optim.*, 6, p. 446-460, 1996.
- [158] R. J. Vanderbei. LOQO: an interior point code for quadratic programming. Interior point methods. *Optim. Methods Softw.*, 11/12, p. 451-484, 1999.
- [159] Y. J. Wang; P. S. Fei. A infeasible Mizuno-Todd-Ye type algorithm for convex quadratic programming with polynomial complexity. *Num. Func. Anal. and Optim.*, 28, p. 487-502, 2007.
- [160] Y. J. Wang; F. M. Ma; J. Z. Zhang. A nonsmooth L-M method for solving the generalized nonlinear complementarity problem over a polyhedral cone. *Appl. Math. Optim.*, 52, p. 73-92, 2005.
- [161] Y. J. Wang; X. Z. Zhang; L. Q. Qi. Unconstrained optimization reformulation of the generalized nonlinear complementarity problem and related method. *Optimization*, 54, p. 563-577, 2005.
- [162] S. J. Wright. *Primal-dual interior-point methods*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
- [163] S. J. Wright; D. Ralph. A superlinear infeasible-interior-point algorithm for monotone complementarity problems. *Math. Oper. Res.*, 21, p. 815-838, 1996.
- [164] S. J. Wright; Y. Zhang. A superquadratic infeasible-interior-point method for linear complementarity problems. *Math. Program.*, 73, p. 269-289, 1996.
- [165] Y. S. Xia; J. Wang. A general projection neural network for solving monotone variational inequalities and related optimization problems. *IEEE Transactions on Neural Networks*, 15, p. 318-328, 2004.
- [166] B. Xiao; P. T. Harker. A nonsmooth Newton method for variational inequalities, II: numerical results. *Math. Program.*, 65, p. 195-216, 1994.

- [167] XpressMP: Dash Optimization. <http://www.dashoptimization.com>.
- [168] T. Yamamoto. Historical developments in convergence analysis for Newton's and Newton-like methods. *J. of Comp. and Appl. Math.*, 124, p. 1-23, 2000.
- [169] N. Yamashita; M. Fukushima. Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems. *Math. Program.*, 76, p. 469-491, 1997.
- [170] H. Yamashita; H. Yabe. Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization. *Math. Program.*, 75, p. 377-397, 1996.
- [171] Y. Ye; K. Anstreicher. On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP. *Math. Program.*, v. 62, p. 537-551, 1993.
- [172] Y. Ye; O. Güler; R. A. Tapia; Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$ -iteration algorithm for linear programming. *Math. Program.*, v. 59, p. 151-162, 1993.
- [173] Y. Zhang. On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem. *SIAM J. Optim.*, v. 4, p. 208-227, 1994.
- [174] X. Z. Zhang; H. F. Jiang; Y. J. Wang. A smoothing Newton-type method for generalized nonlinear complementarity problem. *J. of Comp. and Appl. Math.*, v. 212, p. 75-85, 2008.
- [175] X. Z. Zhang; F. M. Ma; Y. J. Wang. A Newton-type algorithm for generalized linear complementarity problem over a polyhedral cone. *Applied Math. and Computation*, 169, p. 388-401, 2005.
- [176] Y. Zhang; R. A. Tapia. Superlinear and quadratic convergence of primal-dual interior-point methods for linear programming revisited. *J. Optim. Theory and Appl.*, v. 73, p. 229-242, 1992.
- [177] Y. Zhang; R. A. Tapia. A superlinearly convergent polynomial primal-dual interior point algorithm for linear programming. *SIAM J. Optim.*, v. 3, p. 118-133, 1993.

- [178] Y. Zhang; R. A. Tapia; J. Dennis Jr. On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms. *SIAM J. Optim.*, v. 2, p. 304-324, 1992.
- [179] Y. Zhang; R. A. Tapia; F. Potra. On the superlinear convergence of interior-point algorithms for a general class of problems. *SIAM J. Optim.*, v. 3, p. 413-422, 1993.
- [180] Y. Zhang; D. Zhang. On polynomiality of the Mehrotra-type predictor-corrector interior-point algorithms. *Math. Program.*, v. 68, p. 303-318, 1995.
- [181] D. Zhang, Y. Zhang. A Mehrotra-type predictor-corrector algorithm with polynomiality and Q -subquadratic convergence. *Annals of Oper. Res.*, v. 62, p. 131-150, 1996.

