

EUDA MARA DA SILVA FERREIRA

**ALGORITMOS PARA O PROBLEMA DE EQUILÍBRIO
APLICADOS AO PROBLEMA DE EQUILÍBRIO DE NASH**

CURITIBA

2013

EUDA MARA DA SILVA FERREIRA

**ALGORITMOS PARA O PROBLEMA DE EQUILÍBRIO
APLICADOS AO PROBLEMA DE EQUILÍBRIO DE NASH**

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática do Departamento de Matemática, Setor de Ciências Exatas e do Departamento de Construção Civil, Setor de Tecnologia, Universidade Federal do Paraná.

Orientador: Prof. Dr. Luiz Carlos Matioli

CURITIBA

2013

F383a

Ferreira, Euda Mara da Silva

Algoritmos para o problema de equilíbrio aplicados ao Problema de Equilíbrio de Nash / Euda Mara da Silva Ferreira. – Curitiba, 2013.
101f. : il. color. ; 30 cm.

Tese (doutorado) - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-graduação em Métodos Numéricos Métodos Numéricos em Engenharia, 2013.

Orientador: Luiz Carlos Matioli.

Bibliografia: p. 99-101.

1. Teoria dos Jogos. 2. Equilíbrio econômico. 3. Algoritmos. I. Universidade Federal do Paraná. II. Matioli, Luiz Carlos. III. Título.

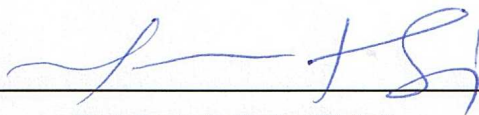
CDD: 519.3

TERMO DE APROVAÇÃO

EUDA MARA DA SILVA FERREIRA


“Algoritmos para o Problema de Equilíbrio Aplicados ao Problema de Equilíbrio de Nash”

Tese aprovada como requisito parcial para obtenção do grau de Doutor no Curso de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, pela seguinte banca examinadora:



Prof. Dr. Luiz Carlos Matioli.

(Orientador) Membro do PPGMNE/UFPR



Prof.ª Dr.ª Susana Scheimberg de Makler.

Membro do PESC-UFRJ



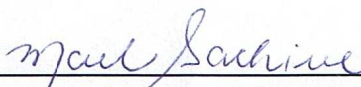
Prof. Dr. Juliano de Bem Francisco.

Membro do Dep. de Matemática da UFSC



Prof.ª Dr.ª Solange Regina dos Santos.

Membro do Dep. de Matemática da UNESPAR



Prof.ª Dr.ª Mael Sachine.

Membro do Dep. de Matemática da UFPR

Curitiba, 20 de setembro 2013.

Dedicatória

Dedico a DEUS, por todas as vitórias alcançadas. Em especial ao meu esposo, Misael, meus filhos Abner e Joabe, pelo apoio, carinho e compreensão em todos os momentos.

Agradecimentos

A Deus pela oportunidade de realizar este Curso. Ao meu esposo Misael, meus filhos, Abner e Joabe, por serem compreensíveis e me apoiarem durante o período que me dediquei à realização deste trabalho. Aos meus queridos pais, Ival e Aparecida pelo carinho, incentivo e apoio constante. Também agradeço ao meu orientador Professor Dr Luiz Carlos Matioli, pela dedicação e por me ajudar a ampliar os meus conhecimentos. Enfim, a todos que, direta ou indiretamente, contribuíram para a realização desta tese.

Sumário

Lista de Figuras	vii
Lista de Algoritmos	viii
Resumo	xi
Abstract	xii
1 INTRODUÇÃO	1
2 PROBLEMA DE EQUILÍBRIO DE NASH GENERALIZADO	4
2.1 Exemplo de um GNEP.....	5
3 CONCEITOS FUNDAMENTAIS	7
3.1 BUSCA LINEAR.....	7
3.2 MÉTODO DE CAUCHY.....	8
3.3 MÉTODO DE NEWTON.....	9
3.3.1 Método de Newton para Zeros de Função.....	9
3.3.2 Método de Newton para Minimização Irrestrita.....	11
3.4 MÉTODOS QUASE-NEWTON.....	13
3.4.1 Métodos Quase-Newton para Resolução de Sistema não Linear.....	13
3.4.2 Método Quase-Newton para Minimização Irrestrita.....	15
3.5 MÉTODO DE GAUSS-NEWTON.....	16
3.6 MÉTODO BARZILAI E BORWEIN.....	17
3.7 REGIÃO DE CONFIANÇA.....	19

3.8	MÉTODO DE LEVENBERG-MARQUARDT	21
3.9	CONDIÇÕES DE KKT	23
3.10	MÉTODOS DE PENALIDADES	25
3.10.1	Penalidades Internas (ou Métodos de Barreiras)	26
3.10.2	Penalidades Externas	27
3.11	MÉTODOS DE LAGRANGEANOS AUMENTADOS	28
3.11.1	Problema de Minimização com Restrições de Igualdade	28
3.11.2	Problemas de Minimização com Restrições de Desigualdade	30
3.12	SUBGRADIENTE	35
3.13	MÉTODOS SUBGRADIENTES	39
3.13.1	Direção	39
3.13.2	Escolha do Tamanho do Passo	40
3.13.3	Critério de Parada	41
3.13.4	Método Subgradiente Básico	41
3.13.5	Método Subgradiente quando f^* é conhecido	42
3.13.6	Método de Dilatação do Espaço - Shor	43
3.13.7	Método Subgradiente com Dilatação do Espaço na Direção da Diferença de Dois Subgradients Sucessivos	44
3.13.8	Método de Dilatação do Espaço na Direção do Subgradiente no Ponto Atual	46
3.13.9	Método de Elipsóides	46
4	REFORMULAÇÕES DO NEP E GNEP	48
4.1	FUNÇÃO NIKAIDO-ISODA (NI-função)	48
4.2	O CASO JUNTAMENTE CONVEXO	52
4.3	DESIGUALDADE VARIACIONAL (VI)	52
4.3.1	Condições de KKT e Desigualdade Variacional (VI)	53
4.4	REFORMULAÇÃO VIA PONTO PROXIMAL	55

5	ALGORITMOS PARA A SOLUÇÃO DO NEP E GNEP	59
5.1	MÉTODOS PRÁTICOS	59
5.2	MÉTODO DE PENALIDADE EXATA	60
5.3	MÉTODO DE NEWTON LOCAL	61
5.4	MÉTODOS DE LAGRANGEANO AUMENTADO PARA PROBLEMA DE EQUILÍBRIO	62
5.4.1	Lagrangeano Aumentado para Problema de Equilíbrio	64
5.4.2	Versão Linearizada do Lagrangeano Aumentado	68
6	METODOLOGIA PROPOSTA	71
6.1	APROXIMAÇÃO QUADRÁTICA DO TERMO NÃO DIFERENCIÁVEL ...	72
6.2	MÉTODO DE NEWTON PARA RESOLUÇÃO DO SUBPROBLEMA GERADO PELO ALGORITMO DE EQUILÍBRIO	76
6.3	MÉTODO SUBGRADIENTE PARA RESOLUÇÃO DO SUBPROBLEMA GERADO PELO ALGORITMO DE EQUILÍBRIO ...	80
7	RESULTADOS NUMÉRICOS	86
7.1	PROBLEMAS MENORES	87
7.2	PROBLEMAS MAIORES	94
8	CONSIDERAÇÕES FINAIS	97
	Referências	99

Lista de Figuras

Figura 1	FAMÍLIA DE FUNÇÕES PENALIDADES	33
Figura 2	SUBGRADIENTES	37
Figura 3	APROXIMAÇÃO QUADRÁTICA	74

Lista de Algoritmos

Algoritmo	MÉTODO DE CAUCHY PARA O CASO NÃO QUADRÁTICO ..	8
Algoritmo	MÉTODO DE NEWTON PARA SISTEMAS DE EQUAÇÕES ...	10
Algoritmo	MÉTODO DE NEWTON PURO PARA MINIMIZAÇÃO	11
Algoritmo	MÉTODO DE NEWTON PARA MINIMIZAÇÃO COM BUSCA .	12
Algoritmo	MÉTODO QUASE-NEWTON PARA SOLUÇÃO DE SISTEMA NÃO LINEAR	14
Algoritmo	MÉTODO QUASE-NEWTON PARA PROBLEMA IRRESTRITO	15
Algoritmo	MÉTODO DE GAUSS-NEWTON	17
Algoritmo	MÉTODO GRADIENTE DE BARZILAI E BORWEIN	19
Algoritmo	REGIÃO DE CONFIANÇA	20
Algoritmo	MÉTODO DE LEVENBERG-MARQUARDT	22
Algoritmo	MÉTODO DE BARREIRAS	26
Algoritmo	PENALIDADE EXTERNA	28

Algoritmo	LAGRANGEANO AUMENTADO CLÁSSICO PARA PROBLEMA DE MINIMIZAÇÃO COM RESTRIÇÃO DE IGUALDADE	29
Algoritmo	LAGRANGEANO AUMENTADO PARA PROBLEMA DE MINIMIZAÇÃO COM RESTRIÇÃO DE DESIGUALDADE E PENALIDADE DA FAMÍLIA P	34
Algoritmo	MÉTODO BÁSICO DO SUBGRADIENTE	41
Algoritmo	MÉTODO SHOR DE DILATAÇÃO DE ESPAÇO	44
Algoritmo	MÉTODO DE DILATAÇÃO DO ESPAÇO NA DIREÇÃO DO SUBGRADIENTE NO PONTO ATUAL	46
Algoritmo	MÉTODO ELIPSÓIDES	47
Algoritmo	MÉTODO NÃO LINEAR DE JACOBI	59
Algoritmo	MÉTODO NÃO LINEAR DE GAUSS-SEIDEL	60
Algoritmo	PENALIDADE EXATA	61
Algoritmo	MÉTODO LAGRANGEANO AUMENTADO EXATO (EMLA) ..	65
Algoritmo	MÉTODO LAGRANGEANO AUMENTADO EXTRAGRADIENTE INEXATO (IALEM)	65
Algoritmo	MÉTODO DE PONTO PROXIMAL EXTRAGRADIENTE INEXATO (IPPEM)	66

Algoritmo	MÉTODO DE LAGRANGEANO AUMENTADO EXTRAGRADI- ENTE INEXATO LINEARIZADO (LIALEM)	68
Algoritmo	PROBLEMA DE EQUILÍBRIO COMBINADO COM MÉTODO DE NEWTON PARA RESOLVER O SUBPROBLEMA	78
Algoritmo	PROBLEMA DE EQUILÍBRIO COMBINADO COM GRADIENTE PARA RESOLVER O SUBPROBLEMA	82
Algoritmo	PROBLEMA DE EQUILÍBRIO COMBINADO COM O MÉTODO SUBGRADIENTE PARA RESOLVER O SUBPROBLEMA	83

Resumo

Nesta pesquisa são apresentados dois novos algoritmos para resolução do Problema de Equilíbrio de Nash (NEP), ambos baseados na resolução de um sistema não linear $G(x) = 0$, sendo $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ contínua, mas não diferenciável em todos os pontos do domínio. Assim, a origem desses métodos é o artigo de IUSEM e NASRI (2007a) em que foi introduzido o método de Lagrangeano Aumentado, para a resolução de um Problema de Equilíbrio geral, do qual o Problema de Equilíbrio de Nash é um caso particular. Existem algumas dificuldades com relação à solução do sistema $G(x) = 0$, a saber: a falta de diferenciabilidade e de convexidade. O que pode ser garantido é a continuidade da $G(x)$. Para superar as dificuldades apresentadas serão desenvolvidas duas metodologias diferentes para resolver $G(x) = 0$. A primeira consiste na suavização do termo não diferenciável para tornar as funções que definem $G(x) = 0$ continuamente diferenciáveis. Após a suavização, será aplicado o Método de Newton para resolver o sistema não linear. A segunda consiste em resolver o sistema $G(x) = 0$ por meio de um problema de otimização, ou seja, ao invés de resolver $G(x) = 0$ será resolvido o problema $\text{minimizar}\{f(x) : x \in \mathbb{R}^n\}$ com $f(x) = \frac{1}{2}\|G(x)\|^2$. Para isto, será utilizado um procedimento similar ao Método do Gradiente. Primeiramente, suaviza-se as funções em $G(x)$ que são não diferenciáveis, como realizado para o Método de Newton, aplica-se o Método do Gradiente com busca Barzilai Borwein para resolver $G(x) = 0$. Uma vez que o Método do Gradiente tem convergência lenta, do ponto de vista computacional pode até não convergir, este será substituído por Métodos Subgradientes para resolver o sistema $G(x) = 0$. Portanto, a principal contribuição deste trabalho é a apresentação de duas novas metodologias para a resolução do Problema de Equilíbrio de Nash. A primeira, baseada no Método de Newton, e a segunda, em Métodos Subgradientes para resolver um sistema não linear e não diferenciável $G(x) = 0$.

Palavras-chave: Problema de Equilíbrio de Nash; Problema de Equilíbrio de Nash Generalizado; Métodos Subgradientes; Método de Newton; Problema de Equilíbrio; Algoritmos para solução de um NEP.

Abstract

This research presents two new algorithms in order to resolve the Nash Equilibrium Problem (NEP), both are based on the resolution of a none linear $G(x) = 0$ system, where $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuous, but not differentiable in all points of domain. Therefore, the origin of these methods is the article of IUSEM e NASRI (2007a) where the method of increased lagrangian was introduced, for solving a Generalized Nash Equilibrium Problem, where the equilibrium is a particular case. There are certain difficulties in relation with the solution of the $G(x) = 0$ system, such as, the lack of differentiability and convexity, however the continuity of $G(x)$ is certain. To overcome the presented difficulties, two different methodologies shall be developed to resolver $G(x) = 0$. The first one consists in smoothing the nondifferentiable term, for the functions that define $G(x) = 0$ become continuously differentiable. After the smoothing a Newton method shall be applied to resolve the none-linear system. And, the second one consists in resolving the $G(x) = 0$ system through a problem of optimization, that is, instead of resolving $G(x) = 0$ the problem of $minimize\{f(x) : x \in \mathbb{R}^n\}$ shall be resolved with $f(x) = \frac{1}{2}||G(x)||^2$. Thereby, a similar procedure to the gradient method shall be used. Firstly, the functions in $G(x)$ that are differentiable, shall be smoothed, as held in the Newton method, by applying the method of gradient with Barzilai Borwein research to resolve $G(x) = 0$. Once the method of gradient has a slow convergence and from a computational point of view might even not converge, this method shall be replaced by Subgradient Methods to resolve the $G(x) = 0$ system. Thence, the main contribution of this paper is the presentation of two new methodologies for the resolution of the Nash Equilibrium Problem, where the first is based on the Newton Method and the second on the Subgradient Methods for resolving a none-linear and not differentiable $G(x) = 0$.

Keywords : Nash Equilibrium Problem; Generalized Nash Equilibrium Problem; Subgradient Methods; Newton Method; Equilibrium Problem; Algorithms for a NEP solution.

1 INTRODUÇÃO

Neste trabalho será abordado o Problema de Equilíbrio de Nash (NEP) e o Problema de Equilíbrio de Nash Generalizado (GNEP), suas reformulações e seus principais métodos de solução, pois, ultimamente, esse assunto tem sido alvo de muita atenção de pesquisadores na área de economia, matemática, informática e engenharia, apresentando-se como uma excelente ferramenta para solução de problemas com n-jogadores.

A origem da teoria do equilíbrio iniciou-se com o matemático e economista Cournot (FACCHINEI; KANZOW, 2010a), com o estudo ‘Oligopolio’, sobre várias empresas, que ofertavam produtos homogêneos, deixando de ser o preço o diferencial entre elas, e sim a qualidade do produto, divulgação da marca e a dificuldade na entrada de novas empresas na competição de mercado. Esta teoria se fortaleceu nos trabalhos de Von Neumann, Von Neumann e Morgenstern (FACCHINEI; KANZOW, 2010a), em que resolveram jogos não-cooperativos com estratégias puras ou mistas e soma zero. Ou seja, num jogo com dois jogadores, um ganha e outro perde. A soma dos ganhos dos dois jogadores é igual a zero. A teoria desenvolvida por Von Neumann e Morgenstern proporcionou uma solução para tais jogos, onde se encontra um ponto de equilíbrio entre os interesses dos dois jogadores, de forma que nenhum saia perdendo totalmente, os chamados jogos cooperativos. John Forbes Nash Junior ampliou a teoria de Von Neumann e Morgenstern desenvolvida em ‘Teoria dos Jogos e o Comportamento Econômico’ (NASH, 1950) e demonstrou matematicamente o equilíbrio para jogos de n-pessoas com estratégias não-cooperativas. Ou seja, não há acordos ou colaboração entre os jogadores. Este estudo proporcionou a Nash entrar para história com uma tese de apenas 27 páginas que revolucionou a área matemática da Teoria dos Jogos com o título: Non-cooperative Games (Jogos não cooperativos) com a qual obteve o título de doutor em 1950, pela Universidade de Princeton. Sua tese, mais tarde ficou conhecida como ‘Equilíbrio de Nash’, e lhe rendeu o prêmio Nobel da Economia, em 1994 (JOHN, 2011). O Problema de Equilíbrio de Nash (NEP) foi amplamente aprovado ser eficiente, flexível e rico de consequências (FACCHINEI; KANZOW, 2010a).

A teoria do Equilíbrio de Nash foi generalizada por Debreu e formalmente abordada em seu trabalho, ‘Um teorema de existência social de equilíbrio’, em 1952, quando

foi introduzido o termo GNEP - Problema do Equilíbrio de Nash Generalizado. Este trabalho serviu de base matemática para Arrow e Debreu escreverem o artigo, 'Existência de um Equilíbrio para Economia Competitiva', em 1954 (FACCHINEI; KANZOW, 2010a).

Após o trabalho de Arrow e Debreu, pesquisadores concentraram seus estudos na existência de equilíbrio, porém utilizando hipóteses fracas como unicidade ou singularidade local. A partir da década de 1990, o GNEP passou a ser aplicado fora do âmbito econômico, refletindo conseqüentemente na necessidade de algoritmos para seu cálculo. Foi nesta época que começaram as aplicações interessantes na área da engenharia das estruturas de telecomunicações, de informática para a análise de mercados e de cenários de poluição. Nesta década houve dois nomes que se destacaram, sendo Robinson, em 1993, com o trabalho, 'Os preços-sombra para as medidas de eficácia', modelo Linear e modelo Geral, em que otimiza-se modelos de combates com formulações de um GNEP, e Scotti, em 1995, que introduz o GNEP no estudo e solução de problemas complexos com o uso de técnicas de programação não-linear (FACCHINEI; KANZOW, 2010a).

O GNEP é um estudo interdisciplinar, sua aplicação pode ser encontrada nas áreas de economia, matemática, engenharia, ciências da computação e pesquisa operacional. Devido à sua origem matemático-econômica, há várias aplicações no meio econômico e na Teoria dos Jogos, porém pode ser aplicado em situações em que as estratégias de decisão de um depende das estratégias de decisão dos outros envolvidos e o objetivo é encontrar um equilíbrio. Ou seja, a melhor estratégia, de forma que nenhum dos envolvidos possa melhorar unilateralmente a sua.

Esta tese apresenta algoritmos para solução dos problemas de Equilíbrio de Nash e Equilíbrio de Nash Generalizado. Para o desenvolvimento deste trabalho serão apresentados estudos já desenvolvidos sobre o NEP e o GNEP, sendo o Problema de Equilíbrio de Nash o objetivo principal desta pesquisa. Poderá ser observado que existem mais pesquisas sobre a existência de soluções, do que boas implementações. Com o objetivo de avançar a pesquisa, serão propostos dois novos algoritmos baseados na resolução de um sistema não linear $G(x) = 0$, com $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ contínua mas não diferenciável em todos os pontos do domínio, sendo o primeiro algoritmo baseado, no Método de Newton, e o segundo em Métodos Subgradientes. Para ambos os métodos, devido à falta de diferenciabilidade e de convexidade da $G(x)$, embora sua continuidade possa ser garantida, será realizada uma suavização do termo não diferenciável com objetivo de tornar $G(x)$ continuamente diferenciável. A origem desses algoritmos é o artigo de IUSEM e NASRI (2007a) em que foi introduzido o método de Lagrangeano Aumentado, para a resolução de um Problema de Equilíbrio geral, em que o Problema Equilíbrio de Nash é um caso particular.

Este trabalho será desenvolvido com base em pesquisa bibliográfica e implementação computacional, assim como a proposta de novos algoritmos.

A distribuição do trabalho será da seguinte forma, o capítulo a seguir apresenta a definição dos Problemas de Equilíbrio de Nash Generalizado e Equilíbrio de Nash; no Capítulo 3, os conceitos básicos de programação não linear e Métodos Subgradientes, que auxiliarão no entendimento do NEP, GNEP e da metodologia proposta. No Capítulo 4, apresenta-se as reformulações do GNEP, sendo as mais importantes a reformulação Nikaido-Isoda e a reformulação Desigualdade Variacional (VI). Este assunto é tratado com mais detalhes por: FACCHINEI e KANZOW (2010a); NABETANI (2008); HEUSINGER e KANZOW (2009), HEUSINGER e KANZOW (2008) e HEUSINGER (2009). Neste mesmo capítulo, será apresentada a reformulação via Ponto Proximal, que é uma teoria mais geral de Equilíbrio (IUSEM; SOSA, 2007), em que o NEP é um caso particular. No capítulo 5, serão apresentados alguns algoritmos de solução para o GNEP apresentados por FACCHINEI e KANZOW (2010a), sendo eles: método não linear de Jacobi, método não linear de Gauss-Seidel, método de penalidade exata e método de Newton Local. Também, serão apresentados os métodos de Lagrangeano Aumentado para Problemas de Equilíbrio (IUSEM; NASRI, 2007a). No sexto capítulo, é descrita a metodologia proposta por este trabalho e, por fim, no sétimo capítulo, serão apresentados os resultados numéricos.

2 PROBLEMA DE EQUILÍBRIO DE NASH GENERALIZADO

O Problema de Equilíbrio de Nash Generalizado é uma teoria estudada por vários pesquisadores. Na literatura podem ser encontrados vários termos para ela, por exemplo: pseudo-jogo, problema de equilíbrio social, programação de equilíbrio, restrição do problema de equilíbrio, economia e abstrato. Porém, o termo mais aceito é Problema de Equilíbrio de Nash Generalizado (GNEP) (FACCHINEI; KANZOW, 2010a).

O GNEP consiste em N jogadores, cada jogador é representado por j e controla a variável $x^j \in \mathbb{R}^{n_j}$. Designa-se por x o vetor formado por todas as variáveis de decisão ou jogadas estratégicas:

$$x := (x^1, \dots, x^N).$$

Considere $n := \sum_{j=1}^N n_j$, e $x^{-j} := (x^1, \dots, x^{j-1}, x^{j+1}, \dots, x^N)$ as variáveis de decisão de todos os outros jogadores, exceto o jogador j .