Apêndice A

Operações Tensoriais

As operações tensoriais consideradas neste trabalho tiveram como ponto de partida o livro de Santaló [147]. Também foi de grande valia o trabalho de Kolda [106].

Se $\mathcal{T} \in \mathbb{R}^{n \times n \times n}$ é um tensor de terceira ordem, então os elementos de \mathcal{T} são representados por \mathcal{T}_{kij} , $k, i, j = 1, 2, \dots, n$, e as operações envolvendo este tensor e vetores $u, v \in \mathbb{R}^n$ são dadas por:

$\mathcal{T}u \in \mathbb{R}^{n \times n}$ é uma matriz cujos elementos são:

$$(\mathcal{T}u)_{kj} = \sum_{i=1}^n \mathcal{T}_{kij} u_i, \quad k, j = 1, 2, \dots, n.$$

$\mathcal{T}uv \in \mathbb{R}^n$ é um vetor cujos elementos são:

$$(\mathcal{T}uv)_k = \sum_{i=1}^n \sum_{j=1}^n \mathcal{T}_{kij} u_i v_j, \quad k = 1, 2, \dots, n.$$

No caso em que $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, então $F(x) \equiv (F_1(x), F_2(x), \dots, F_n(x))^T$ onde cada $F_i : \mathbb{R}^n \rightarrow \mathbb{R}$, assim $F''(x) \equiv \mathcal{T} \in \mathbb{R}^{n \times n \times n}$. Neste caso podemos escrever:

$$F''(x) u v = \begin{bmatrix} u^T \nabla^2 F_1(x) v \\ u^T \nabla^2 F_2(x) v \\ \vdots \\ u^T \nabla^2 F_n(x) v \end{bmatrix} \quad \text{e} \quad F''(x) u = \begin{bmatrix} u^T \nabla^2 F_1(x) \\ u^T \nabla^2 F_2(x) \\ \vdots \\ u^T \nabla^2 F_n(x) \end{bmatrix}.$$

As mesmas expressões valem para o caso em que $\mathcal{T} \equiv H''(w) \in \mathbb{R}^{n_1 \times n_1 \times n_1}$ e $u, v \in \mathbb{R}^{n_1}$, no Capítulo 5:

$$H''(w) u v = \begin{bmatrix} u^T \nabla^2 H_1(w) v \\ u^T \nabla^2 H_2(w) v \\ \vdots \\ u^T \nabla^2 H_{n_1}(w) v \end{bmatrix} \in \mathbb{R}^{n_1} \quad \text{e} \quad H''(w) u = \begin{bmatrix} u^T \nabla^2 H_1(w) \\ u^T \nabla^2 H_2(w) \\ \vdots \\ u^T \nabla^2 H_{n_1}(w) \end{bmatrix} \in \mathbb{R}^{n_1 \times n_1}.$$

Apêndice B

Tabelas

Neste apêndice apresentamos as tabelas completas dos resultados com os diferentes pontos iniciais utilizados, para todo o conjunto de problemas-teste GNCP, e que geraram as estatísticas mínima, média e máxima das tabelas incluídas na seção 5.4. Os casos em que houve parada por falta de progresso ($\|w_k - w_{k-1}\|_1 < \epsilon$) estão identificados com *, e na tabela imediatamente seguinte apresentamos os valores correspondentes de $\|H_k\|_1$ para fins de acompanhamento dos resultados.