Para enfatizar todas as variáveis de decisão em x , será utilizado (x^j, x^{-j}) em vez de x . Note-se ainda que (x^j, x^{-j}) é o vetor $x := (x^1, \dots, x^j, \dots, x^N)$ e que, a notação (x^j, x^{-j}) não significa que os elementos de x são reordenados de tal forma que x^j se torna o primeiro.

Para cada jogador há uma função objetivo $\theta_j : \mathbb{R}^n \rightarrow \mathbb{R}$ que irá depender tanto da variável x^j do jogador j , bem como das variáveis de decisão x^{-j} dos outros jogadores. A função θ_j é muitas vezes chamada de função utilidade do jogador j , função de retorno ou função de perda, isso dependerá da aplicação do GNEP.

O conjunto $X_j(x^{-j}) \subseteq \mathbb{R}^{n_j}$ representa as jogadas estratégicas, chamado de conjunto viável ou estratégias do jogador j . O objetivo do jogador j é levar em conta as jogadas estratégicas x^{-j} e escolher uma estratégia x^j que resolve o problema de minimização:

$$\begin{aligned} \min_{x^j} \quad & \theta_j(x^j, x^{-j}) \\ \text{s.a} \quad & x^j \in X_j(x^{-j}). \end{aligned} \tag{2.1}$$

O conjunto solução para o problema (2.1) é denotado por $S_j(x^{-j})$. Seja o conjunto $S(x) := \prod_{(j=1)}^N S_j(x^{-j})$, diz-se que \bar{x} é uma solução, ou seja $\bar{x} \in S(\bar{x})$, e portanto \bar{x} é um ponto fixo do mapeamento de todos os pontos do conjunto S.

Um ponto \bar{x} é, portanto, um equilíbrio se nenhum jogador puder melhorar sua função objetivo, ou seja, alterar unilateralmente x^j para qualquer outro ponto viável.

Se o conjunto $X_j(x^{-j})$ não depender das jogadas estratégicas rivais, tem-se $X_j(x^{-j}) = X_j$ para algum conjunto $X_j \subseteq \mathbb{R}^{n_j}$ com $j = 1, \dots, N$. Então, o GNEP passa a ser um Problema de Equilíbrio de Nash (NEP) padrão. Portanto, a diferença entre o Problema de Equilíbrio de Nash (NEP) e o GNEP está no conjunto das estratégias de decisão, ou seja, o conjunto solução de um NEP é um subconjunto do conjunto de todas as soluções de um GNEP (FACCHINEI; KANZOW, 2010a).

A formulação mais comum do Problema de Equilíbrio de Nash Generalizado é com restrições de desigualdade. Portanto, pode-se denotar o conjunto $X_j(x^{-j})$ por

$$X_j(x^{-j}) = \{x^j \in \mathbb{R}^{n_j} : g(x^j, x^{-j}) \leq 0\}, \tag{2.2}$$

em que $g(\cdot, x^{-j}) : \mathbb{R}^{n_j} \rightarrow \mathbb{R}^{m_j}$ representa o conjunto das restrições de todos os jogadores. As restrições de igualdade podem ser facilmente introduzidas em (2.2).

2.1 Exemplo de um GNEP

O exemplo apresentado a seguir foi extraído do artigo FACCHINEI e KANZOW (2010a), p.178.

Exemplo 2.1 *Considere um jogo com dois jogadores, $N = 2$ com $n_1 = 1$ e $n_2 = 1$. Assim, cada jogador controla uma variável. Por simplicidade define-se $x_1^1 = x^1$ e $x_1^2 = x^2$. Assumindo que os Problemas dos jogadores são:*

$$\begin{aligned} \min_{x^1} (x^1 - 1)^2 & & \min_{x^2} (x^2 - \frac{1}{2})^2 \\ \text{s.a} \quad x^1 + x^2 \leq 1, & & \text{s.a} \quad x^1 + x^2 \leq 1. \end{aligned}$$

A solução ótima é dada pelos conjuntos:

$$S_1(x^2) = \begin{cases} 1, & \text{se } x^2 \leq 0 \\ 1 - x^2, & \text{se } x^2 \geq 0 \end{cases} \quad \text{e} \quad S_2(x^1) = \begin{cases} \frac{1}{2}, & \text{se } x^1 \leq \frac{1}{2} \\ 1 - x^1, & \text{se } x^1 \geq \frac{1}{2}. \end{cases}$$

É possível verificar que as soluções deste problema são dadas por $(\alpha, 1 - \alpha)$ para cada $\alpha \in [\frac{1}{2}, 1]$. Observe que o problema possui restrições de desigualdade, com infinitas soluções, e a variável de decisão de um jogador é influenciada pela variável de decisão do jogador rival.

Para apresentar as principais reformulações do GNEP, assim como algoritmos para resolvê-los, serão necessários conceitos matemáticos considerados básicos para pesquisadores em programação matemática, mas que não são conhecidos por aqueles que não são necessariamente da área. Estes conceitos são de fundamental importância para entender a continuação deste trabalho. Sendo assim, no capítulo seguinte, serão desenvolvidos conceitos importantes e que serão fundamentais para o desenvolvimento dos demais tópicos da tese, bem como da metodologia proposta.

3 CONCEITOS FUNDAMENTAIS

O objetivo deste capítulo é introduzir alguns conceitos de programação não linear e Métodos Subgradientes que serão utilizados no desenvolvimento deste trabalho. A seguir, serão apresentados: Método do Gradiente, os métodos de Newton e Quase-Newton, Método Barzilai e Borwein, Região de Confiança, o método de Levenberg-Marquardt, bem como as condições de KKT, Penalidades, Lagrangeanos Aumentados e métodos Subgradientes. Estes conceitos visam a solução de problemas de minimização irrestrita, problemas de minimização com restrições e a solução de sistemas não lineares.

3.1 BUSCA LINEAR

No problema de minimização irrestrita, ou seja,

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & x \in \mathbb{R}^n, \end{aligned} \tag{3.1}$$

cada iteração do método de busca linear é realizado, calculando a direção de descida $d^k \in \mathbb{R}^n$, a partir de um ponto $x^k \in \mathbb{R}^n$, e o tamanho do passo α_k nessa direção. Tal iteração é dada por

$$x^{k+1} = x^k + \alpha_k d^k.$$

Definição 3.1 *Considere uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$, um ponto $x \in \mathbb{R}^n$ e uma direção $d \in \mathbb{R}^n$. Diz-se que d é uma direção de descida para f , a partir de x , quando existe $\epsilon > 0$ tal que*

$$f(x + \alpha d) < f(x),$$

para todo $\alpha \in (0, \epsilon]$.

O resultado adiante apresenta uma condição suficiente para uma direção ser de descida.

Teorema 3.2 (MARTINEZ; SANTOS, 1998) *Se $\nabla f(x)^T d < 0$, então d é uma direção de descida para f , a partir de x .*

3.2 MÉTODO DE CAUCHY

O Método de Cauchy, também chamado de Método do Gradiente, é um dos mais conhecidos e citados na literatura na resolução do problema (3.1). Neste método, a direção de descida d^k é oposta ao gradiente da função f no ponto x^k . Esta escolha é justificada pelo fato de que esta direção é a que fornece o máximo decréscimo em f .

Para o Método de Cauchy tem-se dois casos, quando a função objetivo não é quadrática e quando a função objetivo é quadrática. A seguir é apresentado o algoritmo para o caso não quadrático.

Algoritmo 3.1 MÉTODO DE CAUCHY PARA O CASO NÃO QUADRÁTICO

Dado: $x^0 \in \mathbb{R}^n$

$k = 0$

Enquanto $\nabla f(x^k) \neq 0$

Defina $d^k = -\nabla f(x^k)$

Obtenha $\alpha_k > 0$ tal que $f(x^k + \alpha_k d^k) < f(x^k)$

Faça $x^{k+1} = x^k + \alpha_k d^k$

$k = k + 1$

Fim.

O subproblema de determinar α_k , minimizador de f na direção de d^k , pode não ter solução e, portanto, o Algoritmo 3.1 nem sempre está bem definido (FRIEDLANDER, 1994). O teorema a seguir não garante a convergência, mas fala da ordem quando a convergência ocorre.

Teorema 3.3 (FRIEDLANDER, 1994) *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^2$. Seja \bar{x} um minimizador local de f , tal que a matriz $\nabla^2 f(\bar{x})$ é definida positiva. Se o Algoritmo 3.1 está bem*

definido $\forall k \in \mathbb{N}$ e a sequência $\{x^k\}$ gerada por ele converge para \bar{x} , então a sequência $\{f(x^k)\}$ converge linearmente para $\{f(\bar{x})\}$ com taxa não superior a $\left[\frac{L-l}{L+l}\right]^2$, onde L e l são o maior e o menor autovalor da $\nabla^2 f(\bar{x})$, respectivamente.

Para o caso quadrático, considere:

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c$$

com $A \in \mathbb{R}^{n \times n}$ definida positiva, b e $c \in \mathbb{R}^n$, então existe um único $\bar{x} \in \mathbb{R}^n$, que é minimizador global de f . Neste caso, em que $A > 0$ o passo α_k que fornece o menor valor de f partindo de x^k na direção d^k é:

$$\alpha_k = \frac{-\nabla f(x^k)^T d^k}{(d^k)^T A d^k}. \quad (3.2)$$

O algoritmo para o caso quadrático é o mesmo que o Algoritmo 3.1 para o caso não quadrático, a alteração está no cálculo do tamanho do passo dado por (3.2).

3.3 MÉTODO DE NEWTON

Nesta seção é introduzido inicialmente o Método de Newton para zeros de função. Em seguida, amplia-se uma extensão dos conceitos do método para o problema de minimização irrestrita.

3.3.1 Método de Newton para Zeros de Função

O Método de Newton para solução de sistema de equações é baseado no polinômio de Taylor de primeira ordem. O problema consiste em encontrar $x \in \mathbb{R}^n$, tal que $F(x) = 0$ com $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F \in C^1$. Considere:

$$F(x) = \begin{pmatrix} F_1(x) \\ F_2(x) \\ \vdots \\ F_n(x) \end{pmatrix}.$$

Então,

$$F'(x) = \nabla F(x) = \begin{pmatrix} \nabla F_1(x)^T \\ \nabla F_2(x)^T \\ \vdots \\ \nabla F_n(x)^T \end{pmatrix} = J(x).$$

A aproximação de Taylor de primeira ordem de $F(x)$ na vizinhança de x^k , é dada por:

$$F(x) \cong F(x^k) + \nabla F(x^k)(x - x^k).$$

O algoritmo de Newton, para resolução de sistemas de equações utiliza, a cada iteração, a aproximação de Taylor de primeira ordem de F numa vizinhança de x^k , de modo que, o próximo iterando, x^{k+1} é obtido resolvendo

$$F(x) \cong F(x^k) + J(x^k)(x - x^k) = 0.$$

Se $J(x^k)$ é não-singular o problema anterior tem uma única solução (MARTINEZ; SANTOS, 1998) e a iteração ou passo de Newton consiste em resolver o seguinte sistema:

$$\begin{cases} d^k = -(J(x^k))^{-1}F(x^k) \\ x^{k+1} = x^k + d^k \end{cases}$$

em que d^k é a direção de Newton.

Convergência: Se $J(x^k)$ é não-singular e o ponto inicial $x^0 \in \mathbb{R}^n$ está suficientemente próximo da solução \bar{x} , o Método de Newton gera uma sequência $\{x^k\}$ que converge para \bar{x} com convergência quadrática (MARTINEZ; SANTOS, 1998).

Algoritmo 3.2 MÉTODO DE NEWTON PARA SISTEMAS DE EQUAÇÕES

Dado: $x^0 \in \mathbb{R}^n$ e $\varepsilon > 0$

$k = 0$

Enquanto $\|F(x^k)\| \geq \varepsilon$ (ou algum outro critério)

Calcular:

$$x^{k+1} = x^k - (J(x^k))^{-1}F(x^k)$$

$$k = k + 1$$

Fin.

O algoritmo do Método de Newton pode variar dependendo do ponto inicial. Quando o método depender, que o ponto inicial esteja suficientemente próximo da solução, é chamado Newton Puro, pois não tem busca e sua convergência é quadrática. Já quando o algoritmo de Newton converge independentemente que x^0 esteja próximo da solução do problema, é chamado Newton Globalizado. Neste caso, é necessário um método de busca unidirecional para obter o comprimento do passo e a convergência será no máximo superlinear. Além disto, é necessário uma função de mérito para decidir o decréscimo da função. Em geral, usa-se a função $f(x) = \frac{1}{2}\|F(x)\|^2$ como função de mérito.

3.3.2 Método de Newton para Minimização Irrestrita

Considere o problema (3.1) de minimização irrestrita e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ duas vezes continuamente diferenciável. Um ponto $\bar{x} \in \mathbb{R}^n$ é minimizador global de um problema de minimização irrestrita se, e somente se, para todo $x \in \mathbb{R}^n$, $f(x) \geq f(\bar{x})$. Por outro lado, $\bar{x} \in \mathbb{R}^n$ é um minimizador local se existir uma vizinhança V de \bar{x} , tal que $f(x) \geq f(\bar{x})$, para todo x em V . O ponto que resolve o problema (3.1) satisfaz a condição $\nabla f(x) = 0$. O Método de Newton, para resolver o problema de minimização irrestrita, é o mesmo utilizado para resolver o sistema de equações $\nabla f(x) = 0$ (NOCEDAL; WRIGHT, 1999). Neste caso, ao aplicar o Método de Newton tem-se:

$$F(x^k) = \nabla f(x^k), \quad J(x^k) = \nabla^2 f(x^k)$$

então

$$\nabla f(x) \cong f(x^k) + \nabla^2 f(x^k)(x - x^k).$$

Denotando $\nabla^2 f(x) = H(x)$, com $H(x)$ não singular, então o passo de Newton puro para o problema de minimização irrestrita é dado por:

$$\begin{cases} d^k = -(H(x^k))^{-1} \nabla f(x^k) \\ x^{k+1} = x^k + d^k \end{cases} \quad (3.3)$$

Algoritmo 3.3 MÉTODO DE NEWTON PURO PARA MINIMIZAÇÃO

Dado: $x^0 \in \mathbb{R}^n$, $\varepsilon > 0$

$k = 0$

Enquanto $\|\nabla f(x^k)\| \geq \varepsilon$

Calcular:

$$x^{k+1} = x^k - (H(x^k))^{-1} \nabla f(x^k)$$

$k = k + 1$

Fim.

Convergência: Se a função f é quadrática e $H(x)$ é invertível, então o Método de Newton resolve o problema (3.1) em uma iteração. Senão, para cada iteração k deverá ser encontrado um ponto estacionário para o modelo quadrático de f em torno de x^k (GONZAGA, 2004). Para o Método de Newton Puro a convergência é quadrática e depende que x^0 pertença à vizinhança de \bar{x} .

Se x^0 não estiver suficientemente próximo de \bar{x} , o Método de Newton pode não obter um minimizador, ou nem mesmo convergir. No caso de $x \in \mathbb{R}^n$ e $H(x)$ ser definida positiva para a iteração k , $H(x^k)$ é sempre não-singular e a direção de Newton $d^k = -(H(x^k))^{-1} \nabla f(x^k)$ é de descida, mas o comprimento dos passos pode ser muito longos.

Para garantir a convergência do Método de Newton independente do ponto inicial, utiliza-se o algoritmo de Newton com busca unidirecional. Se $H(x)$ é contínua e definida positiva para todo $x \in \mathbb{R}^n$, o algoritmo de Newton com busca é globalmente convergente e chamado de ‘Newton Globalizado’. A seguir será apresentado o algoritmo de Newton com busca. Esta busca pode ser Exata ou Inexata para obter o comprimento do passo.

Algoritmo 3.4 MÉTODO DE NEWTON PARA MINIMIZAÇÃO COM BUSCA

Dado: $x^0 \in \mathbb{R}^n$, $\varepsilon > 0$

$k = 0$

Enquanto $\|\nabla f(x^k)\| \geq \varepsilon$

Calcular: $d^k = -(H(x^k))^{-1} \nabla f(x^k)$

Determinar: $\alpha_k > 0$ tal que $f(x^k + \alpha_k d^k) < f(x^k)$

Fazer $x^{k+1} = x^k + \alpha_k d^k$

$k = k + 1$

Fim.

Quando $\alpha_k = 1$, para todo k , o algoritmo passa a ser de Newton Puro.

Convergência: Para o método de Newton com busca, a convergência pode chegar a superlinear. No método com busca há perda na velocidade de convergência. Porém, a vantagem está na garantia de convergência.

Observação: O problema de determinar as raízes de um sistema de equações não lineares $F(x) = 0$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ pode ser transformado em um problema de minimização, ao utilizar o seguinte procedimento:

$$\min_{x \in \mathbb{R}^n} f(x) = \min_{x \in \mathbb{R}^n} \frac{1}{2} \|F(x)\|_2^2. \quad (3.4)$$

3.4 MÉTODOS QUASE-NEWTON

O Método de Newton é ‘bom’, porém ‘caro’, no sentido de necessitar calcular hessiana a cada iteração. Para contornar este problema foram desenvolvidos os Métodos Quase-Newton que fazem uso da aproximação para a hessiana. Neste caso, B_k é uma aproximação da matriz hessiana $H(x^k)$, obtida iterativamente ao utilizar informações geradas pelos passos precedentes do algoritmo (GONZAGA, 2004), e gera uma sequência x^k , que quando converge, tem uma boa taxa de convergência. A iteração do algoritmo é dada por:

$$x^{k+1} = x^k + \alpha_k d^k,$$

com direção de descida $d^k = -B_k \nabla f(x^k)$.

3.4.1 Métodos Quase-Newton para Resolução de Sistema não Linear

O problema de encontrar $x \in \mathbb{R}^n$, tal que $F(x) = 0$ com $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, pelo Método Quase-Newton, consiste em realizar uma aproximação secante de $F(x)$ nos pontos x^k e x^{k+1} conforme MARTINEZ e SANTOS (1998):

$$F(x^{k+1}) - F(x^k) = \nabla F(x^k)(x^{k+1} - x^k), \quad (3.5)$$

então, ao substituir $y^k = F(x^{k+1}) - F(x^k)$ e $s^k = x^{k+1} - x^k$ em (3.5), tem-se

$$\nabla F(x^k) s^k = y^k.$$

A ideia é aproximar $\nabla F(x^k)$ por uma matriz B_k , nos pontos x^k e x^{k+1} . Como consequência obtém-se o sistema:

$$B_k s^k \cong y^k. \quad (3.6)$$

A equação (3.6) é denominada equação secante e pode ser considerada um sistema linear cuja incógnita é a matriz B_k . O sistema tem n^2 variáveis (número de entradas da matriz) e apenas n equações. Deste modo, apenas quando $n = 1$ o sistema poderá ter solução única. Se $n > 1$ e $s^k \neq 0$ haverá uma infinidade de matrizes B_k que satisfazem a equação secante (MARTINEZ; SANTOS, 1998).

A matriz B_k pode ser obtida por meio do método Broyden. A aproximação é feita por uma matriz simétrica de posto um, dada por:

$$B_{k+1} = B_k + \frac{(y^k - B_k s^k)(y^k - B_k s^k)^T}{(y^k - B_k s^k)^T s^k}. \quad (3.7)$$

Dada B_k , se B_{k+1} for determinada por (3.7), então a equação secante é satisfeita. De fato:

$$B_{k+1} s^k = [B_k + \frac{(y^k - B_k s^k)(y^k - B_k s^k)^T}{(y^k - B_k s^k)^T s^k}] s^k = B_k s^k + (y^k - B_k s^k) = y^k. \quad (3.8)$$

Algoritmo 3.5 MÉTODO QUASE-NEWTON PARA SOLUÇÃO DE SISTEMA NÃO LINEAR

Dado: $x^0 \in \mathbb{R}^n$, $\varepsilon > 0$ e uma matriz $n \times n$ simétrica definida positiva,

por exemplo $B_0 = I_{n \times n}$.

$k = 0$

Enquanto $\|\nabla F(x^k)\| \geq \varepsilon$

Calcular: $d^k = -B_k \nabla F(x^k)$, sendo B_k dada por (3.7).

Fazer: $x^{k+1} = x^k + d^k$

$$k = k + 1$$

Fin.

Como no Método de Newton, os Métodos Quase-Newton também podem não convergir. A convergência destes métodos sem busca são locais e dependem do ponto inicial estar próximo da solução do problema. Uma alternativa de globalizar os Métodos Quase-Newton para resolver Sistemas não Lineares é usar a estratégia que foi utilizada no Método de Newton para resolver Sistemas não Lineares. Ver a relação (3.4).

3.4.2 Método Quase-Newton para Minimização Irrestrita

Os métodos secantes para minimização irrestrita aproximam a hessiana da função objetivo, resolvendo o sistema na forma:

$$B_k s^k = y^k \quad (3.9)$$

onde B_k é uma aproximação da hessiana de f , $y^k = \nabla f(x^{k+1}) - \nabla f(x^k)$ e $s^k = x^{k+1} - x^k$.

Observe que a equação (3.9) fornece informação sobre B por meio do ∇f em dois pontos. Esta equação é conhecida como equação secante. Considere B_k a aproximação da hessiana que satisfaz a equação secante. Um dos meios de obter B_k é o Método Broyden apresentado na equação (3.8).

Uma outra forma de se obter B_k é pelo Método Broyden, Fletcher, Goldfarb, Shanno (BFGS), conhecido como a atualização secante mais popular para minimização irrestrita. Ele utiliza atualização de posto 2, e com isso obtém simultaneamente a satisfação da equação secante e matrizes definidas positivas (GONZAGA, 2004). A fórmula de atualização pelo Método BFGS da matriz B_k é dada por:

$$B_{k+1} = B_k + \frac{y^k (y^k)^T}{(y^k)^T s^k} - \frac{B_k s^k (s^k)^T B_k}{(s^k)^T B_k s^k}. \quad (3.10)$$

A direção no Método Quase-Newton para minimização irrestrita é obtida ao substituir a hessiana em (3.3) por B_k , isto é

$$d^k = -B_k \nabla f(x^k).$$

Algoritmo 3.6 MÉTODO QUASE-NEWTON PARA PROBLEMA IRRESTRITO

Dados: $x^0 \in \mathbb{R}^n$, $\varepsilon > 0$ e uma matriz $n \times n$ simétrica definida positiva, por exemplo, $B_0 = I_{n \times n}$.

$k = 0$

Enquanto $\|\nabla f(x^k)\| \geq \varepsilon$

Calcular: $d^k = -B_k \nabla f(x^k)$

Determinar: α_k tal que $f(x^k + \alpha_k d^k) < f(x^k)$

Faça: $x^{k+1} = x^k + \alpha_k d^k$

Obter B_{k+1} , tal que $B_{k+1} s^k = y^k$, por exemplo

B_{k+1} dada em (3.10).

Fin.

3.5 MÉTODO DE GAUSS-NEWTON

O Método de Gauss-Newton é muito utilizado para o problema de quadrados mínimos, no sentido de minimizar a soma dos quadrados dos desvios no ajuste de dados, ou seja, quando

$$f(x) = \frac{1}{2} \|F(x)\|^2 = \frac{1}{2} \sum_{i=1}^n F_i(x)^2$$

sendo $F(x) = (F_1(x), \dots, F_n(x))^T$ e $F_i : \mathbb{R}^n \rightarrow \mathbb{R}$, para todo $i = 1, \dots, n$.

No caso de mínimos quadrados, as funções F_i apresentam estruturas particulares, ou seja, elas representam os resíduos nos dados ajustados. O que justifica o desenvolvimento de um novo método, que não o de Newton ou Quase-Newton, já estudados em seções anteriores.

O Método de Gauss-Newton também é considerado uma modificação do método de Newton para encontrar o mínimo de uma função, propõe uma forma mais simples de encontrar a hessiana. O Método de Gauss-Newton explora a estrutura do gradiente $\nabla f(x)$ e da hessiana $\nabla^2 f(x)$, aqui o termo $\sum_{i=1}^n F_i(x)^T \nabla^2 F_i(x)$ é desprezado (NOCEDAL; WRIGHT, 1999). Ou seja:

$$\nabla^2 f(x) = \sum_{i=1}^n \nabla F_i(x) \nabla F_i(x)^T + \sum_{i=1}^n F_i(x)^T \nabla^2 F_i(x) \cong JF(x)^T JF(x) = B_k$$

e

$$\nabla f(x) = JF(x)^T F(x)$$

logo, obtém d^k ao resolver

$$JF(x)^T JF(x) d^k = -JF(x)^T F(x)$$

ou em notação abreviada $B_k d^k = -\nabla f(x^k)$.

Algoritmo 3.7 MÉTODO DE GAUSS-NEWTON

Dados: $x^0 \in \mathbb{R}^n$, $\varepsilon > 0$

$k = 0$

Enquanto $\|\nabla f(x^k)\| \geq \varepsilon$

Calcular: $B_k d^k = -\nabla f(x^k)$, com

$$B_k = JF(x)^T JF(x)$$

Determinar: $\alpha_k > 0$ tal que $f(x^k + \alpha_k d^k) < f(x^k)$

Faça: $x^{k+1} = x^k + \alpha_k d^k$

$k = k + 1$

Fim.

3.6 MÉTODO BARZILAI E BORWEIN

Neste método, BARZILAI e BORWEIN (1988) propuseram uma forma alternativa para o cálculo do tamanho do passo no Método do Gradiente, empregando dois novos tamanhos de passo para o uso em conjunto com a direção do Método de Cauchy. Utilizaram, para isto, a aproximação da equação secante, como será apresentado nesta seção.

Este método, usualmente chamado de Método do Gradiente de Barzilai e Borwein,

ou simplesmente Método BB, se comparado ao Método do Gradiente, possui menor custo computacional, pois não requer o armazenamento de matrizes e nem buscas lineares a cada iteração. Além disso, a nova escolha do tamanho do passo acelera a taxa de convergência do método, sendo ele, menos sensível ao condicionamento. Barzilai e Borwein provaram que o algoritmo possui convergência R-superlinear para o caso bidimensional quadrático. Com base no Método do Gradiente, considere a seguinte equação:

$$x^{k+1} = x^k - \alpha_k \nabla f(x^k),$$

que pode ser escrita como:

$$x^{k+1} = x^k - B_k \nabla f(x^k)$$

com $B_k = \alpha_k I$ e $I \in \mathbb{R}^{n \times n}$.

De acordo com BARZILAI e BORWEIN (1988), a ideia central do método é obter uma matriz B_k com propriedade Quase-Newton, ou seja, satisfazendo a equação secante (3.6) definida na seção 3.4. Deste modo,

$$B_k \Delta x_{k-1} = \Delta f_{k-1}$$

com $\Delta x_{k-1} = x^k - x^{k-1}$ e $\Delta f_{k-1} = \nabla f(x^k) - \nabla f(x^{k-1})$.

Nestes termos, α_k é calculado tal que:

$$\begin{aligned} \min_{\alpha \in \mathbb{R}} p(\alpha) &= \frac{1}{2} \|\Delta x_{k-1} - \alpha \Delta f_{k-1}\|^2 \end{aligned} \quad (3.11)$$

Para solucionar (3.11) basta fazer $p'(\alpha) = 0$,

$$\begin{aligned} (\Delta x_{k-1} - \alpha \Delta f_{k-1})^T (-\Delta f_{k-1}) &= 0 \\ -(\Delta x_{k-1})^T \Delta f_{k-1} + \alpha_k (\Delta f_{k-1})^T (\Delta f_{k-1}) &= 0 \end{aligned} \quad (3.12)$$

$$\alpha_k = \frac{(\Delta x_{k-1})^T \Delta f_{k-1}}{(\Delta f_{k-1})^T (\Delta f_{k-1})}.$$

A equação (3.12) fornece o tamanho do passo, denominado de passo espectral. Por simetria, podemos minimizar $\frac{1}{2} \|\alpha \Delta x_{k-1} - \Delta f_{k-1}\|^2$ com relação a α . O tamanho do passo espectral correspondente passa a ser:

$$\alpha_k = \frac{(\Delta x_{k-1})^T \Delta x_{k-1}}{(\Delta x_{k-1})^T (\Delta f_{k-1})}.$$

Algoritmo 3.8 MÉTODO GRADIENTE DE BARZILAI E BORWEIN

Dado: $x^0 \in \mathbb{R}^n$

$k = 0$

Enquanto $\nabla f(x^k) \neq 0$

Se $k = 0$, *encontre* α_0 *por meio de busca linear.*

Senão calcular:

$$\alpha_k = \frac{(\Delta x_{k-1})^T \Delta f_{k-1}}{(\Delta f_{k-1})^T (\Delta f_{k-1})} \text{ ou } \alpha_k = \frac{(\Delta x_{k-1})^T \Delta x_{k-1}}{(\Delta x_{k-1})^T (\Delta f_{k-1})}$$

Faça: $x^{k+1} = x^k - \alpha_k \nabla f(x^k)$

$k = k + 1$

Fim.

O método não requer cálculos matriciais e nem busca linear, exceto para $k = 0$.

3.7 REGIÃO DE CONFIANÇA

A busca nos métodos de região de confiança é diferente dos apresentados anteriormente. A ideia da busca linear consiste em fixar uma direção de descida d e determinar o tamanho do passo nesta direção para reduzir a função objetivo. Já no método de Região de Confiança, define-se um modelo da função objetivo e uma região em torno do ponto corrente, na qual o modelo é confiável. Calcula-se um minimizador aproximado do modelo na região de confiança. Caso este ponto forneça uma redução razoável no valor da função objetivo, ele é aceito e repete-se o processo. Caso contrário, pode ser que o modelo não represente adequadamente a função (RIBEIRO; KARAS, 2013). Em outras palavras, no modelo de região de confiança, determina-se, primeiro, o tamanho do passo, e depois a direção. Em cada iteração é resolvido um subproblema quadrático em certa região.

Considere uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^2$ e um ponto $x^k \in \mathbb{R}^n$, o modelo quadrático de f em torno de x^k definido por:

$$q_k(x) = f(x^k) + \nabla f(x^k)^T (x - x^k) + \frac{1}{2} (x - x^k)^T B_k (x - x^k),$$

sendo $B_k \in \mathbb{R}^{n \times n}$ a hessiana de f ou qualquer outra matriz simétrica que satisfaça $\|B_k\| \leq \epsilon$, para alguma constante $\epsilon > 0$, independente de k .

Espera-se que o modelo definido anterior aproxime bem a função f numa vizinhança de x^k . Considere $\Delta_k > 0$ e a região

$$\{x \in \mathbb{R}^n : \|x - x^k\| \leq \Delta_k\},$$

em que o modelo é confiável, considere $d = x - x^k$ e $m_k(d) = q_k(x^k + d)$.

O método de região de confiança está dividido em duas partes. Primeiro resolve o subproblema:

$$\begin{aligned} \min \quad & m_k(d) = f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T B_k d \\ \text{s.a} \quad & \|d\| \leq \Delta_k, \end{aligned}$$

desta forma, se obtém o passo d^k . A segunda parte consiste em avaliar o passo, com objetivo de que o ponto $x^k + d^k$ proporcione uma redução na função objetivo que seja, no mínimo, uma fração da redução do modelo (RIBEIRO; KARAS, 2013). A variação da função e do modelo são dados por:

$$\text{ared} = f(x^k) - f(x^k + d^k), \text{ redução efetiva ou real na função objetivo.}$$

$$\text{pred} = m(0) - m(d^k), \text{ redução prevista pelo modelo.}$$

Um passo é aceito se $\frac{\text{ared}}{\text{pred}} > \delta$ com $\delta \in (0, 1)$.

A seguir é apresentado o algoritmo completo.

Algoritmo 3.9 REGIÃO DE CONFIANÇA

Dados: $x^0 \in \mathbb{R}^n$, $\epsilon > 0$, $\Delta > 0$ (raio inicial), $0 < \delta_1 < \delta_2 < 1$, $\beta \in (0, 1)$.

$k = 1$

Enquanto $\|\nabla f(x^k)\| \geq \epsilon$

Construa o modelo:

$$m_k(d) = f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T B_k d$$

Calcule d , *solução de:*

$$\min \quad m_k(d); \text{ s.a} \|d\| \leq \Delta$$

Calcular: $r = \frac{\text{ared}}{\text{pred}}$, *onde*

$$ared = f(x^k) - f(x^k + d^k) \text{ e}$$

$$pred = m(0) - m(d^k).$$

Se $r < \delta_1$ (rejeitar o passo e reduzir Δ)

$$\text{Faça: } x^{k+1} = x^k$$

$$\text{Calcular: } \Delta = \beta\Delta$$

Senão

se $r > \delta_2$ (aceitar o passo e aumentar Δ)

$$\text{Faça: } x^{k+1} = x^k + d^k$$

$$\text{Calcular: } \Delta = \frac{1}{\beta}\Delta$$

Senão (aceitar o passo e manter Δ)

$$\text{Faça: } x^{k+1} = x^k + d^k$$

Fim

Fim.

3.8 MÉTODO DE LEVENBERG-MARQUARDT

O Método de Levenberg-Marquardt, assim como o Método de Gauss-Newton, é utilizado para minimizar uma função, que é expressa pela soma dos quadrados de diversas funções não lineares, e utiliza uma região de confiança. O uso de uma região de confiança evita um dos pontos fracos de Gauss-Newton, ou seja, o seu comportamento quando o Jacobiano $JF(x^k)$ é de classificação deficiente (NOCEDAL; WRIGHT, 1999). Para encontrar a direção d^k , no passo k , é necessário resolver a cada iteração o subproblema

$$\begin{aligned} \min_d \quad & \frac{1}{2} \|JF(x^k)d + F(x^k)\|^2 \\ \text{s.a} \quad & \|d\| \leq \Delta_k, \end{aligned} \tag{3.13}$$

com $\Delta_k \in \mathbb{R}_{++}$ o raio da região de confiança e $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

O teorema abaixo fornece condições para que o vetor d^k seja a solução do subproblema (3.13).

Teorema 3.4 *O vetor d^k é uma solução do subproblema (3.13) de região de confiança, para algum $\Delta > 0$, se, e somente se, existe um escalar $\lambda \geq 0$ tal que*

$$[JF(x^k)^T JF(x^k) + \lambda^k I]d^k = -JF(x^k)^T F(x^k), \quad (3.14)$$

$$\lambda(\Delta - \|d\|) = 0.$$

O Método de Levenberg-Marquardt é uma forma de garantir um decréscimo no valor da função objetivo. Propõe a soma de uma parcela λI à matriz hessiana aproximada pelo Método de Gauss-Newton $B_k = JF(x)^T JF(x)$, de modo que $(B_k + \lambda^k I)$ seja definida positiva, em que $\lambda \geq 0$ é um escalar, denominado parâmetro de Levenberg-Marquardt, I é a matriz identidade de ordem n e d^k uma direção de descida. Portanto, a equação (3.14) pode ser escrita da seguinte forma:

$$(B_k + \lambda^k I)d^k = -JF(x^k)^T F(x^k).$$

Quando $\lambda = 0$, o Método de Levenberg-Marquardt se reduz ao Método de Gauss-Newton. Se $\lambda > 0$ tal que a matriz $(B_k + \lambda^k I)$ é definida positiva, conforme MEDEIROS (2008) Δ_k pode ser definido como:

$$\Delta_k = \| -(B_k + \lambda^k I)^{-1} JF(x^k)^T F(x^k) \|. \quad (3.15)$$

No Método Levenberg-Marquardt o raio Δ_k é dado implicitamente em função de λ^k pela equação (3.15).

No algoritmo a seguir, considere $f(x) = \frac{1}{2}\|F(x)\|^2$.

Algoritmo 3.10 *MÉTODO DE LEVENBERG-MARQUARDT*

Dados: $x^0 \in \mathbb{R}^n$, $\varepsilon > 0$, $\Delta_{min} > 0$, $\Delta_{max} > \Delta_{min}$ e $\Delta_0 \in [\Delta_{min}, \Delta_{max}]$

$k = 0$

Enquanto $\|\nabla f(x^k)\| \geq \varepsilon$

Calcular: $d^k = \operatorname{argmin}_d \frac{1}{2}\|JF(x^k)d + F(x^k)\|^2$ *s.a* $\|d\| \leq \Delta_k$

Escolha Δ_{k+1} *conveniente, tal que* $\Delta_{min} < \Delta_{k+1} < \Delta_{max}$

Se $f(x^k + d^k) < f(x^k)$

Faça: $x^{k+1} = x^k + d^k$

Senão

Faça: $x^{k+1} = x^k$

Fim

$k = k + 1$

Até a convergência.

O algoritmo para o subproblema (3.13) poderá ser resolvido por vários métodos, ver NOCEDAL e WRIGHT (1999).

3.9 CONDIÇÕES DE KKT

As condições de Karush, Kuhn e Tucker (KKT) são condições de otimalidade para problemas com restrições, a ser definido na sequência, que sob certas hipóteses, devem ser satisfeitas pelos minimizadores do problema (SOBRAL, 2008). Essas condições são uma generalização do Método dos Multiplicadores de Lagrange. Considere inicialmente o problema de minimização dado por:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) \leq 0 \quad e \quad h(x) = 0, \end{aligned} \tag{3.16}$$

com $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, funções continuamente diferenciáveis. O conjunto viável do problema (3.16) é definido por:

$$\Omega = \{x \in \mathbb{R}^n : g(x) \leq 0 \quad e \quad h(x) = 0\}.$$

Assim, um ponto x é um ponto viável de (3.16) se pertencer ao conjunto Ω . A seguir será definido o sistema KKT e sua formalização matemática.

Definição 3.5 *O sistema KKT associado ao problema (3.16), em um ponto $(\bar{x}, \bar{\lambda}, \bar{\mu})$ é*

- 1) $\nabla L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$
- 2) $\bar{\mu}_i g_i(\bar{x}) = 0, \quad i = 1, \dots, p$
- 3) $g(\bar{x}) \leq 0$
- 4) $h(\bar{x}) = 0$
- 5) $\bar{\mu} \geq 0$
- 6) $\bar{\lambda} \in \mathbb{R}^n$.

Na equação $\nabla L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$, ∇L é o gradiente, em x , da função Lagrangeana definida como:

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j g_j(x).$$

Para o problema de minimização, só com restrições de desigualdade,

$$\begin{aligned} \min \quad & f(x); \\ \text{s.a} \quad & g(x) \leq 0 \end{aligned} \tag{3.17}$$

o sistema KKT em um ponto $(\bar{x}, \bar{\mu})$ é dado por:

- 1) $\nabla L(\bar{x}, \bar{\mu}) = 0$
 - 2) $\bar{\mu}_i g_i(\bar{x}) = 0, \quad i = 1, \dots, p$
 - 3) $g(\bar{x}) \leq 0$
 - 4) $\bar{\mu} \geq 0$.
- (3.18)

Para a existência das condições de KKT ao qual o problema está associado, no nosso caso o problema (3.16), é preciso de mais uma hipótese, ou seja, que o problema satisfaça uma condição chamada de qualificação. Existem várias condições de qualificação na literatura, por exemplo: regularidade (também conhecida como independência linear dos gradientes), Slater, Mangasarian-Fromovitz, quase-regularidade, dentre outras. No decorrer deste trabalho serão necessárias apenas as condições de qualificação de Slater e Mangasarian-Fromovitz que estão definidas a seguir.

Condição de Qualificação de Slater e de Mangasarian-Fromovitz (para o problema (3.16)), de acordo com EUSTÁQUIO (2007).

Definição 3.6 *A condição de qualificação de Slater é verificada se h é afim e g é convexa*

e existe $\tilde{x} \in \Omega$ tal que $g(\tilde{x}) < 0$ e $h(\tilde{x}) = 0$.

Definição 3.7 A condição de qualificação de Mangasarian-Fromovitz (MFCQ) é satisfeita em \bar{x} quando os gradientes das restrições de igualdade são linearmente independentes e existir um vetor $d \in \mathbb{R}^n$ tal que $\nabla h(\bar{x})^T d = 0$ e $\nabla g_j(\bar{x})^T d < 0$, $j = 1, \dots, p$.

Teorema 3.8 Condições de Otimalidade de KKT

Se um ponto $\bar{x} \in \Omega$ é uma solução do problema (3.16) e satisfaz uma condição de qualificação, então existem $\bar{\lambda} \in \mathbb{R}^m$ e $\bar{\mu} \in \mathbb{R}^p$, tais que

$$\begin{aligned} -\nabla f(\bar{x}) &= \sum_{j=1}^m \bar{\lambda}_j \nabla h_j(\bar{x}) + \sum_{j=1}^p \bar{\mu}_j \nabla g_j(\bar{x}) \\ \bar{\mu}_j &\geq 0, \quad j = 1, \dots, p \\ \bar{\mu}_j g_j(\bar{x}) &= 0, \quad j = 1, \dots, p \\ g_j(\bar{x}) &\leq 0, \quad j = 1, \dots, p \\ h_j(\bar{x}) &= 0, \quad j = 1, \dots, m. \end{aligned}$$

3.10 MÉTODOS DE PENALIDADES

Esses métodos são utilizados para resolver problemas de minimização com restrições. Essas restrições podem ser de igualdade e/ou desigualdade.

Métodos de Penalidades são baseados em penalizações e métodos iterativos, em que a cada iteração o problema restrito (3.16) é transformado em um problema irrestrito e pode ser resolvido pelos métodos estudados anteriormente.

Assim, tem-se uma nova função objetivo para cada iteração, ou seja, a cada iteração busca-se a resolução de um novo subproblema irrestrito, dado por:

$$P_k = f(x) + rP_1(g(x)) + \rho P_2(h(x)) = P(x, r, \rho).$$

A solução \bar{x} do problema original estará relacionada com a solução da sequência de subproblemas irrestritos. As funções de penalidade podem ser divididas em dois tipos:

- Funções de Penalidade Barreira ou Penalidade Interna (pontos interiores);
- Funções de Penalidade Externa (pontos exteriores).

A seguir serão apresentados estes métodos de penalidades.

3.10.1 Penalidades Internas (ou Métodos de Barreiras)

Na Penalidade Interna, a função objetivo é modificada ao adicionar um termo funcional que tende a infinito, quando o ponto se aproxima da fronteira do conjunto viável, forma assim uma espécie de barreira intransponível (MARTINEZ; SANTOS, 1998). Há dois exemplos de Penalidade Interna bem conhecidos na literatura e são mais apropriadas para problemas com restrições de desigualdade, no formato de (3.17).

1ª - Função Barreira Logarítmica:

$$x \in \mathbb{R}^n \mapsto B(x) = - \sum_{i=1}^p \log(-g_i(x)) \quad (3.19)$$

2ª Função Barreira Inversa:

$$x \in \mathbb{R}^n \mapsto B(x) = - \sum_{i=1}^p \frac{1}{g_i(x)}. \quad (3.20)$$

A Função Penalidade é dada por:

$$P(x, r) = f(x) + rB(x),$$

em que $B(x)$ é dada por (3.19) ou (3.20), sendo $r > 0$ o parâmetro de penalidade, cuja função é aumentar a penalidade. Nesse caso, o valor de r irá diminuir em todas as iterações.

Algoritmo 3.11 MÉTODO DE BARREIRAS

Dado: $x^0 \in \mathbb{R}^n$, $r^0 > 0$

$k = 0$

Repetir:

$$x^{k+1} \in \operatorname{argmin}\{f(x) + r^k B(x) : x \in \mathbb{R}^n\}$$

$$r^{k+1} = \frac{1}{2} r^k$$

$$k = k + 1$$

Até convergir (critério de parada).

No Algoritmo 3.11, $B(x)$ é a Penalidade Barreira Logarítmica ou a Penalidade

Barreira Inversa. Um critério bastante utilizado para a parada dos algoritmos, para detectar sua convergência (ou uma solução para o problema) é a norma do gradiente da função aproximadamente igual a zero, quando esta é possível de ser calculada (SANT'ANA, 2005). Para o Algoritmo 3.11, o critério de parada é:

$$\|\nabla_x P(x^{k+1}, r^k)\| \leq \varepsilon,$$

em que ε é um valor pré-fixado e $\nabla_x P(x^{k+1}, r^k)$ é o gradiente da função penalizada na variável $x \in \mathbb{R}^n$, no ponto x^{k+1} e $r^k > 0$.

3.10.2 Penalidades Externas

Os Métodos de Penalização Externa podem ser aplicados ao problema no formato geral (3.16).

A solução de um problema penalizado externamente está, geralmente, fora do conjunto viável, mas se aproxima dele quando o termo de penalidade é muito grande (MARTINEZ, 2009).

Para problemas com restrições de igualdade, a penalidade mais utilizada na literatura é a quadrática, apresentada a seguir:

$$x \in \mathbb{R}^n \mapsto p(x) = \frac{1}{2} \|h(x)\|^2 = \frac{1}{2} h(x)^T h(x) = \sum_{i=1}^m (h_i(x))^2$$

e a função Penalidade Externa nesse caso é da forma:

$$P(x, \rho) = f(x) + \rho \sum_{i=1}^m (h_i(x))^2,$$

com $\rho > 0$ fixado. Neste caso, ρ irá aumentar em todas as iterações.