Problemas	n	Método	a		b		c	
			ITER	CPU	ITER	CPU	ITER	CPU
Família 1 $M = P = M^T$ $M = M^T$ $M > 0$	5	N	19	0.017	19	0.003	19	0.004
		PC	7	0.016	7	0.003	7	0.002
	500	N	14	2.544	14	2.767	14	2.768
		PC	8	2.757	8	3.013	8	2.962
Família 2 $M = P$ $M = M^T$ M indefinida	5	N	19	0.002	19	0.002	19	0.003
		PC	7	0.002	7	0.001	7	0.001
	500	N	14	2.621	14	2.763	14	2.673
		PC	8	2.967	8	3.005	8	2.988
Família 3 $M = P$ $M = M^T$ $M \geq 0$	5	N	150	1.548	150	1.446	150	1.639
		PC	100	1.620	100	2.162	100	2.073
	500	N	150	29.080	150	33.276	150	28.878
		PC	100	39.179	100	39.404	100	38.835
Família 4 $M = P$ $M \neq M^T$ $M > 0$	5	N	19	0.002	19	0.002	19	0.001
		PC	7	0.002	7	0.001	7	0.003
	500	N	14	1.727	14	1.545	14	1.738
		PC	8	1.841	8	1.861	8	1.993
Família 6 $M = P$ $M \neq M^T$ $M \geq 0$	5	N	150	1.805	150	1.416	150	1.087
		PC	100	1.842	100	1.703	100	1.661
	500	N	150	17.401	150	17.439	150	19.357
		PC	100	22.976	100	22.988	100	23.433
Família 7 $M \neq P$ $M = M^T > 0$ $P = P^T > 0$	5	N	19	0.002	19	0.002	19	0.002
		PC	7	0.001	7	0.002	7	0.001
	500	N	16	1.769	14	1.646	16	1.734
		PC	8	1.814	8	1.878	9	2.127
Família 10 $M \neq P$ $M = M^T \geq 0$ $P = P^T \geq 0$	5	N	20	0.002	19	0.004	19	0.002
		PC	8	0.003	7	0.001	8	0.001
	500	N	39*	4.758	25*	3.054	38*	4.726
		PC	11*	2.815	16*	3.934	15*	3.353
Família 11 $M \neq P$ $M \neq M^T > 0$ $P \neq P^T > 0$	5	N	21	0.002	20	0.004	20	0.004
		PC	12	0.002	11	0.002	15	0.005
	500	N	150	18.761	36	4.440	27	3.128
		PC	100	22.959	22	5.223	100	23.590
Família 14 $M \neq P$ $M \neq M^T \geq 0$ $P \neq P^T \geq 0$	5	N	21	0.004	150	0.028	21	0.002
		PC	100	0.031	11	0.004	12	0.004
	500	N	150	18.776	150	17.471	150	18.806
		PC	100	24.291	100	23.129	100	23.822

Tabela B.1: Problemas gerados aleatoriamente no ortante positivo.

Família	n	Método	a	b	c
Família 10	500	N	8.80×10^{-9}	7.87×10^{-10}	1.49×10^{-8}
		PC	9.20×10^{-9}	6.70×10^{-10}	2.04×10^{-7}

Tabela B.2: $\|H_k\|_1$ quando o algoritmo parou devido a $\|w_k - w_{k-1}\|_1 < \epsilon$.

n	Método	a		b		c	
		ITER	CPU	ITER	CPU	ITER	CPU
4	N	19	0.008	19	0.006	19	0.006
	PC	6	0.006	6	0.004	6	0.003
40	N	9	0.009	9	0.007	9	0.004
	PC	5	0.007	5	0.007	5	0.003
400	N	7	0.314	7	0.302	7	0.190
	PC	6	0.225	6	0.172	6	0.257
4000	N	7	24.182	7	25.236	7	24.331
	PC	6	20.476	6	21.886	6	20.531
8000	N	7	131	7	134	7	139
	PC	6	114	6	121	6	123
16000	N	7	580	7	579	7	580
	PC	6	505	6	508	6	485

Tabela B.3: Problemas de complementaridade implícita POZ1.