Já, para problemas com restrições de desigualdade uma penalidade utilizada é:

$$x \in \mathbb{R}^n \mapsto p(x) = \sum_{i=1}^p \max\{0, g_i(x)\}$$

esta penalidade é não diferenciável em todo o domínio. Portanto, a solução é considerar:

$$x \in \mathbb{R}^n \mapsto p(x) = \sum_{i=1}^p (\max\{0, g_i(x)\})^2$$

e, deste modo, a função Penalidade Externa é dada por:

$$P(x, r) = f(x) + r \sum_{i=1}^p (\max\{0, g_i(x)\})^2.$$

Algoritmo 3.12 *PENALIDADE EXTERNA*

Dado: $x^0 \in \mathbb{R}^n$, $\rho^0 > 0$

$k = 0$

Repetir:

$$x^{k+1} \in \operatorname{argmin}\{f(x) + \rho^k p(x) : x \in \mathbb{R}^n\}$$

$$\rho^{k+1} > \rho^k \text{ (por exemplo: } \rho^{k+1} = 2\rho^k \text{)}$$

$$k = k + 1$$

Até convergir (critério de parada).

3.11 MÉTODOS DE LAGRANGEANOS AUMENTADOS

Os métodos clássicos de penalidades têm como característica gerar subproblemas mal condicionados. Já os métodos de Lagrangeano Aumentado tentam conciliar o problema de mal condicionamento e de certa forma tentam gerar subproblemas mais fáceis de resolver. São métodos que buscam resolver o problema por meio de um processo iterativo de maneira que, em cada iteração, resolve-se um problema auxiliar irrestrito formado pela função objetivo e um múltiplo das restrições do problema original (MARTINEZ; SANTOS, 1998).

Os métodos de Lagrangeano Aumentado na sua forma original foram introduzidos por POWELL (1969) e HESTENES (1969), para resolver problemas com restrições de igualdade e posteriormente generalizados para problemas com restrições de desigualdade.

3.11.1 Problema de Minimização com Restrições de Igualdade

Considere o problema com restrição de igualdade,

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & h(x) = 0 \end{aligned} \tag{3.21}$$

com $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f, h \in C^1$. A função lagrangeana para (3.21) é definida por:

$$(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m \mapsto l(x, \lambda) = f(x) + \lambda^T h(x) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) \quad (3.22)$$

em que λ é o vetor de multiplicadores de lagrange. Pelas condições necessárias de primeira ordem se \bar{x} é solução de um problema com restrição de igualdade, então existem um vetor de multiplicadores de Lagrange $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3, \dots, \bar{\lambda}_m)^T$ tal que $\nabla l(\bar{x}, \bar{\lambda}) = 0$, sendo l a função Lagrangeana (3.22).

Desta forma, os candidatos a pontos estacionários para o problema com restrição de igualdade devem satisfazer o seguinte sistema de $m+n$ equações não-lineares por $m+n$ variáveis

$$\begin{cases} \nabla f(x) + \lambda^T \nabla h(x) = 0 \\ h_i(x) = 0, \quad i = 1, 2, \dots, m. \end{cases}$$

Definição 3.9 *Diz-se que dois problemas A e B são equivalentes quando toda solução de A é também uma solução de B e vice-versa.*

Pela Definição 3.9 verifica-se que o problema (3.21) é equivalente ao seguinte problema:

$$\begin{aligned} \min \{ f(x) + \lambda^T h(x) \} \\ \text{s.a } h(x) = 0. \end{aligned} \quad (3.23)$$

Com a penalidade quadrática apresentada na seção anterior, tem-se a seguinte função penalizada para o problema (3.23):

$$(x, \lambda, \rho) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} \mapsto L(x, \lambda, \rho) = f(x) + \lambda^T h(x) + \frac{\rho}{2} h(x)^T h(x). \quad (3.24)$$

Neste caso, $\rho > 0$ é o parâmetro de penalidade e λ é o vetor de multiplicadores de Lagrange.

Algoritmo 3.13 *LAGRANGEANO AUMENTADO CLÁSSICO PARA PROBLEMA DE MINIMIZAÇÃO COM RESTRIÇÃO DE IGUALDADE*

Dados: $x^0 \in \mathbb{R}^n, \rho^0 > 0$ e $\lambda^0 \in \mathbb{R}^m$

$k = 0$

Repetir:

$$x^{k+1} \in \operatorname{argmin}\{L(x, \lambda^k, \rho^k) : x \in \mathbb{R}^n\}$$

$$\lambda^{k+1} = \lambda^k + \rho^k h(x^{k+1})$$

$$\text{Se } \|h(x^{k+1})\| > 0.1 \|h(x^k)\|$$

$$\rho^{k+1} > \rho^k \quad (\rho^{k+1} = 2\rho^k)$$

$$\text{Senão } \rho^{k+1} = \rho^k$$

Fim

$$k = k + 1$$

Fim.

onde L é a função Lagrangeano Aumentado (3.24) e ρ o parâmetro de penalidade. O critério de parada mencionado no Algoritmo 3.13 pode ser, por exemplo, $\|\nabla l(x^k, \lambda^k)\| \leq \varepsilon$ e $\|h(x^k)\| \leq \varepsilon$, neste caso $\varepsilon > 0$ é uma precisão pré-fixada e l é a função Lagrangeana (3.22).

3.11.2 Problemas de Minimização com Restrições de Desigualdade

Considere o problema de minimização (3.17), com $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ funções contínuas e diferenciáveis ao menos até primeira ordem.

A função lagrangena para o problema de minimização com Restrição de Desigualdade é dada por:

$$(x, \mu) \in \mathbb{R}^n \times \mathbb{R}^p \mapsto l(x, \mu) = f(x) + \sum_{i=1}^p \mu_i g_i(x)$$

sendo μ o multiplicador de Lagrange.

A generalização do Método de Lagrangeano Aumentado para o problema com restrição de desigualdade foi, primeiramente, proposta por (ROCKAFELLAR, 1976a). O qual sugeriu uma metodologia de acrescentar folga à restrição de desigualdade e então resolver o problema com restrição de igualdade com variáveis de folga não negativas.

A seguir será apresentada em mais detalhes esta metodologia.

Primeiramente, acrescenta-se as variáveis de folga transformando o problema (3.17) na forma a seguir:

$$\begin{aligned} \min \quad & f(x); \\ \text{s.a} \quad & g(x) + v = 0 \\ & v \geq 0, \quad x \in \mathbb{R}^n. \end{aligned} \quad (3.25)$$

Agora, utiliza-se a penalidade quadrática, proposta por HESTENES (1969) e POWELL (1969), nas restrições de igualdade para obter a função Lagrangeano Aumentada associada a (3.25):

$$(x, \lambda, r) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} \mapsto L(x, \lambda, r) = f(x) + \sum_{i=1}^m \lambda_i (g_i(x) + v_i) + r \sum_{i=1}^m (g_i(x) + v_i)^2. \quad (3.26)$$

Assim, o problema a ser resolvido no Método de Lagrangeano Aumentado é dado por:

$$\min \left\{ f(x) + \sum_{i=1}^m \lambda_i (g_i(x) + v_i) + r \sum_{i=1}^m (g_i(x) + v_i)^2 : x \in \mathbb{R}^n, v \in \mathbb{R}_{++}^m \right\}. \quad (3.27)$$

Note que o problema (3.27) para $v \geq 0$, dado x , pode ser resolvido por uma expressão fechada para cada v_i separadamente. Para isto, será desenvolvido os termos em v na expressão (3.27), ou seja:

$$\min \left\{ f(x) + \sum_{i=1}^m (\lambda_i g_i(x) + r g_i^2(x) + r v_i^2 + [\lambda_i + 2r g_i(x)] v_i) : x \in \mathbb{R}^n, v \in \mathbb{R}_{++}^m \right\}. \quad (3.28)$$

Os termos envolvendo v em (3.28) formam uma quadrática dada por $r v_i^2 + (\lambda_i + 2r g_i(x)) v_i$. Essa quadrática possui um mínimo para $v_i \geq 0$ e $r > 0$ dado por

$$\begin{cases} v_i = -g_i(x) - \frac{\lambda_i}{2r} & \text{se } g_i(x) \leq \frac{\lambda_i}{2r}; \quad i = 1, 2, \dots, m \\ v_i = 0 & \text{se } g_i(x) > -\frac{\lambda_i}{2r}. \end{cases}$$

Assim a função Lagrangeano Aumentado associada ao problema (3.25), dada em (3.26) terá a forma:

$$(x, \lambda, r) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} \mapsto L(x, \lambda, r) = f(x) + \sum_{i=1}^m p(g_i(x), \lambda_i, r), \quad (3.29)$$

em que,

$$p(g_i(x), \lambda_i, r) = \begin{cases} -\frac{\lambda_i}{4r} & \text{se } g_i(x) < -\frac{\lambda_i}{2r} \\ \lambda_i g_i(x) + r[g_i(x)]^2 & \text{se } g_i(x) \geq -\frac{\lambda_i}{2r}, \end{cases}$$

ou então,

$$p(g_i(x), \lambda_i, r) = \frac{1}{4r} [(max\{0, \lambda_i + 2r g_i(x)\})^2 - \lambda_i^2]. \quad (3.30)$$

Ao substituir (3.30) em (3.29) obtém-se a função Lagrangeano Aumentado utilizada por ROCKAFELLAR (1976a), dada por:

$$(x, \lambda, r) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} \mapsto L(x, \lambda, r) = f(x) + \frac{1}{4r} \sum_{i=1}^m [(max\{0, \lambda_i + 2r g_i(x)\})^2 - \lambda_i^2], \quad (3.31)$$

sendo $r > 0$ o parâmetro de penalidade.

Desta forma, o Método Lagrangeano Aumentado, aplicado ao problema (3.25), proposto por ROCKAFELLAR (1976a) gera duas seqüências da forma:

Dados $\lambda^0 \in \mathbb{R}^m$ e $r^0 > 0$,

$$x^{k+1} = argmin\{L(x, \lambda^k, r^k) : x \in \mathbb{R}^n\}$$

$$\lambda^{k+1} = max\{0, \lambda^k + 2r^k g_i(x^{k+1})\}, \quad i = 1, 2, \dots, m$$

em que $L(x, \lambda^k, r^k)$ é dada por (3.31).

O parâmetro de penalidade pode ser mantido constante. Caso não seja mantido constante, a seqüência $\{r^k\}$ deve ser não crescente.

A seguir, define-se uma família de funções de penalidades que são coercivas à direita. Elas são usadas para apresentar o algoritmo de Lagrangeano Aumentado (JUSSIANI, 2004).

Definição 3.10 *Define-se uma família P de funções de penalidade*

$$y \in \mathbb{R}, \mu \in \mathbb{R}_{++} \mapsto p(y, \mu) \in \mathbb{R}$$

com domínio de $p(\cdot, \mu) = (-\infty, b)$, $b > 0$, possivelmente $b = +\infty$. Com p satisfazendo as seguintes propriedades, para todo $\mu \in \mathbb{R}_{++}$:

$$(P1) \quad p(0, \mu) = 0;$$

$$(P2) \quad p'(0, \mu) = \mu;$$

$$(P3) \quad p(\cdot, \mu) \text{ é estritamente convexa e diferenciável em } (-\infty, b);$$

$$(P4) \quad \lim_{y \rightarrow b} p'(y, \mu) = +\infty;$$

$$(P5) \quad \lim_{y \rightarrow -\infty} p'(y, \mu) = 0;$$

$p'(y, \mu)$ representa a derivada de p em relação à primeira variável. As características das funções de penalização da família P são:

- Todas passam pela origem e possuem derivada igual a μ na origem;
- São estritamente convexas e diferenciáveis;
- São coercivas à direita.

A figura 1, exibe o comportamento das funções penalidade da família P .

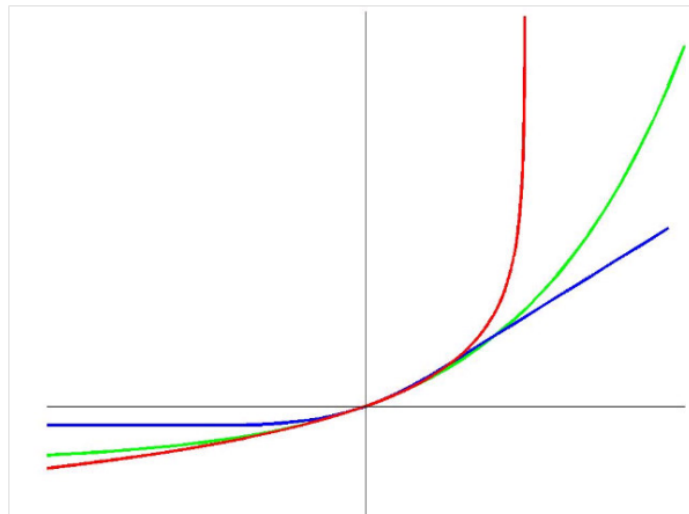


FIGURA 1: FAMÍLIA DE FUNÇÕES PENALIDADES
 FONTE: SANT'ANA (2005, p. 52)

A função Lagrangeano Aumentado para o problema (3.17) pode ser escrita da seguinte forma:

$$(x, \mu, r) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}_{++} \mapsto L(x, \mu, r) = f(x) + r \sum_{i=1}^p p\left(\frac{g_i(x)}{r}, \mu_i\right). \quad (3.32)$$

Sendo $r > 0$ o parâmetro de penalização e p uma função de penalidade da família P . No algoritmo de Lagrangeano Aumentado com restrição de desigualdade, a cada iteração, dados $r^k > 0$, $\mu^k \in \mathbb{R}_+^p$, então x^{k+1} é determinado por:

$$x^{k+1} \in \operatorname{argmin} \left\{ f(x) + r^k \sum_{i=1}^p p\left(\frac{g_i(x)}{r^k}, \mu_i^k\right) : x \in \mathbb{R}^n \right\}.$$

Não é conhecido μ^k antecipadamente. Por isso, será fornecido, na primeira iteração, um valor arbitrário para o vetor μ (normalmente vetor nulo ou vetor de uns). Então, esses valores são atualizados de modo a satisfazer a primeira e a última relação de KKT dadas em (3.18) (JUSSIANI, 2004). Ao derivar, em relação à variável x , a função Lagrangeano Aumentado, tem-se:

$$\nabla_x L(x, \mu, r) = \nabla f(x) + \sum_{i=1}^p p'\left(\frac{g_i(x)}{r^k}, \mu_i^k\right) \nabla g_i(x),$$

sendo μ atualizado por:

$$\mu_i^{k+1} = p'\left(\frac{g_i(x)}{r^k}, \mu_i^k\right), \quad \text{para } i = 1, 2, \dots, p.$$

A seguir é apresentado o algoritmo de Lagrangeano Aumentado, com penalidade da família P .

Algoritmo 3.14 *LAGRANGEANO AUMENTADO PARA PROBLEMA DE MINIMIZAÇÃO COM RESTRIÇÃO DE DESIGUALDADE E PENALIDADE DA FAMÍLIA P*

Dado: $x^0 \in \mathbb{R}^n, \mu \in \mathbb{R}_+^p, r^0 > 0$

$k = 0$

Repetir:

$$x^{k+1} \in \operatorname{argmin} \{L(x, \lambda^k, \rho^k) : x \in \mathbb{R}^n\}$$

$$\text{Atualizar: } \mu_i^{k+1} = p' \left(\frac{g_i(x)}{r^k}, \mu_i^k \right); \quad i = 1, 2, \dots, p$$

$$k = k + 1$$

Até convergir (critério de parada).

Fim.

Em que L é a função Lagrangeano Aumentado (3.32) e r^k o parâmetro de penalidade. Neste caso r^k pode ser mantido constante e igual a 1.

3.12 SUBGRADIENTE

Nesta seção são apresentados os Métodos de Subgradiente, com objetivo de utilizá-los para resolução do NEP. Os Métodos Subgradientes foram originalmente desenvolvidos por Shor nas décadas de 60 e 70. Uma função convexa não é necessariamente diferenciável. Para minimizar uma função não diferenciável, N.Z. Shor sugere um método análogo ao do gradiente, substituindo o gradiente por um subgradiente arbitrário da função num certo ponto (POLYAK, 1987).

A seguir é apresentada a definição de Subgradiente, os teoremas e lemas relacionados à sua definição, para melhor compreensão dos Métodos Subgradientes, baseado em POLYAK (1987).

Definição 3.11 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa e um vetor $s \in \mathbb{R}^n$, para o qual*

$$f(y) \geq f(x) + s^T(y - x), \quad (3.33)$$

para todo $y \in \mathbb{R}^n$, s é chamado de Subgradiente da função f no ponto x .

O subgradiente não é conhecido na sua totalidade, $\partial f(x)$ é utilizado para representar o conjunto dos subgradientes também denominado subdiferencial.

Teorema 3.12 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa, o conjunto dos subgradientes em qualquer ponto é não vazio, convexo e compacto.*

Demonstração: Ver POLYAK (1987).

Uma condição necessária e suficiente para que um ponto \bar{x} seja um minimizador global da função convexa $f(x)$, é dada pelo teorema a seguir.

Teorema 3.13 (Condição necessária e suficiente para um mínimo global de uma função convexa não-diferenciável) Um ponto \bar{x} é minimizador global de uma função convexa f se, e somente se, satisfaz

$$0 \in \partial f(\bar{x}),$$

em que 0 é o vetor nulo em \mathbb{R}^n .

Demonstração:

Condição necessária - Seja $\bar{x} \in \mathbb{R}^n$ o minimizador global de f . Então, pela definição de mínimo global, tem-se $f(y) \geq f(\bar{x}) + 0^T(y - \bar{x}) \quad \forall y \in \mathbb{R}^n$, isto significa que 0 é o subgradiente de f em \bar{x} .

Condição suficiente - Se 0 é subgradiente em \bar{x} , então por (3.33) $f(y) \geq f(\bar{x}) + 0^T(y - \bar{x}) = f(\bar{x}) \quad \forall y \in \mathbb{R}^n$. ■

Observações importantes:

1. Se f é diferenciável em x , então seu gradiente, $\nabla f(x)$, é o único subgradiente de f em x ;
2. Uma função convexa sempre possui um subgradiente.

Exemplo 3.14 Calcular o subgradiente da função $f(x) = |x|$.

Se $x > 0$ então a derivada de $f(x)$ em relação a x é $\partial f(x) = 1$.

Se $x < 0$ então a derivada de $f(x)$ em relação a x é $\partial f(x) = -1$.

Por outro lado, para $x = 0$, qualquer reta passando pela origem, com coeficiente angular no intervalo $[-1, 1]$, desempenha o papel de reta tangente ao gráfico de f no ponto de abscissa $x = 0$ e, conseqüentemente, a derivada da função módulo neste ponto não existe no sentido usual. Logo, existe uma família de retas, com coeficiente angular no intervalo $[-1, 1]$, que fazem o papel de retas tangentes ao gráfico de f no ponto de abscissa $x = 0$. Considere duas destas retas dadas por $r_1(x) = 0,5x$ e $r_2(x) = 0,2x$, com coeficientes angulares iguais a $0,5$ e $0,2$, respectivamente, mostradas na figura a seguir.

Ao utilizar a definição dada por (3.33) em $x = 0$ no exemplo acima, tem-se:

$$|y| \geq 0 + s^T(y - 0), \tag{3.34}$$

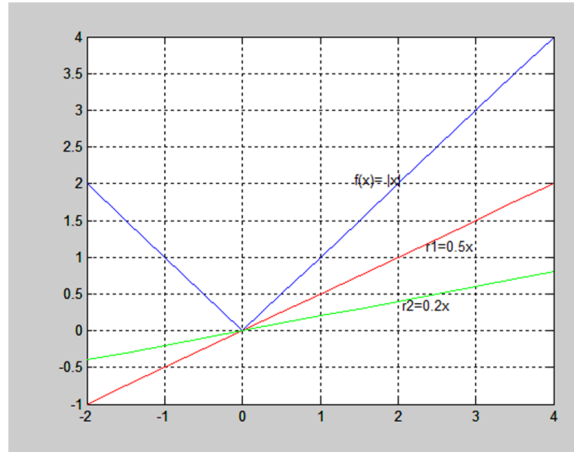


FIGURA 2: SUBGRADIENTES

$\forall y \in \mathbb{R}^n$.

Se $y \geq 0$, então $|y| = y$, logo a desigualdade (3.34) pode ser escrita por $y \geq sy$, podendo ser reescrita na forma $(1 - s)y \geq 0$. Como $y \geq 0$, para que esta desigualdade seja satisfeita é necessário que $(1 - s) \geq 0$, o que implica em $s \leq 1$, ou seja $s \in (-\infty, 1]$. Por outro lado: se $y < 0$, então $|y| = -y$ e a desigualdade em (3.34) pode ser escrita por $y \leq -sy$, podendo ser reescrita na forma $(s + 1)y \leq 0$. Como $y < 0$ para que esta desigualdade seja satisfeita é necessário que $(s + 1) \geq 0$, o que implica em $s \geq -1$, ou seja $s \in [-1, \infty)$. Logo, o conjunto dos subgradientes de f em $x = 0$ é dado pela intersecção dos conjuntos $s \in (-\infty, 1]$ e $s \in [-1, \infty)$, ou seja $\partial f(0) = [-1, 1]$.

Portanto $0 \in \partial f(0) = [-1, 1]$, então, pelo teorema 3.13, a origem é um ponto de mínimo global da função módulo.

Quando o subgradiente é conhecido, pode-se calcular a derivada direcional, a qual é um valor escalar que representa a derivada de um campo escalar ao longo de um vetor.

Definição 3.15 Uma função convexa f em um ponto arbitrário x possui uma derivada direcional em qualquer direção. Considere d um vetor unitário em \mathbb{R}^n , a derivada direcional de f no ponto x e na direção d é definida por:

$$f'(x; d) = \lim_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t}$$

se o limite existir.

Na definição de derivada direcional o vetor d deve ser unitário. Caso contrário a derivada direcional não dependeria somente do ponto e da direção, mas também do comprimento do vetor.

Lema 3.16 Para quaisquer x, d

$$f'(x; d) = \max_{s \in \partial f(x)} (s, d).$$

Demonstração: Ver (POLYAK, 1987).

Lema 3.17 Considere $f_1(x)$ e $f_2(x)$ funções convexas, $f(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$ e $\alpha_1, \alpha_2 \geq 0$, então

$$\partial f(x) = \alpha_1 \partial f_1(x) + \alpha_2 \partial f_2(x).$$

Demonstração: Ver (POLYAK, 1987).

O lema a seguir apresenta uma regra para calcular o subgradiente do máximo de várias funções.

Lema 3.18 Seja

$$f(x) = \max_{1 \leq i \leq m} f_i(x),$$

onde as funções $f_i(x)$ são convexas. Então

$$\partial f(x) = \text{Conv} \bigcup_{i \in I(x)} \partial f_i(x), \quad I(x) = \{i : f_i(x) = f(x)\}.$$

Demonstração: Ver POLYAK (1987).

Lema 3.19 Seja A uma matriz $m \times n$, $\phi(y)$ uma função convexa no \mathbb{R}^m e $f(x) = \phi(Ax)$, $x \in \mathbb{R}^n$. Então

$$\partial f(x) = A^T \partial \phi(Ax).$$

Demonstração: Ver POLYAK (1987).

3.13 MÉTODOS SUBGRADIENTES

Considere o problema irrestrito,

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & x \in \mathbb{R}^n \end{aligned} \tag{3.35}$$

em que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa e em geral não diferenciável, o processo iterativo gera uma sequência de pontos da seguinte forma:

$$x^{k+1} = x^k - \alpha_k s^k \tag{3.36}$$

onde x^k é a k -ésima iteração, s^k um subgradiente qualquer de f no ponto x^k e $\alpha_k > 0$ é o k -ésimo tamanho do passo.

Um subgradiente de f no ponto x é um vetor qualquer s que satisfaz a inequação (3.33). Quando s^k é um subgradiente de f no ponto x^k escreve-se $s^k \in \partial f(x^k)$.

Um ponto \bar{x} é solução do problema (3.35), se $0 \in \partial f(\bar{x})$ (ver teorema 3.13).

Teorema 3.20 (POLYAK, 1987) *O ponto \bar{x} é mínimo global de uma função convexa, ou seja, qualquer sequência converge a ele. O conjunto limitado do ponto mínimo X^* é fracamente estável, ou seja, qualquer sequência convergente tem limite de pontos que pertencem a X^* .*

O Método do Subgradiente não é necessariamente de descida, ou seja, o valor da função f pode aumentar (BOYD; MUTAPCIC, 2008). Esta é uma grande dificuldade encontrada neste método.

3.13.1 Direção

Os Métodos Subgradientes são iterativos e provém do Método de Cauchy, onde encontra-se uma direção d^k tal que $f(x^k + d^k) < f(x^k)$ (TANAKA, 2011). A intenção é escrever o problema irrestrito (3.35) onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função convexa não necessariamente diferenciável, na forma

$$\begin{aligned} \min \quad & f(x^k + d) - f(x^k) \\ \text{s.a} \quad & d \in \mathbb{R}^n, \end{aligned} \tag{3.37}$$

que tem uma solução $d^k \neq 0$. Devido a expansão de primeira ordem de Taylor, tem-se

$$f(x^k + d) - f(x^k) = f'(x^k; d) + \|d\|\varepsilon(d),$$

onde $\varepsilon(d) \rightarrow 0$ quando $\|d\| \rightarrow 0$ e a identidade $f'(x^k; d) = \nabla f(x^k)^T d$. Fazendo uma analogia com o gradiente, tem-se:

$$\begin{cases} \min & f'(x^k; d) \\ \text{s.a} & \|d\| \leq 1 \end{cases} \iff \begin{cases} \min & \nabla f(x^k)^T d \\ \text{s.a} & \|d\| \leq 1, \end{cases}$$

onde a restrição adicional $\|d\| \leq 1$ é necessária, pois a função $f'(x^k; \cdot)$ é positivamente homogênea. Assim, no caso diferenciável a direção de Cauchy tem a forma $d^k = -\frac{\nabla f(x^k)}{\|\nabla f(x^k)\|}$.

Por outro lado, no caso não diferenciável, por analogia ao Método de Cauchy, considera-se a direção d^k dada por:

$$d^k = -\frac{s^k}{\|s^k\|}.$$

Porém, tal direção não é necessariamente de descida, a escolha adequada do tamanho do passo pode garantir a convergência do método. A seguir, são apresentadas algumas regras para a escolha do tamanho dos passos, encontradas na literatura.

3.13.2 Escolha do Tamanho do Passo

A escolha do tamanho do passo α_k em cada iteração nesse método é fundamental, pois se os passos forem muito pequenos, x^k se aproximará lentamente do ponto ótimo. Por outro lado, se forem excessivamente longos, o método poderá oscilar em torno da solução.

A seguir, algumas regras básicas para escolha do tamanho do passo.

1. Tamanho do passo constante:

$\alpha_k = \alpha$, é positivo e constante, independente da iteração k .

2. Passo da Série Divergente:

Neste caso o tamanho do passo tende a zero (POLYAK, 1987).

$$\alpha_k > 0, \quad \lim_{k \rightarrow \infty} \alpha_k = 0 \quad \text{e} \quad \sum_{k=1}^{\infty} \alpha_k = \infty.$$

Um exemplo típico segundo BOYD e MUTAPCIC (2008) é $\alpha_k = \frac{c}{\sqrt{k}}$ com $c \geq 0$. Outros exemplos de sequências sugeridas por (POLYAK, 1987) são:

$$\alpha_k = \frac{\alpha}{k+c}, \quad \alpha_k = \frac{\alpha}{k^\rho}, \quad 0 < \rho \leq 1 \text{ e } \alpha_k = \frac{\alpha}{k \ln k}.$$

3. Quando se conhece o ótimo da função (f^*):

Polyak propõe uma escolha do tamanho do passo quando se conhece a solução f^* , sendo o passo dado por: $\alpha_k = \frac{r[f(x^k) - f^*]}{\|s^k\|^2}$ no qual $0 < r < 2$. O método converge com velocidade de uma progressão geométrica. Quando f^* for desconhecido, utiliza-se \bar{f} como a melhor estimativa de f^* , desta forma o tamanho do passo é dado por $\alpha_k = \frac{r[f(x^k) - \bar{f}]}{\|s^k\|^2}$, com $0 < r < 2$.

3.13.3 Critério de Parada

Uma das dificuldades dos Métodos Subgradientes é estabelecer o critério de parada, pois os subgradientes são encontrados arbitrariamente e, por isso, não possuem informações sobre a condição de otimalidade. Descobrir que $0 \in \partial f(x)$ pode ser uma tarefa inviável pois, na maioria das vezes, é muito difícil encontrar o subdiferencial $\partial f(x)$. Um critério de parada para o Método Subgradiente pode ser dado pelo número máximo de iterações atingido ou pela condição $x^{k+1} \approx x^k$, ou ainda se $f(x^{k+1}) \approx f(x^k)$ (SOUZA, 2008).

3.13.4 Método Subgradiente Básico

Este método foi apresentado por POLYAK (1987), para uma função convexa f ao considerar $\partial f(x^k)$ o subgradiente no ponto x^k . Para este método, a direção é dada por $d^k = -\frac{s^k}{\|s^k\|}$ e o tamanho do passo pela sequência $\{\alpha_k\}$, com as condições: $\alpha > 0$, $\alpha_k \rightarrow 0$ e $\sum_{k=0}^{\infty} \alpha_k = \infty$. Na subseção 3.13.2, foram apresentados alguns exemplos que satisfazem esta condição.

A seguir, o algoritmo básico do Método Subgradiente (POLYAK, 1987):

Algoritmo 3.15 MÉTODO BÁSICO DO SUBGRADIENTE

1º passo: Escolher $x^0 \in \mathbb{R}^n$ e tomar $k = 0$

2º passo: Obter: $s^k \in \partial f(x^k)$ e $f(x^k)$.

Critério de parada: $0 \in \partial f(x^k)$

3º passo: Escolher α_k apropriado, tal que: $\alpha > 0$, $\alpha_k \rightarrow 0$ e $\sum_{k=0}^{\infty} \alpha_k = \infty$.

4º passo: Definir a direção como $d^k = -\frac{s^k}{\|s^k\|}$.

5º passo: Definir $x^{k+1} = x^k + \alpha_k d^k$ e $k = k + 1$. Retorne para o 2º passo.

3.13.5 Método Subgradiente quando f^* é conhecido

O método consiste em escolher o tamanho do passo dado por:

$$\alpha_k = \frac{f(x^k) - f^*}{\|\partial f(x^k)\|^2},$$

quando o valor ótimo da função f^* é conhecido (POLYAK, 1987). Portanto, tem-se uma variação do Algoritmo básico do subgradiente 3.15 no 5º passo, onde x^{k+1} é dado por:

$$x^{k+1} = x^k - \frac{f(x^k) - f^*}{\|\partial f(x^k)\|^2} \partial f(x^k). \quad (3.38)$$

Teorema 3.21 (POLYAK, 1987) *Seja f uma função convexa em \mathbb{R}^n , cujo conjunto de pontos minimizadores é não vazio. Então, na equação (3.38) tem-se que $x^k \rightarrow \bar{x} \in X^*$. Além disso, a estimativa da taxa de convergência, para uma função f arbitrária, é dada por:*

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

Para função que possui um mínimo, o método converge com velocidade de uma progressão geométrica.

Demonstração: Ver (POLYAK, 1987).

Quando f^* for desconhecido, POLYAK (1987) sugere uma modificação no método. É aplicado o processo iterativo para um valor \bar{f} que representa a melhor estimativa de f^* , ou seja:

$$x^{k+1} = x^k - \frac{f(x^k) - \bar{f}}{\|\partial f(x^k)\|^2} \partial f(x^k).$$

3.13.6 Método de Dilatação do Espaço - Shor

É um método baseado em algoritmos para minimização de funções convexas em que o movimento na direção do subgradiente é combinado com uma dilatação de espaço ao longo desta direção (MINOUX, 1986). Considere a k -ésima iteração do Método Subgradiente onde é realizado a substituição da variável x por $B_k y$, sendo B_k uma matriz não singular $n \times n$ e, em seguida, aplica-se a Definição 3.11 de subgradiente na função $\varphi(y) = f(B_k y)$. O que ocorre é um deslocamento no espaço das variáveis x de acordo com a direção:

$$d = -B_k B_k^T s^k$$

onde s^k é um subgradiente de f no ponto x^k . Ou seja, para todo $x \in \mathbb{R}^n$:

$$f(x) \geq f(x^k) + (s^k)^T (x - x^k),$$

ao fazer $x = B_k y$ tem-se:

$$f(B_k y) \geq f(B_k y^k) + (s^k)^T (B_k y - B_k y^k), \quad \forall y \in \mathbb{R}^n$$

ou

$$\varphi(y) \geq \varphi(y^k) + (s^k)^T B_k (y - y^k), \quad \forall y \in \mathbb{R}^n.$$

Isto prova que $((s^k)^T B_k)^T = B_k^T s^k$ é um subgradiente de φ em $y^k = B_k^{-1} x^k$. Considere α_k o tamanho do passo na k -ésima iteração do algoritmo 3.15 aplicado a função $\varphi(y)$, a próxima iteração é dada por:

$$y^{k+1} = y^k - \alpha_k B_k^T s^k$$

e, conseqüentemente, $x^{k+1} = B_k y^{k+1}$, portanto a próxima iteração de x^{k+1} será dada por

$$x^{k+1} = B_k y^k - \alpha_k B_k B_k^T s^k = x^k - \alpha_k B_k B_k^T s^k.$$

A matriz $B_k = R_{a_0} d_0 R_{a_1} d_1 \dots R_{a_{k-1}} d_{k-1}$ é atualizada da seguinte forma:

$$B_{k+1} = B_k R_{a_k}(d_k)$$

onde $R_{a_k}(d)$ é o operador de dilatação do espaço ao longo da direção d com coeficiente de dilatação a_k de modo que $\|d\| = 1$, dado por:

$$R_a(d) = (a - 1)dd^T + I, \quad (3.39)$$

a matriz I é uma matriz identidade de ordem $n \times n$. A seguir, é apresentado o algoritmo para o Método Shor de Dilatação do Espaço.

Algoritmo 3.16 *MÉTODO SHOR DE DILATAÇÃO DE ESPAÇO*

1º passo: Considere x^0 o ponto inicial, α_0 o tamanho do passo nessa iteração, s^0 um subgradiente no ponto x^0 , $B_0 = I$ (matriz identidade) e $k = 0$.

2º passo: Obter: $s^k \in \partial f(x^k)$ e $f(x^k)$.

Critério de parada: $0 \in \partial f(x^k)$

3º passo: Escolher $a_k \geq 0$ e α_k apropriado.

4º passo: Definir a direção como $d^k = -B_k B_k^T s^k$.

5º passo: Calcular $x^{k+1} = x^k - \alpha_k B_k B_k^T s^k$.

6º passo: Atualizar $B_{k+1} = B_k R_{a_k}(d_k)$ e $k = k + 1$. Retorne para o 2º passo.

Pode haver várias versões desse método, basta escolher várias regras para determinar a direção d^k e as sequências $\{\alpha_k\}$ e $\{a_k\}$ (MINOUX, 1986). Os Métodos de Dilatação do Espaço são mais flexíveis do que o Método Básico do Subgradiente, pois além de depender da sequência de tamanho dos passos $\{\alpha_k\}$, depende também da sequência de coeficientes de dilatação $\{a_k\}$.

A seguir, são apresentadas algumas variações do Método de Dilatação do Espaço.

3.13.7 Método Subgradiente com Dilatação do Espaço na Direção da Diferença de Dois Subgradientes Sucessivos

O Método Subgradiente com Dilatação do Espaço na Direção da Diferença de Dois Subgradientes Sucessivos é considerado um dos métodos mais eficientes para solução de problemas de otimização não diferenciável (MINOUX, 1986). Este método é baseado no Método de Dilatação do Espaço ao considerar a direção na diferença de dois subgradientes sucessivos.

Suponha que após a k -ésima iteração do Método Subgradiente seja realizada uma substituição de variável $x = B_k y$, onde matriz B_k é não singular e $n \times n$. Em seguida, aplica-se o Método Subgradiente na função $\varphi(y) = f(B_k^T y)$, considerando:

- $a_k = a$ com $(0 < a < 1)$, $\forall k$;
- $d^k = \frac{r^k}{\|r^k\|}$, com $r^k = B_k^T [s^{k+1} - s^k]$;
- α_k é encontrado de modo a minimizar f na direção $-B_k B_k^T s^k$.

A atualização de B_k é dada por $B_{k+1} = B_k R_{a_k}(d_k)$. A direção $-B_k B_k^T s^k$ nem sempre é uma direção de descida. Se for este o caso, então $x^{k+1} = x^k$ e o algoritmo não pode mais avançar. Uma forma de contornar este problema é proposto por Shor ao substituir o 2º passo do algoritmo 3.16 por:

1. Na iteração k , o corrente ponto é x^k . Encontrar x^{k+1} por meio da fórmula

$$x^{k+1} = x^k - \alpha_k B_k B_k^T s^k,$$

onde α_k é o minimizador de f na direção $-B_k B_k^T s^k$.

Se $\alpha_k = 0$, então

$$\begin{aligned} x^{k+1} &\leftarrow x^k \\ x^k &\leftarrow x^{k-1} \\ y^{k+1} &\leftarrow y^k \\ y^k &\leftarrow y^{k-1}. \end{aligned}$$

Se $\alpha_k \neq 0$, então determine s^{k+1} , o subgradiente de f no ponto x^{k+1} e atualize B_{k+1} conforme o item a seguir.

2. Dado:

$$d^k = \frac{r^k}{\|r^k\|} \quad \text{com} \quad r^k = B_k^T [s^{k+1} - s^k]$$

atualize

$$B_{k+1} = B_k R_{a_k}(d^k).$$

Assim, mesmo que o algoritmo permaneça em um mesmo ponto durante um certo número de iterações sucessivas, as matrizes B_k continuam sendo atualizadas pela dilatação do espaço na direção da diferença entre os dois últimos subgradientes. O algoritmo irá avançar assim que encontrar a direção de descida (MINOUX, 1986).

3.13.8 Método de Dilatação do Espaço na Direção do Subgradiente no Ponto Atual

Neste método a cada iteração k deve se escolher o coeficiente de dilatação a_k , a direção d^k e o tamanho do passo α_k , tal que:

- $a_k = a < 1, \quad (\forall k)$;
- $d^k = \frac{B_k^T s^k}{\|B_k^T s^k\|}$;
- $\alpha_k = \rho \frac{f(x^k) - m}{\|B_k^T s^k\|^2}$.

Onde ρ é o coeficiente de relaxação que satisfaz $0 < \rho \leq 2$ e m é uma aproximação do valor ótimo da função f^* .

Algoritmo 3.17 MÉTODO DE DILATAÇÃO DO ESPAÇO NA DIREÇÃO DO SUBGRADIENTE NO PONTO ATUAL

1º passo: Escolher $x^0 \in \mathbb{R}^n$ e tomar $k = 0$

2º passo: Obter: $s^k \in \partial f(x^k)$ e $f(x^k)$.

Critério de parada: $0 \in \partial f(x^k)$

3º passo: Escolher $a_k = a < 1 (\forall k)$

3º passo: Escolher $\alpha_k = \rho \frac{f(x^k) - m}{\|B_k^T s^k\|^2}$

4º passo: Definir a direção como $d^k = \frac{B_k^T s^k}{\|B_k^T s^k\|}$.

5º passo: Definir $x^{k+1} = x^k - \alpha_k d^k$ e $k = k + 1$. Retorne para o 2º passo.

3.13.9 Método de Elipsóides

Considere R o limite superior de $\|x^0 - x^*\|$, que representa a distância entre o ponto inicial e o ponto ótimo solução do problema de minimização. O Método de Elipsóides consiste em escolher as sequências:

- $a_k = a = \sqrt{\frac{n-1}{n+1}} \forall k$, onde n é a dimensão do espaço;
- $d^k = \frac{B_k^T s^k}{\|B_k^T s^k\|}$;
- $\alpha_k = \frac{R}{n+1} \frac{\beta^k}{\|B_k^T s^k\|}$, com $\beta = \frac{n}{\sqrt{n^2-1}}$.

O método pode ser representado geometricamente da seguinte forma:

Considere a bola de centro x^1 e raio R , ou seja, $H_1 = S_1 = \{x; \|x - x^1\| \leq R\}$. Encontra-se um subgradiente no ponto x^1 , se $g(x^1) = 0$ então x^1 é um mínimo de f . Caso contrário, forma-se um hiperplano $s^1(x - x^1) = 0$, assim $x^* \in \bar{S} \equiv S_1 \cap \{x; s^1(x - x^1) \leq 0\}$. Determina-se o volume mínimo do elipsóide de \bar{S} . O centro do novo elipsóide é $x^2 = x^1 - \alpha_2 d^1$ no qual $d^1 = \frac{s^1}{\|s^1\|}$. Realiza-se a dilatação do espaço na direção d^1 . No novo espaço a imagem de H_2 é $\{y; \|y - y^2\| \leq R\beta\}$, no qual, $y^2 = R_a(s^1)x^2$, ou seja, tem-se uma esfera de raio $R\beta$ e centro y^2 . Em seguida, aplica-se o algoritmo no novo espaço, assim sucessivamente, até que o critério de parada seja satisfeito (SOUZA, 2008).

Algoritmo 3.18 MÉTODO ELIPSÓIDES

Considere, $H_k = B_k B_k^T$:

1º passo: Seja x^0 o ponto inicial e s^0 o subgradiente de f no ponto x^0 . $H_0 = RI$, sendo R o limite superior de $\|x^0 - x^*\|$ e I a matriz identidade $n \times n$.

2º passo: Seja x^k o ponto na iteração k e s^k o subgradiente neste ponto.

Define-se x^{k+1} por:

$$x^{k+1} = x^k - \frac{1}{n+1} \frac{H_k s^k}{\sqrt{(s^k)^T H_k s^k}}$$

e H_{k+1} atualizada por:

$$H_{k+1} = \frac{n^2}{n^2-1} \left(H_k - \frac{2}{n+1} \frac{(H_k s^{k+1})(H_k)^T}{(s^{k+1})^T H_k s^{k+1}} \right).$$

3º passo: Se o critério de parada for satisfeito: PARE. Caso contrário, faça

$k = k + 1$ e retorne ao 2º passo.

4 REFORMULAÇÕES DO NEP E GNEP

No Capítulo 2, foram apresentadas as definições do Problema de Equilíbrio de Nash e do Problema de Equilíbrio de Nash Generalizado. Neste capítulo, serão apresentadas algumas reformulações do GNEP. Essas reformulações são, na verdade, formulações equivalentes do GNEP ou subclasses do GNEP, com o objetivo de reduzir problemas extremamente difíceis, pois o GNEP, na sua forma geral, é muito difícil, em problemas mais fáceis. As reformulações apresentadas a seguir são consideradas recentes, e são baseadas no artigo de FACCHINEI e KANZOW (2010a). Também será apresentada a reformulação para o Problema de Equilíbrio de Nash via Ponto Proximal.

Para mais clareza na exposição dos conteúdos, a seguir é reescrito o problema GNEP

$$\begin{aligned} \min_{x^j} \quad & \theta_j(x^j, x^{-j}) \\ \text{s.a} \quad & x^j \in X_j(x^{-j}), \end{aligned} \tag{4.1}$$

sendo θ_j, x^j, x^{-j} e X_j definidos no Capítulo 2.

4.1 FUNÇÃO NIKAIDO-ISODA (NI-função)

A função Nikaido-Isoda ou Ky-Fan é uma das mais importantes reformulações do GNEP. Foi introduzida inicialmente para provar a existência de um equilíbrio de Nash e tornou-se uma ferramenta para o desenvolvimento de métodos numéricos para a solução do Problema de Equilíbrio de Nash Generalizado (HEUSINGER, 2009) e (FACCHINEI; KANZOW, 2010a). A função Nikaido-Isoda é dada por:

$$\Psi(x, y) = \sum_{j=1}^N [\theta_j(x^j, x^{-j}) - \theta_j(y^j, x^{-j})]. \tag{4.2}$$

A função Ψ depende da função objetivo de cada jogador j , não depende dos

espaços estratégicos. Suponha que x e y são dois pontos viáveis para o GNEP, cada alteração representa uma melhoria da função objetivo do jogador j quando ele muda de ação de x^j para y^j , enquanto todos os outros jogadores permanecem com a escolha x^{-j} . O equilíbrio de GNEP é caracterizado pela impossibilidade de obter alguma melhoria para qualquer possível escolha y , isto é, essencialmente, o conteúdo do seguinte teorema (FACCHINEI; KANZOW, 2010a).

Teorema 4.1 *Sejam Ψ uma NI-função,*

$$X(x) := \prod_{j=1}^N X_j(x^{-j}) \quad e \quad \hat{V}(x) := \sup_{y \in X(x)} \Psi(x, y). \quad (4.3)$$

Então, valem as seguintes propriedades:

- (a) $\hat{V}(x) \geq 0$ para todo $x \in X(x)$.
- (b) $\hat{V}(\bar{x}) = 0$ e $\bar{x} \in X(\bar{x})$ se, e somente se, \bar{x} for solução do GNEP.

O Teorema 4.1 caracteriza as soluções de um GNEP como o conjunto de pontos $\bar{x} \in X(\bar{x})$, tais que $0 = \hat{V}(\bar{x}) \leq \hat{V}(x)$, para todo $x \in X(\bar{x})$. Pode-se dizer que \bar{x} é uma solução do GNEP se, e somente se, é um minimizador global com valor zero da função objetivo do problema

$$\begin{aligned} \min \quad & \hat{V}(x) \\ \text{s.a} \quad & x \in X(x). \end{aligned} \quad (4.4)$$

Esse problema é chamado de ‘quase otimização do problema de restrição’, pois não é um problema de otimização padrão, uma vez que a viabilidade depende da variável x .

Esta reformulação não é de grande interesse prático, pois tem algumas desvantagens do ponto de vista matemático e prático, mesmo para o NEP padrão. Portanto, dado um vetor x , o supremo em (4.3) pode não existir a menos sob hipóteses adicionais (como a compacidade de $X(x)$). Por outro lado, esse supremo, se existir, normalmente não é alcançado em um único ponto, que, por sua vez, implica que o mapeamento \hat{V} e a reformulação de otimização correspondentes (4.4) não são, em geral, diferenciáveis (HEUSINGER; KANZOW, 2009).