n	Método	a		b		c	
		ITER	CPU	ITER	CPU	ITER	CPU
4	N	19	0.006	19	0.003	19	0.003
	PC	6	0.003	6	0.004	6	0.002
	PCM	6	0.003	6	0.003	6	0.002
40	N	9	0.009	9	0.006	9	0.005
	PC	5	0.007	5	0.004	6	0.005
	PCM	5	0.007	5	0.003	5	0.003
400	N	7	0.203	7	0.216	8	0.225
	PC	6	0.203	6	0.171	6	0.187
	PCM	6	0.187	6	0.172	6	0.191
4000	N	7	23.502	7	23.885	7	24.097
	PC	6	21.403	6	22.386	6	19.516
	PCM	6	21.114	6	20.707	6	22.148
8000	N	7	132	7	133	7	131
	PC	6	111	6	108	6	109
	PCM	6	116	6	115	6	115
16000	N	7	562	7	564	7	572
	PC	6	519	6	524	6	494
	PCM	6	492	6	500	6	510

Tabela B.4: Problemas de complementaridade implícita POZ2.

n	Método	a		b		c	
		ITER	CPU	ITER	CPU	ITER	CPU
4	N	18	0.005	18	0.006	18	0.003
	PC	4	0.003	5	0.003	5	0.003
	PCM	4	0.001	5	0.003	5	0.003
40	N	8	0.017	8	0.020	8	0.019
	PC	5	0.016	5	0.015	5	0.015
	PCM	5	0.015	5	0.015	5	0.016
400	N	6	0.281	7	0.187	6	0.203
	PC	5	0.234	5	0.172	6	0.156
	PCM	5	0.171	6	0.196	6	0.156
4000	N	6	26.00	7	27.97	7	29.03
	PC	6	25.79	6	24.93	6	24.96
	PCM	6	25.27	6	24.82	6	25.02
8000	N	6	129.9	7	152.7	7	150.5
	PC	6	132.7	6	131.3	6	130.9
	PCM	6	131.5	6	132.5	6	132.0
16000	N	6	529.0	7	631.5	7	620.2
	PC	6	534.3	6	539.5	6	546.8
	PCM	6	529.7	6	513.7	6	550.8

Tabela B.5: Problema de complementaridade implícita POZ3.

(n, q, s)	Método	a		b		c	
		ITER	CPU	ITER	CPU	ITER	CPU
(10,5,1)	N	25	0.015	25	0.014	25	0.015
	PCstd	11	0.011	12	0.013	18	0.015
	PCmod	10	0.015	10	0.010	17	0.015
(10,15,1)	N	30	0.063	29	0.062	29	0.047
	PCstd	11	0.015	14	0.031	12	0.031
	PCmod	12	0.032	12	0.031	13	0.031
(100,50,5)	N	64	0.889	53	0.702	60	1.061
	PCstd	18	0.515	16	0.577	20	0.702
	PCmod	22	0.639	20	0.562	24	0.780
(100,100,5)	N	65	2.871	65	2.776	67	2.464
	PCstd	16	0.874	16	0.905	17	1.076
	PCmod	20	1.123	22	1.216	23	1.311
(100,150,5)	N	90*	4.970	92*	5.026	97	4.992
	PCstd	24*	2.037	28*	2.370	23*	2.093
	PCmod	27	2.481	30*	2.714	37	3.198
(500,250,25)	N	111*	273	135*	311	132*	293
	PCstd	24*	80	27*	89	34	91
	PCmod	36*	125	47*	163	42*	139

Tabela B.6: Família 1, matrizes indefinidas e não simétricas.

(n, q, s)	Método	a	b	c
(100,150,5)	N	1.03×10^{-10}	1.14×10^{-10}	-
	PCstd	2.29×10^{-10}	1.32×10^{-8}	1.62×10^{-8}
	PCmod	-	1.48×10^{-10}	-
(500,250,25)	N	7.21×10^{-10}	2.05×10^{-10}	1.99×10^{-10}
	PCstd	4.73×10^{-8}	2.67×10^{-7}	-
	PCmod	1.23×10^{-6}	4.77×10^{-6}	6.87×10^{-6}

Tabela B.7: $\|H_k\|_1$ quando o algoritmo parou devido a $\|w_k - w_{k-1}\|_1 < \epsilon$ (família 1).