Para solucionar essas deficiências, HEUSINGER e KANZOW (2009) propõem em seu artigo a ideia de regularização para o GNEP, considerando $\sigma > 0$ um parâmetro

fixado e definindo regularização da função Nikaido-Isoda por

$$(x, y) \in \mathbb{R}^N \times \mathbb{R}^N \mapsto \Psi_\sigma(x, y) := \sum_{j=1}^N \left[\theta_j(x^j, x^{-j}) - \theta_j(y^j, x^{-j}) - \frac{\sigma}{2} \|x^j - y^j\|^2 \right]. \quad (4.5)$$

Logo, tem-se a seguinte função objetivo para mapeamento $x \in X(x)$:

$$V_\sigma(x) = \sup_{y \in X(x)} \Psi_\sigma(x, y) = \sum_{j=1}^N \left\{ \theta_j(x^j, x^{-j}) - \min_{y^j \in X_j(x^{-j})} \left[\theta_j(y^j, x^{-j}) + \frac{\sigma}{2} \|x^j - y^j\|^2 \right] \right\}.$$

Veja, no teorema a seguir, as propriedades do mapeamento $V_\sigma(x)$ (HEUSINGER; KANZOW, 2009).

Teorema 4.2 *A função regularizada $V_\sigma(x)$ tem as seguintes propriedades:*

- a) $V_\sigma(x) > 0$ para todo $x \in X(x)$.
- b) \bar{x} é um equilíbrio de Nash generalizado se, e somente se $\bar{x} \in X(\bar{x})$ e $V_\sigma(\bar{x}) = 0$.
- c) Para cada $x \in X(x)$, existe um único vetor $y_\sigma(x) = (y_\sigma^1(x), \dots, y_\sigma^N(x))$, tal que

$$\operatorname{argmin}_{y^j \in X_j(x^{-j})} \left[\theta_j(y^j, x^{-j}) + \frac{\sigma}{2} \|x^j - y^j\|^2 \right] = y_\sigma^N(x). \quad (4.6)$$

Pelo teorema anterior, encontrar a solução para o GNEP é equivalente ao mínimo global do seguinte problema de minimização restrita

$$\begin{aligned} \min \quad & V_\sigma(x) \\ \text{s.a} \quad & x \in X(x), \end{aligned} \quad (4.7)$$

porém, a não diferenciabilidade do mapeamento $V_\sigma(x)$ ainda é uma desvantagem para a aplicação de métodos de otimização do problema (4.7).

A função Nikaido-Isoda é utilizada por HEUSINGER e KANZOW (2009) na apresentação de uma reformulação suave de otimização restrita do GNEP, que tem propriedades muito diferentes da apresentada anteriormente. Esta reformulação é considerada boa, porém, ela não dá uma caracterização completa de todas as soluções da GNEP, mas fornece uma caracterização dos equilíbrios de Nash chamados normalizados.

Definição 4.3 *Um vetor $\bar{x} \in X(x)$ é chamado um Equilíbrio de Nash normalizado do*

GNEP, se mantém $\sup_{y \in X(x)} \Psi(\bar{x}, y) = 0$, onde Ψ é a função Nikaido-Isoda na forma (4.2).

Observe que um Equilíbrio de Nash normalizado é também solução do GNEP. Ressalta-se que a recíproca nem sempre se verifica.

A reformulação suave do GNEP é baseado na regularização da função Nikaido-Isoda $\Psi_\sigma(x, y)$ em (4.5). Com base nesse mapeamento define-se:

$$\begin{aligned} \hat{V}_\sigma(x) &:= \sup_{y \in X(x)} \Psi_\sigma(x, y) = \\ &= \max_{y \in X(x)} \sum_{j=1}^N \left[\theta_j(x^j, x^{-j}) - \theta_j(y^j, x^{-j}) - \frac{\sigma}{2} \|x^j - y^j\|^2 \right] = \\ &= \max_{y \in X(x)} \left[\Psi(x, y) - \frac{\sigma}{2} \|x^j - y^j\|^2 \right]. \end{aligned} \quad (4.8)$$

Veja, no teorema a seguir, as propriedades do mapeamento $\hat{V}_\sigma(x)$ (HEUSINGER; KANZOW, 2009).

Teorema 4.4 *A função regularizada $\hat{V}_\sigma(x)$ tem as seguintes propriedades:*

- a) $\hat{V}_\sigma(x) \geq 0$ para todo $x \in X(x)$.
- b) \bar{x} é um Equilíbrio de Nash Generalizado se, e somente se $\bar{x} \in X(\bar{x})$ e $\hat{V}_\sigma(\bar{x}) = 0$.
- c) Para cada $x \in X(x)$, existe um único vetor $\hat{y}_\sigma(x)$ tal que

$$\operatorname{argmax} \left[\Psi(x, y) - \frac{\sigma}{2} \|x - y\|^2 \right] = \hat{y}_\sigma(x), \quad (4.9)$$

e $\hat{y}_\sigma(x)$ é contínuo em x .

Este teorema caracteriza o Equilíbrio de Nash normalizado de um GNEP como o mínimo global do problema de otimização restrita a seguir

$$\begin{aligned} \min \quad & \hat{V}_\sigma(x) \\ \text{s.a} \quad & x \in X(x). \end{aligned} \quad (4.10)$$

Ao contrário da reformulação em (4.7), não há uma reformulação de todos os Equilíbrios de Nash Generalizados. A função regularizada $\hat{V}_\sigma(x)$ é continuamente diferenciável, desde que as funções θ_j sejam continuamente diferenciáveis para cada jogador $j = 1, \dots, N$. Veja mais informações em HEUSINGER e KANZOW (2009) e na tese de PhD de HEUSINGER (2009).

4.2 O CASO JUNTAMENTE CONVEXO

Um GNEP é juntamente convexo desde que satisfaça a Hipótese de Convexidade, ou seja, para cada jogador j e todos os outros jogadores x^{-j} , a função objetivo $\theta_j(\Delta, x^{-j})$ é convexa e o conjunto $X_j(x^{-j})$ é fechado e convexo.

Definição 4.5 *Um GNEP é juntamente convexo, se para algum $X \subseteq \mathbb{R}^n$ fechado e convexo, para todos $j = 1, \dots, N$, o conjunto $X_j(x^{-j})$ é dado como:*

$$X_j(x^{-j}) = \{x^j \in \mathbb{R}^{n_j} : (x^j, x^{-j}) \in X\}. \quad (4.11)$$

Quando os conjuntos $X_j(x^{-j})$ são definidos explicitamente por um sistema de desigualdades como em (2.2), então, verifica-se que (4.11) é equivalente à condição de $g^1 = g^2 = \dots = g^N := g$ e que $g(x)$ (componente a componente) é convexo em relação a todas as variáveis x . Além disso, neste caso, é óbvio que $X = \{x \in \mathbb{R}^n : g(x) \leq 0\}$ (FACCHINEI; KANZOW, 2010a).

Esta classe de problemas foi primeiramente estudada por Rosen, ver (ROSEN, 1965) em que para jogos com n -jogadores as restrições para cada jogador, bem como a sua função objetivo, pode depender da estratégia de cada jogador.

4.3 DESIGUALDADE VARIACIONAL (VI)

A Desigualdade Variacional também é uma das importantes reformulações do Problema de Equilíbrio de Nash Generalizado. Ultimamente tem sido alvo de muita atenção de pesquisadores, para a sua solução. Para melhor compreensão, considere o Exemplo 4.6 extraído do trabalho, ‘Inequações Variacionais Teoria e Aplicações’, de CHUQUIPOMA (2010).

Exemplo 4.6 *Se f é uma função continuamente diferenciável sobre o intervalo $I = [a, b]$, e x_0 é um ponto onde o mínimo ocorre, então três casos podem ocorrer:*

se $a < x_0 < b$ então $f'(x_0) = 0$;

se $x_0 = a$ então $f'(x_0) \geq 0$;

se $x_0 = b$ então $f'(x_0) \leq 0$.

Essas condições necessárias podem ser resumidas como o problema de encontrar $x_0 \in I$ tal que:

$$f'(x_0)(y - x_0) \geq 0, \quad \forall y \in I,$$

tal desigualdade será chamada de *Inequação Variacional*.

Considere o problema de minimização restrita

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & g(x) \leq 0, \end{aligned}$$

e o conjunto X definido na Seção 4.2, o problema de desigualdade variacional (*PVI*) corresponde a encontrar $\bar{x} \in X$ tal que:

$$\nabla f(\bar{x})^T (y - \bar{x}) \geq 0; \quad \forall y \in X.$$

Ao aplicar esse conceito para o GNEP em (2.1), tem-se

$$F(x)^T (y - x) \geq 0, \quad \forall y \in X(x)$$

com

$$F(x) := (\nabla_{x^j} \theta_j(x))_{j=1}^N \in \mathbb{R}^n \quad e \quad X(x) := \prod_{j=1}^N X_j(x^{-j}).$$

Teorema 4.7 *Seja um GNEP com função objetivo θ_j convexa e C^1 . Então, cada solução para o PVI (em que X é o conjunto na definição de convexidade conjunta e, como usual $F(x) := (\nabla_{x^j} \theta_j(x))_{j=1}^N$), é também uma solução do GNEP.*

4.3.1 Condições de KKT e Desigualdade Variacional (VI)

Em um GNEP o problema restrito (3.17) é reformulado do seguinte modo:

$$\begin{aligned} \min_{x^j} \quad & \theta_j(x^j, x^{-j}) \\ \text{s.a} \quad & g^j(x^j, x^{-j}) \leq 0 \quad e \quad x^j \in \mathbb{R}^{n_j}. \end{aligned} \tag{4.12}$$

No conjunto $X_j(x^{-j}) = \{x^j \in \mathbb{R}^{n_j} : g^j(x^j, x^{-j}) \leq 0\}$ as funções envolvidas são

C^1 e pode ser escrita as condições de KKT, apresentadas na seção 2.7 para o problema de cada jogador j dado por (4.12). As condições de KKT dos jogadores dá as condições de KKT do GNEP. Se para cada jogador j a função objetivo e as restrições têm condição de qualificação de Mangasarian-Fromowitz e a qualificação de Slater, então existe um vetor $\lambda^j \in \mathbb{R}^m$ de multiplicadores tal que,

$$\begin{aligned} \nabla_{x^j} \theta_j(x^j, x^{-j}) + \nabla_{x^j} g^j(x^j, x^{-j})^T \lambda^j &= 0, \\ \lambda_i^j g^j(x^j, x^{-j}) &= 0 \quad i = 1, \dots, m \quad e \quad j = 1, \dots, n, \\ \lambda_i^j &\geq 0, \quad i = 1, \dots, m, \quad e \quad j = 1, \dots, n, \\ g^j(x^j, x^{-j}) &\leq 0, \quad j = 1, \dots, n. \end{aligned} \tag{4.13}$$

sendo $g^j : \mathbb{R}^{n_j} \rightarrow \mathbb{R}^{m_j}$ (FACCHINEI; KANZOW, 2010a).

A solução do sistema KKT também será solução para o GNEP, segundo o teorema a seguir.

Teorema 4.8 (FACCHINEI; KANZOW, 2010a) *Seja um GNEP dado por (4.12) e assuma que todas as funções envolvidas são continuamente diferenciáveis.*

a) *Seja \bar{x} um equilíbrio do GNEP em que todos os subproblemas do jogador satisfazem uma condição de qualificação. Então, existe um $\bar{\lambda}$ que, juntamente com \bar{x} resolve o sistema (4.13).*

b) *Suponha que $(\bar{x}, \bar{\lambda})$ resolve o sistema (4.13) e que o GNEP satisfaça a hipótese de convexidade. Então \bar{x} é um ponto de equilíbrio do GNEP.*

Por outro lado, considere o correspondente PVI dado no Teorema 4.7. As condições de KKT deste VI são dadas por (FACCHINEI; KANZOW, 2010a):

$$\begin{aligned} F(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) &= 0 \\ g_i(\bar{x}) &\leq 0 \\ \lambda_i &\geq 0 \\ \lambda_i g_i(\bar{x}) &= 0, \quad \forall i = 1, \dots, m. \end{aligned} \tag{4.14}$$

O sistema KKT (4.14) é satisfeito somente se o sistema de KKT (4.13) satisfizer uma condição de qualificação e os multiplicadores forem iguais para todos os jogadores j . A solução para o GNEP é um equilíbrio variacional somente se os multiplicadores forem os mesmos para todos os jogadores j , conforme o teorema a seguir.

Teorema 4.9 *Considere o GNEP com g, θ_j convexas e também C^1 . Então, os seguintes itens são satisfeitos:*

1. *Seja \bar{x} uma solução para o PVI tal que as condições KKT em (4.14) são satisfeitas para os mesmos multiplicadores $\bar{\lambda}$. Então, \bar{x} é uma solução do GNEP, e as correspondentes condições de KKT em (4.13) estão satisfeitas com $\lambda^1 = \dots = \lambda^N = \bar{\lambda}$.*
2. *Por outro lado, suponha que \bar{x} é uma solução do GNEP de forma que as condições de KKT em (4.13) estão satisfeitos com $\bar{\lambda}^1 = \dots = \bar{\lambda}^N$. Então, $(\bar{x}, \bar{\lambda})$ com $\bar{\lambda} := \bar{\lambda}^1$ é um ponto de KKT e \bar{x} é uma solução de PVI.*

Em situações práticas, em que os jogadores possuem diferentes funções objetivo e muitas vezes com suas próprias restrições, os multiplicadores das restrições compartilhadas em um GNEP não são idênticos.

No artigo ‘Abordagem de Desigualdade Variacional para o Problema de Equilíbrio de Nash Generalizado’, do professor NABETANI (2008), são propostas duas parametrizações para PVI, o preço-dirigido e os métodos de parametrização de recursos direcionados que constroem famílias parametrizadas de PVI’s relacionadas com a GNEP.

4.4 REFORMULAÇÃO VIA PONTO PROXIMAL

Nas seções anteriores, foram apresentadas importantes reformulações de GNEP. Nesta seção, será vista uma reformulação um pouco mais geral, pois é baseada em um Problema de Equilíbrio mais geral que o Problema de Equilíbrio de Nash. Ou seja, o Problema de Equilíbrio de Nash é um caso particular do Problema de Equilíbrio tratado nesta seção. Esta reformulação foi proposta no artigo de IUSEM e SOSA (2007), sobre um método de Ponto Proximal para problemas de equilíbrio em espaços de Hilbert. Neste artigo, os autores propõem e analisam um método de Ponto Proximal para problemas de equilíbrio em espaço de Hilbert o qual estende o já conhecido método de Ponto Proximal para desigualdade variacional. Baseado nos resultados deste artigo são apresentados em IUSEM e NASRI (2007a) algoritmos de Lagrangeano Aumentado para o Problema de Equilíbrio. A descrição dos algoritmos de Lagrangeano Aumentado será apresentada no capítulo seguinte que trata dos métodos.

Nos próximos parágrafos, serão descritos os principais resultados mostrados no artigo de IUSEM e SOSA (2007). Inicialmente define-se o Problema de Equilíbrio tratado no artigo.

Seja H um espaço de Hilbert. Considere um conjunto convexo $K \subset H$ e $f : K \times K \rightarrow R$ tal que:

P1: $f(x, x) = 0$ para todo $x \in K$.

P2: $f(\cdot, y) : K \rightarrow R$ é semicontínua superior, para todo $y \in K$.

P3: $f(x, \cdot) : K \rightarrow R$ é convexa e semicontínua inferior, para todo $x \in K$.

O Problema de Equilíbrio, denotado por $EP(f, K)$, consiste em encontrar $x^* \in K$ tal que $f(x^*, y) \geq 0$ para todo $y \in K$. O conjunto solução de $EP(f, K)$ será denotado como $S(f, K)$.

Para os resultados apresentados pelos autores anteriormente citados é necessário uma certa propriedade de monotocidade de f . Para isso, os autores consideram as seguintes propriedades alternativas:

P4: $f(x, y) + f(y, x) \leq 0$ para todos $x, y \in K$.

P5: Sempre que $f(x, y) \geq 0$, com $x, y \in K$, se verifica que $f(y, x) \leq 0$.

P6: Existe $\theta \geq 0$ tal que $f(x, y) + f(y, x) \leq \theta \|x - y\|^2$ para todos $x, y \in K$.

Uma função f que satisfaz a propriedade P6 será chamada de θ -sobmonótona.

Os mesmos autores apresentam uma sequência de lemas e proposições que utilizam na demonstração dos teoremas principais. Aqui serão tratados apenas os teoremas, para isso é apresentado o algoritmo de Ponto Proximal tratado no artigo.

Considere f satisfazendo P1-P3, P5 e P6 e θ a constante de sobmonotonicidade de f . Dada uma sequência de parâmetros de regularização $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$, para algum $\bar{\gamma} > \theta$ e dado $x^0 \in K$. O método de Ponto Proximal para o Problema de Equilíbrio consiste em encontrar x^{k+1} como a única solução de $EP(f_k, K)$ em que $f_k : K \times K \rightarrow \mathbb{R}$ é definida como

$$f_k(x, y) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle. \quad (4.15)$$

Nota 4.10 Se $f(x, y) = \sup_{u \in T(x)} \langle u, y - x \rangle$, para algum operador monótono maximal ponto-conjunto $T : H \rightarrow P(H)$, então, a sequência definida por (4.15) é precisamente aquela gerada pelo método de Ponto Proximal para o problema de desigualdade variacional $VIP(T, K)$ apresentado em ROCKAFELLAR (1976a) e ROCKAFELLAR (1976b).

Para a convergência do método de Ponto Proximal para Problema de Equilí-

brio (que será chamado daqui em diante por PPEP) será necessária a noção de solução assintótica para o Problema de Equilíbrio.

Definição 4.11 *Diz-se que $\{z^k\} \in K$ é uma sequência assintoticamente resolvente para $EP(f_k, K)$ se $\lim_{k \rightarrow \infty} f_k(z^k, y) \geq 0$ para todo $y \in K$.*

No próximo teorema, é apresentado a convergência da sequência gerada pelo método PPEP. Antes, será apresentada a definição de função fracamente semicontínua superior que será usada na hipótese do teorema.

Definição 4.12 *Dizemos que f é fracamente semicontínua superior se qualquer sequência $\{y^k\}$ de elementos em K tal que $y^k \rightarrow y$, com $y \in K$ então $\limsup f(y^k) \leq f(y)$.*

Teorema 4.13 *Considere $EP(f, K)$ e f satisfazendo P1-P3, P5 e P6. Se $EP(f, K)$ tem soluções, então para todo $x^0 \in K$:*

- i. A sequência $\{x^k\}$ gerada por PPEP é limitada e $\lim_{k \rightarrow \infty} \|x^{k+1} - x^k\| = 0$;*
- ii. $\{x^k\}$ é uma sequência assintoticamente resolvente para $EP(f, K)$;*
- iii. Se $f(\cdot, y)$ é fracamente semicontínua superior para todo y em K , então $\{x^k\}$ é fracamente convergente para alguma solução \bar{x} de $EP(f, K)$.*

O teorema é atrativo no sentido que garante a convergência para uma solução do Problema de Equilíbrio. O próximo passo é relaxar a hipótese fracamente semicontínua superior no item (iii) do teorema. Para isso, os autores IUSEM e SOSA (2007) mostram alguns resultados intermediários que não serão apresentados aqui, neste trabalho.

É importante salientar que para enfraquecer a hipótese de fracamente semicontínua superior serão necessárias outras hipóteses. A primeira delas é a reformulação do Problema de Equilíbrio como desigualdade variacional, e a segunda é a extensão de monotonicidade de $f(x, \cdot)$ para um conjunto aberto contendo K .

O teorema a seguir livra-se da hipótese de fracamente semicontínua superior do Teorema (4.13) item (iii).

Teorema 4.14 *Se f satisfaz P1-P4, $EP(f, K)$ tem solução, e para todo x em K , $f(x, \cdot)$ pode ser estendida para um conjunto aberto $W \supset K$, e desta forma preserva sua convexidade. Então, a sequência $\{x^k\}$, gerada pelo algoritmo PPEP, é fracamente convergente para a solução de $EP(f, K)$ para todo x^0 em K .*

Com os resultados anteriores, é finalizada esta seção e uma continuação desse assunto é apresentado em IUSEM e NASRI (2007a), em que os autores discorrem sobre o Método de Lagrangeano Aumentado para resolver o Problema de Equilíbrio no caso de dimensão finita. Como o artigo apresenta um algoritmo e a versão linearizada para implementação computacional, será realizada a apresentação deste artigo no Capítulo 5 que trata dos métodos para resolver o Problema de Equilíbrio de Nash.

5 ALGORITMOS PARA A SOLUÇÃO DO NEP E GNEP

Neste capítulo, serão apresentados algoritmos de solução para GNEP e NEP, uma vez que o Problema de Equilíbrio de Nash é um caso particular do Problema de Equilíbrio e o conjunto solução de um NEP é um subconjunto do conjunto de todas as soluções de um GNEP. Tendo como base os conceitos de programação não linear, vistos no Capítulo 3. Os algoritmos a seguir são para GNEP geral e GNEP conjunto convexo, conforme FACCHINEI e KANZOW (2010a). Eles não são totalmente adequados, há necessidade de estudos sobre a convergência para o GNEP. O único caso de resultados mais confiáveis, é aquele em que o GNEP tem um conjunto de estrutura convexa.

5.1 MÉTODOS PRÁTICOS

Os métodos de Jacobi e Gauss-Seidel são métodos ‘naturais’ de decomposição. Considere o GNEP geral onde o subproblema do jogador j é dado por (2.1), ou seja,

$$\begin{aligned} \min_{x^j} \quad & \theta_j(x^j, x^{-j}) \\ \text{s.a} \quad & x^j \in X_j(x^{-j}). \end{aligned}$$

Primeiro será descrito o método não linear de Jacobi, segundo FACCHINEI e KANZOW (2010a).

Algoritmo 5.1 MÉTODO NÃO LINEAR DE JACOBI

Passo 1: Escolha um ponto de partida $x^0 = (x^{0,1}, \dots, x^{0,N})$ e defina $k = 0$.

Passo 2: Se x^k satisfaz um critério de terminação adequado: PARE.

Passo 3: Para $j = 1, \dots, N$, compute uma solução $x^{k+1,j}$, de

$$\begin{aligned} \min_{x^j} \quad & \theta_j(x^j, x^{k,-j}) \\ \text{s.a} \quad & x^j \in X_j(x^{k,-j}). \end{aligned}$$

Passo 4: Definir $x^{k+1} := (x^{k+1,1}, \dots, x^{k+1,N})$, $k \leftarrow k + 1$, e retorne ao passo 1.

Em cada iteração k , o Algoritmo 5.1 resolve N problemas de otimização. Para cada $j = 1, \dots, N$ a função objetivo

$$\theta_j(x^{k,1}, \dots, x^{k,j-1}, x^{k,j}, x^{k,j+1}, \dots, x^{k,N})$$

tem de ser minimizada sobre todos $x^j \in X_j(x^{-j})$, enquanto que todas as variáveis dos outros jogadores $\mu \neq j$ são fixas. No entanto, este algoritmo não utiliza as informações mais recentes, uma vez computado x^j , já tem as novas variáveis que podem ser usadas. No algoritmo de Gauss-Seidel, essas novas variáveis são utilizadas tanto na função objetivo θ_j como no conjunto viável. Desta forma, o método tem a seguinte forma, segundo FACCHINEI e KANZOW (2010a):

Algoritmo 5.2 *MÉTODO NÃO LINEAR DE GAUSS-SEIDEL*

Passo 1: Escolha um ponto de partida $x^0 = (x^{0,1}, \dots, x^{0,N})$ e defina $k = 0$.

Passo 2: Se x^k satisfaz um critério de terminação adequado: PARE.

Passo 3: Para $j = 1, \dots, N$, compute uma solução $x^{k+1,j}$, de

$$\begin{aligned} & \min_{x^j} \theta_j(x^{k+1,1}, \dots, x^{k+1,j-1}, x^j, x^{k,j+1}, \dots, x^{k,N}); \\ & \text{s.a. } x^j \in X_j(x^{k+1,1}, \dots, x^{k+1,j-1}, x^{k,j+1}, \dots, x^{k,N}) \end{aligned}$$

Passo 4: Definir $x^{k+1} := (x^{k+1,1}, \dots, x^{k+1,N})$, $k \leftarrow k + 1$, e retorne ao passo 1.

Nos Algoritmos 5.1 e 5.2 os resultados não são confiáveis, pois a sequência (desde que ela exista) pode convergir para um ponto \bar{x} que é apenas um ponto de acumulação da sequência $\{x^k\}$. Porém, esses algoritmos são simples e fáceis de implementar, o que explica a sua popularidade entre os profissionais (FACCHINEI; KANZOW, 2010a).

5.2 MÉTODO DE PENALIDADE EXATA

A seguir se analisará um algoritmo baseado em Penalidade Exata para um GNEP conforme o problema (4.12). O objetivo é transformar o problema GNEP em n problemas de Nash irrestritos, não diferenciáveis, usando uma abordagem de penalidade. Considere $\rho_j > 0$ um parâmetro de penalidade e o problema penalizado, em que cada jogador j tem por objetivo resolver o problema de otimização:

$$\min_{x^j} P_j(x; \rho_j) := \theta_j(x^j, x^{-j}) + \rho_j \|g_+^j(x^j, x^{-j})\| \quad (5.1)$$

sendo $\|\cdot\|$ a norma Euclidiana, $X_j(x^{-j}) = \{x^j \in \mathbb{R}^{n_j} : g^j(x^j, x^{-j}) \leq 0\}$ o conjunto viável e $g_+^j(x^j, x^{-j}) = \max\{0, g^j(x^j, x^{-j})\}$ é a penalidade. Note que o problema (5.1) é sem restrição e não diferenciável. A seguir, o algoritmo de Penalidade Exata segundo FACCHINEI e KANZOW (2010).

Algoritmo 5.3 PENALIDADE EXATA

Passo 1: Escolha $x^0 \in \mathbb{R}^n$, $\rho_j > 0$ e $c_j \in (0, 1)$ para $j = 1, \dots, N$, defina $k = 0$.

Passo 2: Se x^k satisfaz um critério de terminação adequado: PARE.

Passo 3: Faça $I^k := \{j : x^{k,j} \notin X_j(x^{k,-j})\}$. Para cada $j \in I^k$, se

$$\|\nabla_{x^j} \theta_j(x^{k,j}, x^{k,-j})\| > c_j [\rho_j \|(\nabla_{x^j} \|g_+^j(x^{k,j}, x^{k,-j})\|)]$$

então faça $\rho_j = 2\rho_j$.

Passo 4: Calcule a solução do problema penalizado (5.1), faça $k \leftarrow k+1$ e retorne ao passo 1.

O que é observado, sob adequadas condições, após um número finito de possíveis atualizações do parâmetro de penalidade, a solução do problema penalizado (5.1) é também uma solução GNEP original.

5.3 MÉTODO DE NEWTON LOCAL

A convergência no Método de Newton está sempre garantida para uma certa vizinhança que contém a solução do problema (ou raiz da função) e sua convergência é quadrática. Para a solução de um GNEP não é diferente, se assume que o ponto de partida está suficientemente próximo da solução. Por este motivo, o Método de Newton também é chamado de Método Local. Neste caso, será apresentada uma modificação no Método de Newton que em FACCHINEI e KANZOW (2010a) é chamado de Método Levenberg-Marquardt, aplicado ao sistema KKT $\Phi(x, \lambda) = 0$ dado em (4.14).

Na solução do GNEP, pelo Método de Levenberg-Marquardt, a iteração é dada por:

$$(x^{k+1}, \lambda^{k+1}) := (x^k, \lambda^k) + d^k$$

sendo $d^k = (\Delta x^k, \Delta \lambda^k)$ solução do sistema linear:

$$[J\Phi(x^k, \lambda^k)^T J\Phi(x^k, \lambda^k) + \alpha(x^k, \lambda^k)I]d = -J\Phi(x^k, \lambda^k)^T \Phi(x^k, \lambda^k),$$

com $\alpha(x^k, \lambda^k) := \|\Phi(x^k, \lambda^k)\|$ e $J\Phi(x^k, \lambda^k)$ a ‘Jacobiana Generalizada’. Como $\alpha > 0$, a matriz dos coeficientes $[J\Phi(x^k, \lambda^k)^T J\Phi(x^k, \lambda^k) + \alpha(x^k, \lambda^k)I]$ é definida positiva e portanto d é uma direção de descida. A Jacobiana generalizada utilizada no Método de Newton é uma matriz fácil de calcular, se a função Φ é continuamente diferenciável em torno do ponto de interesse.

5.4 MÉTODOS DE LAGRANGEANO AUMENTADO PARA PROBLEMA DE EQUILÍBRIO

Nesta seção, serão apresentados os algoritmos desenvolvidos pelos autores IUSEM e NASRI (2007a) para o Problema de Equilíbrio geral e, como caso especial, o Problema de Equilíbrio de Nash. A análise de convergência que os autores apresentam neste artigo está fortemente ligada ao trabalho de Métodos de Ponto Proximal para o Problema de Equilíbrio apresentada em IUSEM e SOSA (2007), o qual foi descrito no capítulo anterior que tratava das reformulações do Problema de Equilíbrio.

O Problema de Equilíbrio tratado, nesta seção, é o mesmo presente no capítulo anterior, com todas as notações. Desta forma, serão apresentadas as versões exata e inexata dos Métodos de Lagrangeano Aumentado para resolver $EP(f, K)$ em \mathbb{R}^n , para o caso no qual o conjunto viável K tem a forma:

$$K = \{x \in \mathbb{R}^n : h_i(x) \leq 0 \quad (1 \leq i \leq m)\} \quad (5.2)$$

em que h_i são convexas, para todo $i = 1, 2, \dots, m$.

O Método de Lagrangeano Aumentado já foi visto anteriormente. Aqui será reapresentado no modo em que foi utilizado em IUSEM e NASRI (2007a), ou seja, para o problema de otimização convexa no formato primal:

$$(PE) \quad \min \quad h_0(x) \quad \text{s.a.} \quad h_i(x) \leq 0 \quad (i = 1, 2, \dots, m),$$

em que $h_i : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são convexas, para todo $i = 0, 1, 2, \dots, m$.

A função Lagrangeana associada a (PE) é definida por:

$$(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m \mapsto l(x, \lambda) = h_0(x) + \sum_{i=1}^m \lambda_i h_i(x). \quad (5.3)$$

O problema dual associado ao problema (PE) é:

$$(PD) \quad \min \quad -\psi(\lambda) \quad \text{s.a.} \quad \lambda_i \geq 0 \quad (i = 1, 2, \dots, m),$$

sendo $\psi : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{-\infty\}$ definida por:

$$\psi(\lambda) = \inf \{l(x, \lambda) : x \in \mathbb{R}^n\}.$$

A função Lagrangeano Aumentado associada ao problema (PE) é dada por:

$$(x, \lambda, \gamma) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} \mapsto L(x, \lambda, \gamma) = h_0(x) + \gamma \sum_{i=1}^m \left[\left(\max \left\{ 0, \lambda_i + \frac{h_i(x)}{2\gamma} \right\} \right)^2 - \lambda_i^2 \right],$$

em que \mathbb{R}_{++} é o conjunto dos números reais positivos.

Na iteração k do Método de Lagrangeano Aumentado, aqui proposto, são conhecidos $\gamma^k > 0$, $x^k \in \mathbb{R}^n$ e $\lambda^k \in \mathbb{R}_+^m$, o algoritmo determina $x^{k+1} \in \mathbb{R}^n$ como minimizador de $L(x, \lambda^k, \gamma^k)$ e então atualiza os multiplicadores pela fórmula fechada

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h_i(x^{k+1})}{2\gamma^k} \right\}, \quad (i = 1, 2, \dots, m).$$

A prova de convergência para esta classe de métodos é por meio do problema dual. Supondo que os problemas (PE) e (PD) possuem soluções viáveis e os algoritmos estão bem definidos, mostra-se a convergência do problema dual (PD), em geral, utilizando os métodos de Pontos Proximais e então a equivalência entre o método de Ponto Proximal e Lagrangeano Aumentado. Para o Problema de Equilíbrio, esta será a metodologia utilizada pelos autores de IUSEM e NASRI (2007a). A qual será apresentada a seguir.

5.4.1 Lagrangeano Aumentado para Problema de Equilíbrio

Será assumido que o conjunto K , dado em (5.2), satisfaz alguma condição de qualificação, por exemplo, a condição de Slater dada na Definição 3.6. Além disso, o Problema de Equilíbrio $EP(f, K)$ é o mesmo apresentado na seção 4.4.

A bifunção Lagrangeana para o problema $EP(f, K)$, $L : (\mathbb{R}^n \times \mathbb{R}^m) \times (\mathbb{R}^n \times \mathbb{R}^m) \rightarrow \mathbb{R}$ é definida como (IUSEM; NASRI, 2007a):

$$L((x, \lambda), (y, \mu)) = f(x, y) + \sum_{i=1}^m \lambda_i h_i(y) - \sum_{i=1}^m \mu_i h_i(x). \quad (5.4)$$

Nota 5.1 *É importante destacar que o problema de otimização (PE) pode ser tomado como um caso particular do Problema de Equilíbrio $EP(f, K)$ fazendo $f(x, y) = h_0(y) - h_0(x)$. Neste caso, (5.4) reduz a $L((x, \lambda), (y, \mu)) = l(x, \lambda) - l(x, \mu)$, sendo l a função Lagrangeana (5.3) do problema (PE).*

Na sequência, será introduzido o Método Lagrangeano Aumentado Proximal para $EP(f, K)$. Primeiramente, será apresentada a versão exata e, então, a versão aproximada, assim como a teoria de convergência deste último que possui uma versão possível de implementação computacional.

Considere K o conjunto dado na relação (5.2). Baseado neste, define-se $p_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ como (IUSEM; NASRI, 2007a):

$$p_i(x, y, \lambda, \gamma) = \frac{\gamma}{2} \left[\left(\max \left\{ 0, \lambda_i + \frac{h_i(y)}{\gamma} \right\} \right)^2 - \left(\max \left\{ 0, \lambda_i + \frac{h_i(x)}{\gamma} \right\} \right)^2 \right], \quad (5.5)$$

e a função Lagrangeano Aumentado Proximal $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ para $EP(f, K)$ como:

$$L(x, y, \lambda, z, \gamma) = f(x, y) + \gamma \langle x - z, y - x \rangle + \gamma \sum_{i=1}^m p_i(x, y, \lambda, \gamma).$$

A versão exata do Método de Lagrangeano Aumentado-EMLA para $EP(f, K)$ é dada pelos passos a seguir (note que este é um algoritmo conceitual, pois não está fixado um critério para que o algoritmo pare, pelo fato de este ser a base para o algoritmo em versão implementável).

Algoritmo 5.4 MÉTODO LAGRANGEANO AUMENTADO EXATO (EMLA)

Passo 1: Tome uma sequência limitada $\{\gamma_k\} \subset \mathbb{R}_{++}$, $\lambda_0 \in \mathbb{R}_+^m$ e $k = 0$.

Passo 2: Determine x^{k+1} como a única solução de $EP(L_k, \mathbb{R}^n)$ com L_k dada por

$$L_k(x, y, \lambda^k, x^k, \gamma_k) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle + \sum_{i=1}^m p_i(x, y, \lambda^k, \gamma_k),$$

sendo p_i dada por (5.5).

Passo 3: Atualize a variável dual

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h_i(x^{k+1})}{\gamma_k} \right\}, \quad (i = 1, 2, \dots, m).$$

Passo 4: Faça $k = k + 1$ e volte ao Passo 2.

O próximo algoritmo é a versão inexata do Método de Lagrangeano Aumentado para resolver $EP(f, K)$. Este será chamado IALEM, como em IUSEM e NASRI (2007a).

Algoritmo 5.5 MÉTODO LAGRANGEANO AUMENTADO EXTRAGRADIENTE INEXATO (IALEM)

Passo 1: Tome uma sequência limitada $\{\gamma_k\} \subset \mathbb{R}_{++}$, $(x^0, \lambda_0) \in \mathbb{R}^n \times \mathbb{R}_+^m$, $k = 0$ e uma tolerância $\delta \in (0, 1)$.

Passo 2: Encontre um par (\bar{x}^k, e^k) tal que \bar{x}^k resolve $EP(L_k^e, \mathbb{R}^m)$ com L_k^e definida como:

$$L_k^e(x, y) := f(x, y) + \gamma_k \langle x^k - z, y - x \rangle + \sum_{i=1}^m p_i(x, y, \lambda^k, \gamma_k) - \langle e^k, y - x \rangle,$$

com p_i dada por (5.5) e e^k dado por:

$$\|e^k\| \leq \delta \gamma_k \|\bar{x}^k - x^k\|.$$

Passo 3: Atualize a variável dual

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h_i(\bar{x})}{\gamma_k} \right\}, \quad (i = 1, 2, \dots, m).$$

Passo 4: Se $(x^k, \lambda^k) = (\bar{x}^k, \lambda^{k+1})$ pare. Senão faça:

$$x^{k+1} = \bar{x}^k - \frac{1}{\gamma_k} e^k.$$

Passo 5: Faça $k = k + 1$ e volte ao Passo 2.

Note que o algoritmo EMLA é um caso particular de IALEM ao fazer $e^k = 0$, para todo $k \in \mathbb{N}$.

Na sequência, será apresentado o algoritmo do Método Ponto Proximal Extragradiante Inexato para resolver $EP(f, K)$, que será utilizado como uma ferramenta auxiliar para a análise de convergência do IALEM (IUSEM; NASRI, 2007a).

Algoritmo 5.6 *MÉTODO DE PONTO PROXIMAL EXTRAGRADIANTE INEXATO (IPPEM)*

Passo 1: Tome uma sequência limitada $\{\gamma_k\} \subset \mathbb{R}_{++}$, $(x^0, e^0) \in \mathbb{R}^n \times \mathbb{R}^n$, $k = 0$ e uma tolerância $\delta \in (0, 1)$.

Passo 2: Encontre um par $(\bar{x}^k, e^k) \in \mathbb{R}^n \times \mathbb{R}^n$ tal que \bar{x}^k resolve $EP(f_k^e, K)$ com

$$f_k^e(x, y) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle - \langle e^k, y - x \rangle$$

e

$$\|e^k\| \leq \delta \gamma_k \|\bar{x}^k - x^k\|.$$

Passo 3: Se $\bar{x}^k = x^k$, então pare. Caso contrário faça

$$x^{k+1} = \bar{x}^k - \frac{1}{\gamma_k} e^k.$$

O teorema seguinte mostra que se $EP(f, K)$ tem solução. Então, a sequência $\{x^k\}$ gerada pelo algoritmo IPPEM converge para alguma solução deste. Para tal, serão utilizadas as hipóteses apresentadas na seção 4.4, com $K = \mathbb{R}^n$ na hipótese P6 e também a seguinte hipótese:

$P6'$: Para todo $x^1, \dots, x^q \in \mathbb{R}^n$ e todo $t_1, \dots, t_q \in \mathbb{R}_+$ tal que $\sum_{l=1}^q t_l = 1$, vale

$$\sum_{l=1}^q t_l f(x^l), \sum_{k=1}^q t_k x^k \leq 0.$$

Teorema 5.2 *Considere $EP(f, K)$ satisfazendo as hipóteses P1-P5. Tome uma sequência $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$, para algum $\bar{\gamma} > 0$, sendo θ dada em P4. Seja $\{x^k\}$ a sequência gerada pelo algoritmo IPPEM. Se $EP(f, K)$ tem solução, então $\{x^k\}$ converge para alguma solução x^* de $EP(f, K)$.*

O próximo teorema trata da equivalência das sequências geradas pelos algoritmos IALEM, Método de Lagrangeano Aumentado Inexato para $EP(f, K)$ e IPPEM, Método de Ponto Proximal Extragradiente Inexato para $EP(L, \mathbb{R}^n \times \mathbb{R}_+^m)$ (IUSEM; NASRI, 2007a).

Teorema 5.3 *Considere $EP(f, K)$ que satisfaz P1-P4 sobre $\mathbb{R}^n \times \mathbb{R}^n$, uma sequência $\{\gamma_k\} \subset \mathbb{R}_{++}$ e a tolerância $\delta \in (0, 1)$. Seja $\{(x^k, \lambda^k)\}$ a sequência gerada pelo algoritmo IALEM aplicado a $EP(f, K)$, com o vetor erro associado $e^k \in \mathbb{R}^n$, e seja $\{(\bar{x}^k, \bar{\lambda}^k)\}$ a sequência gerada pelo algoritmo IPPEM aplicado a $EP(L, \mathbb{R}^n \times \mathbb{R}_+^m)$, com vetor erro $(e^k, 0) \in \mathbb{R}^n \times \mathbb{R}^m$, com os mesmos γ_k e δ . Se $(x^0, \lambda^0) = (\bar{x}^0, \bar{\lambda}^0)$, então $\{(x^k, \lambda^k)\} = \{(\bar{x}^k, \bar{\lambda}^k)\}$ para todo k .*

O resultado, a seguir, garante a convergência do algoritmo IALEM (IUSEM; NASRI, 2007a).

Teorema 5.4 *Considere $EP(f, K)$. Assuma que:*

- i. f satisfaça P1-P6 e $P6'$ sobre $\mathbb{R}^n \times \mathbb{R}^n$;*
- ii. O conjunto K é dado por (5.2);*
- iii. A condição de qualificação de Slater tratada na Definição 3.6 vale para K ;*
- iv. $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ para algum $\bar{\gamma} > \theta$, em que θ é a constante de submonotonocidade de f definida em P6.*

Seja $\{(x^k, \lambda^k)\}$ a sequência gerada pelo algoritmo IALEM para resolver $EP(f, K)$. Se $EP(f, K)$ tem solução, então a sequência $\{(x^k, \lambda^k)\}$ converge para algum par ótimo (x^, λ^*) de $EP(f, K)$ e, conseqüentemente, $x^* \in S(f, K)$.*

5.4.2 Versão Linearizada do Lagrangeano Aumentado

Os autores IUSEM e NASRI (2007a) propõem uma versão linearizada do algoritmo de Lagrangeano Aumentado Inexato, o qual será apresentado abaixo. Para isso, assumir que f e h_i , para $i = 1, \dots, m$, são funções diferenciáveis. Ainda, será usado a seguinte notação: $F_x(y) = f(x, y)$, e o gradiente de F_x será denotado como $\nabla F_x : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Ao linearizar a função Lagrangeana (5.4) como uma função de y próximo de $y = x$, obtém-se a função $\bar{L} : (\mathbb{R}^n \times \mathbb{R}^m) \times (\mathbb{R}^n \times \mathbb{R}^m) \rightarrow \mathbb{R}$ (será utilizado \bar{L} como a função linearizada para $EP(f, K)$, para não haver confusão com a função Lagrangeana L) definida como:

$$\bar{L}((x, \lambda), (y, \mu)) = \langle \nabla F_x(x), y - x \rangle + \sum_{i=1}^m \lambda_i \langle \nabla h_i(x), y - x \rangle + \sum_{i=1}^m (\lambda_i - \mu_i) h_i(x).$$

O algoritmo, a seguir, é a versão linearizada do Lagrangeano Aumentado.

Algoritmo 5.7 MÉTODO DE LAGRANGEANO AUMENTADO EXTRAGRADIENTE INEXATO LINEARIZADO (LIALEM)

Passo 1: Tome uma sequência limitada $\{\gamma^k\} \subset \mathbb{R}_{++}$, $k = 0$ e uma tolerância

$\delta \in (0, 1)$.

Passo 2: Dado (x^k, λ^k) defina $\bar{p}_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}_{++}$ como

$$\bar{p}_i(x, y, \lambda, \gamma) = \max \left\{ 0, \lambda_i^k + \frac{h_i(x)}{\gamma} \right\} \langle \nabla h_i(x), y - x \rangle, (i = 1, 2, \dots, m) \quad (5.6)$$

e encontre um par $(\bar{x}^k, e^k) \in \mathbb{R}^n \times \mathbb{R}^n$ tal que \bar{x}^k resolve $EP(\bar{L}_k^e, \mathbb{R}^n)$, em que:

$$\begin{aligned} \bar{L}_k^e(x, y) = & \langle \nabla F_x(x), y - x \rangle + \gamma_k \langle x - x^k, y - x \rangle + \\ & + \gamma_k \sum_{i=1}^m \bar{p}_i(x, y, \lambda^k, \gamma_k) - \langle e^k, y - x \rangle \end{aligned} \quad (5.7)$$

com \bar{p}_i como em (5.6) e e^k dado por:

$$\|e^k\| \leq \delta \gamma_k \|\bar{x}^k - x^k\|. \quad (5.8)$$

Passo 3: Defina

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h_i(\bar{x}^k)}{\gamma_k} \right\}, \quad (i = 1, 2, \dots, m). \quad (5.9)$$

Passo 4: Se $(x^k, \lambda^k) = (\bar{x}^k, \lambda^{k+1})$, então pare. Caso contrário faça

$$x^{k+1} = \bar{x}^k - \frac{1}{\gamma_k} e^k.$$

Para finalizar esta seção, é apresentado, a seguir o teorema que garante a convergência do algoritmo LIALEM, segundo IUSEM e NASRI (2007a).