(n, q, s)	Método	a		b		c	
		ITER	CPU	ITER	CPU	ITER	CPU
(10,5,1)	N	21	0.016	21	0.015	21	0.012
	PCstd	14	0.012	9	0.008	8	0.006
	PCmod	11	0.012	8	0.010	8	0.013
(10,15,1)	N	25	0.062	25	0.047	29	0.046
	PCstd	12	0.031	12	0.032	14	0.046
	PCmod	20*	0.251	19*	0.148	14	0.015
(100,50,5)	N	60	0.811	66	0.952	63	0.998
	PCstd	16	0.624	19	0.531	17	0.640
	PCmod	22	0.624	24	0.702	19	0.484
(100,100,5)	N	84	3.869	80	3.369	82	3.666
	PCstd	19	1.217	19	1.108	21	1.217
	PCmod	29	1.701	31	1.841	31	1.496
(100,150,5)	N	79*	4.335	78*	4.012	77*	4.219
	PCstd	41*	3.436	24*	2.087	30*	2.636
	PCmod	27*	2.636	27*	2.449	24	2.215
(500,250,25)	N	118*	266	128*	279	150*	317
	PCstd	27	87	30	91	24	74
	PCmod	35*	123	41*	138	32*	110

Tabela B.8: Família 2, matrizes indefinidas e simétricas.

(n, q, s)	Método	a	b	c
(10,15,1)	PCmod	1.05×10^{-10}	2.11×10^{-10}	-
(100,150,5)	N	3.49×10^{-10}	1.94×10^{-10}	2.73×10^{-10}
	PCstd	2.70×10^{-8}	1.84×10^{-8}	8.33×10^{-8}
	PCmod	2.33×10^{-9}	2.34×10^{-9}	-
(500,250,25)	N	4.70×10^{-9}	6.57×10^{-10}	2.03×10^{-8}
	PCstd	-	-	-
	PCmod	6.87×10^{-8}	1.66×10^{-10}	5.77×10^{-8}

Tabela B.9: $\|H_k\|_1$ quando o algoritmo parou devido a $\|w_k - w_{k-1}\|_1 < \epsilon$ (família 2).

(n, q, s)	Método	a		b		c	
		ITER	CPU	ITER	CPU	ITER	CPU
(10,5,1)	N	18	0.016	19	0.015	19	0.013
	PCstd	9	0.010	7	0.010	7	0.008
	PCmod	9	0.010	6	0.007	7	0.008
(10,15,1)	N	150	0.131	150	0.147	150	0.109
	PCstd	100	0.125	100	0.109	100	0.125
	PCmod	100	0.125	100	0.110	100	0.062
(100,50,5)	N	150	1.709	150	1.830	150	1.884
	PCstd	100	2.699	100	2.464	100	2.730
	PCmod	100	2.528	100	2.855	100	2.605
(100,100,5)	N	150	5.838	150	5.563	150	5.514
	PCstd	100	6.411	100	6.240	100	5.803
	PCmod	100	5.882	100	5.866	100	5.740
(100,150,5)	N	150	8.967	150	8.796	150	9.264
	PCstd	100	10.796	100	10.639	100	10.202
	PCmod	100	9.906	100	10.280	100	9.921
(500,250,25)	N	150	285	150	284	150	285
	PCstd	100	273	100	271	100	271
	PCmod	100	327	100	319	100	297

Tabela B.10: Família 3, matrizes indefinidas, não simétricas e singulares.

(n, q, s)	Método	a		b		c	
		ITER	CPU	ITER	CPU	ITER	CPU
(10,5,1)	N	27	0.015	28	0.012	28	0.015
	PCstd	14	0.015	14	0.009	13	0.009
	PCmod	13	0.015	13	0.009	12	0.010
(10,15,1)	N	27	0.031	30	0.047	34	0.031
	PCstd	15	0.062	17	0.072	17	0.063
	PCmod	17	0.032	17	0.036	16	0.047
(100,50,5)	N	60	0.702	69	0.895	70	0.912
	PCstd	19	0.457	19	0.546	21	0.516
	PCmod	29*	0.733	24	0.493	28	0.629
(100,100,5)	N	66	2.043	66	2.075	69	2.064
	PCstd	15	0.799	17	0.786	17	0.846
	PCmod	25	1.093	23	0.983	25	1.080
(100,150,5)	N	69	4.384	76	4.508	72	4.165
	PCstd	33*	3.266	39*	4.524	22*	2.387
	PCmod	32*	3.276	36*	3.588	26*	2.683
(500,250,25)	N	95	235	109*	275	128	293
	PCstd	23*	84	66	208	24*	86
	PCmod	46*	163	28	97	34*	122

Tabela B.11: Família 4, matrizes indefinidas, não simétricas e solução sem complementaridade estrita.

(n, q, s)	Método	a	b	c
(100,50,5)	PCmod	3.98×10^{-10}	-	-
(100,150,5)	PCstd	2.52×10^{-9}	4.80×10^{-9}	1.61×10^{-8}
	PCmod	2.92×10^{-7}	3.29×10^{-7}	3.36×10^{-9}
(500,250,25)	N	-	1.73×10^{-10}	-
	PCstd	1.24×10^{-8}	-	1.53×10^{-9}
	PCmod	3.26×10^{-6}	-	3.96×10^{-8}

Tabela B.12: $\|H_k\|_1$ quando o algoritmo parou devido a $\|w_k - w_{k-1}\|_1 < \epsilon$ (família 4).

q	Solução	r	Método	a		b		c	
				ITER	CPU	ITER	CPU	ITER	CPU
3	interior	0.1	N	27	0.018	26	0.009	26	0.009
			PC	10	0.019	11	0.005	17	0.008
		10	N	24	0.004	25	0.004	150	0.039
			PC	10	0.002	9	0.002	100	0.040
	borda	0.1	N	26	0.004	26	0.004	26	0.004
			PC	7	0.001	8	0.001	7	0.001
		10	N	29	0.004	29	0.005	31	0.005
			PC	15	0.003	15	0.003	14	0.006
12	interior	0.1	N	150	0.089	25	0.015	150	0.076
			PC	14	0.010	12	0.009	12	0.009
		10	N	15	0.009	15	0.004	15	0.005
			PC	8	0.003	8	0.003	8	0.003
	borda	0.1	N	18	0.005	15	0.004	20	0.008
			PC	32	0.012	10	0.004	100	0.069
		10	N	15	0.005	17	0.005	17	0.005
			PC	11	0.008	18	0.007	15	0.009
20	interior	0.1	N	15	0.015	11	0.009	16	0.008
			PC	8	0.010	10	0.008	7	0.008
		10	N	14	0.007	16	0.008	15	0.007
			PC	10	0.011	10	0.013	11	0.006
	borda	0.1	N	15	0.014	15	0.014	15	0.008
			PC	11	0.014	13	0.015	11	0.008
		10	N	15	0.007	15	0.007	14	0.014
			PC	10	0.006	10	0.006	9	0.013
40	interior	0.1	N	11	0.022	16	0.019	16	0.019
			PC	10	0.014	8	0.010	7	0.009
		10	N	11	0.014	12	0.022	11	0.014
			PC	7	0.014	7	0.011	7	0.010
	borda	0.1	N	9	0.011	9	0.018	11	0.016
			PC	10	0.023	7	0.008	9	0.022
		10	N	14	0.037	14	0.027	14	0.017
			PC	9	0.026	51	0.129	16	0.024

Tabela B.13: Família 1, T indefinida.

q	Solução	r	Método	a		b		c	
				ITER	CPU	ITER	CPU	ITER	CPU
3	interior	0.1	N	26	0.093	26	0.008	25	0.010
			PC	8	0.032	7	0.003	8	0.003
		10	N	24	0.020	25	0.008	26	0.004
			PC	6	0.003	6	0.001	6	0.003
	borda	0.1	N	26	0.005	26	0.004	26	0.004
			PC	7	0.001	7	0.001	7	0.001
		10	N	25	0.007	26	0.004	27	0.009
			PC	8	0.002	7	0.001	9	0.003
12	interior	0.1	N	15	0.009	14	0.008	14	0.010
			PC	10	0.010	8	0.006	8	0.008
		10	N	14	0.007	14	0.005	15	0.006
			PC	7	0.003	7	0.003	7	0.005
	borda	0.1	N	14	0.005	13	0.026	14	0.004
			PC	10	0.004	7	0.006	8	0.003
		10	N	13	0.005	15	0.010	15	0.009
			PC	7	0.006	8	0.006	8	0.006
20	interior	0.1	N	14	0.006	13	0.007	13	0.006
			PC	8	0.005	8	0.004	8	0.005
		10	N	12	0.012	13	0.009	12	0.013
			PC	7	0.008	8	0.007	7	0.009
	borda	0.1	N	15	0.007	15	0.014	17	0.017
			PC	9	0.007	10	0.012	10	0.008
		10	N	14	0.006	14	0.007	14	0.009
			PC	9	0.005	9	0.005	8	0.010
40	interior	0.1	N	11	0.014	11	0.013	12	0.014
			PC	7	0.011	7	0.021	7	0.021
		10	N	11	0.021	12	0.031	12	0.015
			PC	7	0.022	8	0.012	7	0.011
	borda	0.1	N	15	0.028	18	0.044	17	0.041
			PC	10	0.018	11	0.032	10	0.028
		10	N	14	0.017	16	0.039	15	0.024
			PC	10	0.030	10	0.026	10	0.030

Tabela B.14: Família 2, T definida positiva.

q	Solução	r	Método	a		b		c	
				ITER	CPU	ITER	CPU	ITER	CPU
3	interior	0.1	N	25	0.008	25	0.007	22	0.007
			PC	8	0.003	9	0.004	8	0.003
		10	N	21	0.015	20	0.012	21	0.012
			PC	17	0.009	7	0.010	100	0.024
	borda	0.1	N	25	0.008	25	0.008	24	0.004
			PC	8	0.004	8	0.002	8	0.003
		10	N	24	0.004	150	0.024	150	0.038
			PC	7	0.003	100	0.041	100	0.053
12	interior	0.1	N	150	0.046	150	0.083	150	0.057
			PC	100	0.037	100	0.072	100	0.037
		10	N	13	0.004	13	0.015	14	0.008
			PC	15	0.012	11	0.015	9	0.003
	borda	0.1	N	18	0.006	12	0.007	13	0.004
			PC	8	0.006	9	0.007	9	0.003
		10	N	14	0.006	15	0.012	15	0.009
			PC	8	0.003	11	0.004	12	0.009
20	interior	0.1	N	11	0.010	12	0.010	14	0.012
			PC	10	0.013	9	0.010	10	0.013
		10	N	150	0.058	150	0.076	13	0.006
			PC	8	0.005	100	0.064	100	0.096
	borda	0.1	N	150	0.088	150	0.086	150	0.080
			PC	100	0.057	100	0.055	100	0.053
		10	N	16	0.016	18	0.010	16	0.010
			PC	100	0.068	16	0.011	100	0.080
40	interior	0.1	N	11	0.029	16	0.040	16	0.026
			PC	11	0.034	8	0.024	17	0.050
		10	N	10	0.014	11	0.018	10	0.032
			PC	7	0.012	7	0.015	7	0.032
	borda	0.1	N	14	0.014	150	0.225	150	0.153
			PC	12	0.030	100	0.141	100	0.192
		10	N	17	0.024	16	0.022	14	0.022
			PC	11	0.019	14	0.022	11	0.019

Tabela B.15: Família 4, T semidefinida negativa.