Teorema 5.5 *Considere $EP(f, K)$ e assuma que:*

- i. f satisfaz P1-P6 e P6' sobre $\mathbb{R}^n \times \mathbb{R}^n$;*
- ii. f é continuamente diferenciável;*
- iii. h_i é diferenciável, para $i = 1, 2, \dots, m$;*
- iv. A condição de qualificação de Slater dada na Definição 3.6 é satisfeita para o conjunto K .*

Tome uma sequência $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ para algum $\bar{\gamma} > \theta$, em que θ é a constante de submonotonicidade definida em P6, e uma tolerância $\delta \in (0, 1)$. Seja $\{(x^k, \lambda^k)\}$ a sequência gerada pelo algoritmo LIALEM aplicado a $EP(f, K)$. Se $EP(f, K)$ tem solução então $\{(x^k, \lambda^k)\}$ converge para um par ótimo (x^, λ^*) de $EP(f, K)$ de modo que $x^* \in S(f, K)$. Além disso, se x^k não é uma solução de $EP(f, K)$ e \bar{x}^k é a única solução de $EP(\bar{L}_k^e, \mathbb{R}^n)$ com $e^k = 0$, então existe $\delta_k > 0$ tal que qualquer $x \in B(\bar{x}^k, \delta_k)$ resolve o k -ésimo subproblema do algoritmo LIALEM.*

A seguir, serão apresentadas algumas questões deixadas em aberto pelos autores IUSEM e NASRI (2007a):

1. A implementação dos métodos apresentados no artigo;
2. No passo 2 do algoritmo IALEM, o qual demanda a solução de \bar{x}^k de $EP(\bar{L}_k^e, \mathbb{R}^n)$. Pela aplicação dos resultados teóricos, mostrados no artigo, é obtido que \bar{x}^k pertence a $EP(\bar{L}_k^e, \mathbb{R}^n)$ se, e somente se,

$$\gamma_k[x^k - \bar{x}^k] + e^k \in \partial F_{\bar{x}^k}(\bar{x}^k) + \sum_{i=1}^m \max \left\{ 0, \lambda_i^k + \frac{h_i(\bar{x}^k)}{\gamma_k} \right\} \partial h_i(\bar{x}^k). \quad (5.10)$$

No caso suave a equação (5.10) pode ser resolvida como:

$$0 = \gamma_k[x - x^k] + \nabla F_x(x) + \sum_{i=1}^m \max \left\{ 0, \lambda_i^k + \frac{h_i(x)}{\gamma_k} \right\} \nabla h_i(x). \quad (5.11)$$

A dificuldade está no fato de que o último termo, do lado direito da equação (5.11), é contínuo, mas não diferenciável, devido a presença do máximo no somatório. Neste caso, métodos sem derivadas têm que ser utilizados para encontrar a solução.

3. Um outro método de ponto proximal inexato é apresentado, segundo os autores IUSEM e NASRI (2007a), para resolver $EP(f, K)$ em IUSEM e NASRI (2007b). Neste caso, no algoritmo I no artigo citado anteriormente, ao invés do passo 3 do algoritmo IPPEM, a solução x^k do subproblema é usada para construir um hiperplano H_k o qual separa $x^k \in S(f, K)$, e a próxima iteração x^{k+1} é denominada de projeção de Bregman de x^k sobre H_k . No contexto de dimensão finita a projeção de Bregman é a projeção ortogonal. A convergência do método pode ser encontrada no Teorema 5.5 de IUSEM e NASRI (2007b). Tanto o Lagrangeano Aumentado inexato para $EP(f, K)$, quanto sua versão linearizada, podem ser desenvolvidas via o algoritmo I do artigo citado anteriormente.

No capítulo seguinte, será apresentada a metodologia sugerida pela pesquisa do item 2 acima, em que é proposta a solução do Problema de Equilíbrio de Nash (NEP), baseada na resolução de um sistema não linear, $G(x) = 0$ dado em (5.11) sendo $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ não diferenciável em algum ponto do seu domínio.

6 METODOLOGIA PROPOSTA

Neste trabalho, são propostos algoritmos para resolução do Problema de Equilíbrio de Nash (NEP). São dois novos algoritmos, ambos baseados na resolução de um sistema não linear $G(x) = 0$, sendo $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ contínua mas não diferenciável em todos os pontos do domínio. Assim, a origem desses algoritmos é o artigo de IUSEM e NASRI (2007a), em que foi introduzido o Método de Lagrangeano Aumentado, para a resolução de um Problema de Equilíbrio geral, do qual o Problema de Equilíbrio de Nash é um caso particular.

Conforme descrito no capítulo anterior, sobre o Problema de Equilíbrio, na versão linearizada apresentada no Algoritmo 5.7, o subproblema a ser resolvido neste algoritmo é equivalente à resolução do sistema não linear $G(x) = 0$, sendo $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ dada por:

$$G(x) = \gamma_k[x - x^k] + \nabla F_x(x) + \sum_{i=1}^m \max \left\{ 0, \lambda_i^k + \frac{h_i(x)}{\gamma_k} \right\} \nabla h_i(x) \quad (6.1)$$

em que, $x^k \in \mathbb{R}^n$, $\{\gamma_k\} \subset \mathbb{R}_{++}$ é uma sequência limitada de parâmetros de regularização, $h_i(x), i = 1, \dots, m$ são as restrições do Problema de Equilíbrio, $\lambda_i^k, i = 1, \dots, m$ são os multiplicadores de Lagrange e $\nabla F_x(x)$ obtido ao substituir y por x em $\nabla F_x(y) = \nabla_y f(x, y)$, sendo $f(x, y)$ a bifunção dada pelo Problema de Equilíbrio $EP(f, K)$.

Existem algumas dificuldades com relação à solução do sistema (6.1), a saber: a falta de diferenciabilidade e de convexidade da $G(x)$. Apesar de $h(x)$ ser convexa e $f(x, y)$ também ser convexa na variável x (ver propriedade P3 da Seção 4.4), não temos como garantir que $G(x)$ seja convexa. O que pode ser garantido é a continuidade da $G(x)$.

Desta forma, serão desenvolvidas duas metodologias diferentes para resolver (6.1). A primeira consiste na suavização do termo não diferenciável para tornar as funções que definem $G(x)$ continuamente diferenciável. Após a suavização será aplicado o Método de Newton para resolver o sistema não linear (6.1). A segunda consiste em resolver o sistema (6.1) por meio de um problema de otimização, ou seja, ao invés de resolver $G(x) = 0$ será

resolvido o problema $\text{minimizar}\{f(x) : x \in \mathbb{R}^n\}$ com $f(x) = \frac{1}{2}\|G(x)\|^2$. Para isto, será utilizado um procedimento similar ao Método do Gradiente. Primeiramente, suaviza-se a função $G(x)$, como realizado para o método de Newton, aplica-se o Método do Gradiente com busca Barzilai Borwein para resolver $G(x) = 0$. Uma vez que o Método do Gradiente tem convergência lenta e do ponto de vista computacional pode até não convergir, este será substituído por Métodos Subgradientes para resolver o sistema $G(x) = 0$.

6.1 APROXIMAÇÃO QUADRÁTICA DO TERMO NÃO DIFERENCIÁVEL

Inicialmente será apresentada uma suavização da função $G(x)$ por meio de uma aproximação quadrática dos termos não diferenciáveis desta. Note que as componentes do sistema (6.1) podem ser reescritas na forma:

$$G_i(x) = \begin{cases} \gamma_k[x_i - x_i^k] + (\nabla F_x(x))_i + \sum_{i=1}^m \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) (\nabla h_i(x))_i & \text{se } \lambda_i^k + \frac{h_i(x)}{\gamma_k} > 0 \\ \gamma_k[x_i - x_i^k] + (\nabla F_x(x))_i & \text{caso contrário,} \end{cases}$$

para $i = 1, \dots, m$, $(\nabla F_x(x))_i$ é a componente i de $\nabla F_x(x)$ e $(\nabla h_i(x))_i$ é a componente i de $\nabla h_i(x)$.

Note que a função $G(x)$ só não é diferenciável para os índices i em que:

$$\lambda_i^k + \frac{h_i(x)}{\gamma_k} = 0.$$

Desta forma, o termo que envolve o máximo no somatório da $G(x)$ em (6.1) é o que deve ser suavizado.

Considere a função $p : \mathbb{R} \rightarrow \mathbb{R}$ dada por:

$$p(y) = \max\{0, y\} = \begin{cases} 0, & \text{se } y \leq 0 \\ y, & \text{se } y > 0, \end{cases}$$

para $y \neq 0$, tem-se:

$$p'(y) = \begin{cases} 0, & \text{se } y < 0 \\ 1, & \text{se } y > 0. \end{cases}$$

Note que existe uma não diferenciabilidade em $p(y)$ na origem. Assim será reali-

zada uma suavização no intervalo $[-\eta, \eta]$, com $\eta > 0$ e pequeno. Para tal será considerada uma função quadrática $q(y) = ay^2 + by + c$, com a , b e c a serem determinados de modo que a função ajustada seja continuamente diferenciável em todo \mathbb{R} . Desta forma, considere a função $\bar{p} : \mathbb{R} \rightarrow \mathbb{R}$ defina por:

$$\bar{p}(y) = \begin{cases} 0, & \text{se } y \leq -\eta \\ q(y), & \text{se } -\eta < y \leq \eta \\ y, & \text{se } y > \eta, \end{cases} \quad (6.2)$$

cuja derivada é dada por:

$$\bar{p}'(y) = \begin{cases} 0, & \text{se } y \leq -\eta \\ q'(y), & \text{se } -\eta < y \leq \eta \\ 1, & \text{se } y > \eta. \end{cases}$$

Como se procura $q(y)$ tal que $\bar{p}(y)$ seja continuamente diferenciável, serão exigidas as seguintes condições:

$$p(-\eta) = \bar{p}(-\eta), \quad p(\eta) = \bar{p}(\eta) \text{ e } p'(-\eta) = \bar{p}'(-\eta), \quad p'(\eta) = \bar{p}'(\eta).$$

Note que:

$$p(-\eta) = 0 = \bar{p}(-\eta),$$

está satisfeita por construção.

Desde que $q(y) = ay^2 + by + c$ e $q'(y) = 2ay + b$, para as demais condições impostas ter-se-á:

$$\begin{aligned} p(\eta) = \bar{p}(\eta) & \xrightarrow{\eta > 0} \eta = a\eta^2 + b\eta + c \\ p'(-\eta) = \bar{p}'(-\eta) & \xrightarrow{-\eta < 0} 0 = -2a\eta + b \\ p'(\eta) = \bar{p}'(\eta) & \xrightarrow{\eta > 0} 1 = 2a\eta + b. \end{aligned} \quad (6.3)$$

A solução do sistema (6.3) nas incógnitas a , b e c em função de η é:

$$a = \frac{1}{4\eta}, \quad b = \frac{1}{2} \quad \text{e} \quad c = \frac{\eta}{4},$$

e desta forma tem-se $q(y)$ e $q'(y)$ dadas, respectivamente, por:

$$q(y) = \frac{1}{4\eta}y^2 + \frac{1}{2}y + \frac{\eta}{4} \quad \text{e} \quad q'(y) = \frac{y}{2\eta} + \frac{1}{2}.$$

Note que:

$$\begin{aligned} q(-\eta) &= \frac{1}{4\eta}\eta^2 - \frac{1}{2}\eta + \frac{\eta}{4} = \frac{\eta}{4} + \frac{\eta}{4} - \frac{\eta}{2} = 0 = p(-\eta) \\ q(\eta) &= \frac{1}{4\eta}\eta^2 + \frac{1}{2}\eta + \frac{\eta}{4} = \frac{\eta}{4} + \frac{\eta}{4} + \frac{\eta}{2} = \eta = p(\eta) \\ q'(-\eta) &= -\frac{\eta}{2\eta} + \frac{1}{2} = 0 = p'(-\eta) \\ q'(\eta) &= \frac{\eta}{2\eta} + \frac{1}{2} = 1 = p'(\eta) \end{aligned}$$

portanto, q e q' são contínuas nos extremos dos intervalos.

A Figura, a seguir, representa a aproximação quadrática considerando $\eta = 0,5$ e $\eta = 0,1$.

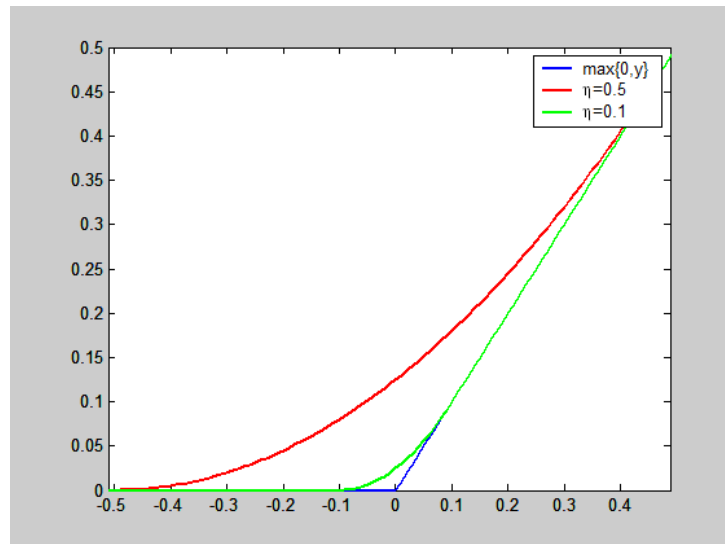


FIGURA 3: APROXIMAÇÃO QUADRÁTICA

Para todo $i = 1, \dots, m$, definimos a função $p_i : \mathbb{R} \rightarrow \mathbb{R}$ como:

$$p_i(h_i(x)) = \max \left\{ 0, \lambda_i^k + \frac{h_i(x)}{\gamma_k} \right\} = \begin{cases} 0, & \text{se } \lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq 0 \\ \lambda_i^k + \frac{h_i(x)}{\gamma_k}, & \text{se } \lambda_i^k + \frac{h_i(x)}{\gamma_k} > 0, \end{cases} \quad (6.4)$$

então, para cada $i = 1, \dots, m$, a aproximação quadrática (6.2) para p_i é dada por:

$$\bar{p}_i(h_i(x)) = \begin{cases} 0, & \text{se } \lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq -\eta \\ \frac{1}{4\eta} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right)^2 + \frac{1}{2} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) + \frac{\eta}{4}, & \text{se } -\eta < \lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq \eta \\ \lambda_i^k + \frac{h_i(x)}{\gamma_k}, & \text{se } \lambda_i^k + \frac{h_i(x)}{\gamma_k} > \eta \end{cases} \quad (6.5)$$

apesar de $p_i : \mathbb{R} \rightarrow \mathbb{R}$, quando se faz $p_i(h_i(x)) = (p_i \circ h_i)(x)$ tem-se uma composição de $p_i : \mathbb{R} \rightarrow \mathbb{R}$ com $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$, portanto $(p_i \circ h_i) : \mathbb{R}^n \rightarrow \mathbb{R}$. Desta forma, a derivada de $p_i \circ h_i$ é o vetor gradiente $\nabla(p_i \circ h_i)(x)$. Portanto, a derivada de $\bar{p}_i(h_i(x))$ dada em (6.5) terá a forma:

$$\bar{p}'_i(h_i(x)) = \begin{cases} 0, & \text{se } \lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq -\eta \\ \frac{1}{2\eta} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) \frac{\nabla h_i(x)}{\gamma_k} + \frac{\nabla h_i(x)}{2\gamma_k}, & \text{se } -\eta < \lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq \eta \\ \frac{\nabla h_i(x)}{\gamma_k}, & \text{se } \lambda_i^k + \frac{h_i(x)}{\gamma_k} > \eta. \end{cases} \quad (6.6)$$

Finalmente, tem-se a função $G(x)$ suavizada:

$$G(x) = \gamma_k[x - x^k] + \nabla F_x(x) + \sum_{i=1}^m \bar{p}_i(h_i(x)) \nabla h_i(x). \quad (6.7)$$

com $\bar{p}_i(h_i(x))$ dada por (6.5).

Desta forma, $G(x)$ dada por (6.7) se torna diferenciável. Além disso, em (6.5), se η for escolhido pequeno a função suavizada fica muito próxima da função verdadeira (ver Figura 3). O Método de Newton pode ser utilizado para resolvê-lo e é apresentado mais à frente.

Proposição 6.1 *Se $\eta \rightarrow 0$ então $\bar{p}_i(h_i(x))$ dada em (6.5) coincide com $p_i(h_i(x))$ dada em (6.4), para todo $i = 1, \dots, m$.*

Prova: *Basta mostrar que se $\eta \rightarrow 0$ o termo quadrático de $\bar{p}_i(h_i(x))$ desaparece.*

Assim,

$$\begin{aligned} & \frac{1}{4\eta} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right)^2 + \frac{1}{2} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) + \frac{\eta}{4} = \\ & \left[\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right] \left[\frac{1}{4\eta} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) + \frac{1}{2} \right] + \frac{\eta}{4} \end{aligned}$$

e, desde que, $-\eta < \lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq \eta$ então $\lambda_i^k + \frac{h_i(x)}{\gamma_k} \rightarrow 0$ quando $\eta \rightarrow 0$. Segue que $\left[\left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) \left(\frac{1}{4\eta} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) + \frac{1}{2} \right) + \frac{\eta}{4} \right] \rightarrow 0$ quando $\eta \rightarrow 0$ e o resultado segue. ■

6.2 MÉTODO DE NEWTON PARA RESOLUÇÃO DO SUB-PROBLEMA GERADO PELO ALGORITMO DE EQUILÍBRIO

Antes de apresentar o algoritmo, algumas explicações se fazem necessárias. Primeiramente, deve ser destacado que o Método de Newton será utilizado para resolver o subproblema gerado pelo Algoritmo 5.7, apresentado no capítulo anterior, para o Problema de Equilíbrio. Como observado anteriormente, a tarefa mais pesada do Algoritmo 5.7 é a resolução do subproblema gerado (5.7), sendo que este deve satisfazer a equação do erro (5.8). Resolver (5.7) satisfazendo (5.8) é equivalente a resolver o sistema não linear $G(x) = 0$, com $G(x)$ dada em (6.1) (ver (IUSEM; NASRI, 2007a)). Desta forma, se η for escolhido pequeno em (6.5) pela Proposição 6.1, ter-se-á mantidos os resultados já provados para o Problema de Equilíbrio.

Na próxima proposição, será provado que escolhendo adequadamente o parâmetro de regularização a Jacobiana de $G(x)$, dada em (6.1), será positiva definida. Para isso, e para os próximos algoritmos a serem apresentados, será necessária a seguinte hipótese:

Hipótese 6.2 *As funções $F_x : \mathbb{R}^n \rightarrow \mathbb{R}$, dada por $F_x = f(x, y)$ e $h_i : \mathbb{R}^m \rightarrow \mathbb{R}$, para $i = 1, \dots, m$, dadas na relação (5.2), são diferenciáveis até segunda ordem.*

Proposição 6.3 *Considere a Hipótese 6.2 satisfeita e $G(x)$ dada em (6.5) e (6.7). Então, a Jacobiana de $G(x)$ é dada por:*

$$JG(x) = \gamma_k I_{n \times n} + \nabla^2 F_x(x) + \sum_{i=1}^m \bar{p}'_i(h_i(x)) \nabla h_i(x) + \sum_{i=1}^m \bar{p}_i(h_i(x)) \nabla^2 h_i(x), \quad (6.8)$$

em que $\bar{p}'_i(h_i(x))$, dada em (6.6), é a derivada de $\bar{p}_i(h_i(x))$ e $I_{n \times n}$ é a matriz identidade. Além disso, para $i \in \{1, 2, \dots, m\}$ tem se:

- (a) Se $\lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq -\eta$ ou $\lambda_i^k + \frac{h_i(x)}{\gamma_k} > \eta$ então $JG(x) > 0$, para todo $x \in \mathbb{R}^n$.
- (b) Se $-\eta < \lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq \eta$ e $x \in \mathbb{R}^n$ estiver fixado, então existe \bar{k} tal que para todo $k > \bar{k}$ $JG(x) > 0$.

Prova: Desde que a Hipótese 6.2 é satisfeita, então (6.8) é obtida derivando se diretamente a expressão $G(x)$ em (6.7). Agora, se para algum $i \in \{1, 2, \dots, m\}$ ocorrer $\left(\lambda_i^k + \frac{h_i(x)}{\gamma_k}\right) \leq -\eta$, então, em (6.5) e (6.6) ter-se-á, respectivamente, $\bar{p}_i(h_i(x)) = 0$ e

$\bar{p}'_i(h_i(x)) = 0$. Segue de (6.8) que $JG(x) = \gamma_k I_{n \times n} + \nabla^2 F_x(x) > 0$ independente de x , pois $\gamma_k > 0$ e F_x é convexa por hipótese na definição do Problema de Equilíbrio. Por outro lado, se para algum $i \in \{1, 2, \dots, m\}$ ocorrer $\left(\lambda_i^k + \frac{h_i(x)}{\gamma_k}\right) > \eta > 0$, então, substituindo $\bar{p}_i(h_i(x))$, dada em (6.5), e $\bar{p}'_i(h_i(x))$, dada em (6.6), na expressão (6.8) ter-se-á:

$$JG(x) = \gamma_k I_{n \times n} + \nabla^2 F_x(x) + \sum_{i=1}^m \frac{\nabla h_i(x)}{\gamma_k} \nabla h_i(x)^T + \sum_{i=1}^m \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) \nabla^2 h_i(x).$$

Desta última expressão, segue que $JG(x) > 0$ independente de x , pois F_x e $h_i(x)$, para $i = 1, \dots, m$, são convexas por hipótese na definição do Problema de Equilíbrio. Portanto, o caso (a) da proposição está provado.

Para provar o caso (b) suponha que para algum $i \in \{1, 2, \dots, m\}$ ocorra $-\eta < \lambda_i^k + \frac{h_i(x)}{\gamma_k} \leq \eta$, então, substituindo $\bar{p}'_i(h_i(x))$, dada em (6.6), na expressão (6.8) ter-se-á:

$$\begin{aligned} JG(x) &= \gamma_k I_{n \times n} + \nabla^2 F_x(x) + \sum_{i=1}^m \left[\frac{1}{2\eta} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) \frac{\nabla h_i(x)}{\gamma_k} + \frac{\nabla h_i(x)}{2\gamma_k} \right] \nabla h_i(x)^T \\ &\quad + \sum_{i=1}^m \bar{p}_i(h_i(x)) \nabla^2 h_i(x). \end{aligned} \quad (6.9)$$

Note que cada termo em (6.9) é uma matriz $n \times n$. A matriz do último termo é positiva semidefinida, pois $h_i(x)$ é convexa para todo $i = 1, \dots, m$ e $\bar{p}_i(h_i(x)) \geq 0$ por construção. O segundo termo também forma uma matriz positiva semidefinida, pois F_x é convexa. O primeiro termo é uma matriz positiva definida. O problema fica no terceiro termo, não pode ser garantido nada pois irá depender do termo $\lambda_i^k + \frac{h_i(x)}{\gamma_k}$, o qual pode ser negativo por pertencer ao intervalo $(-\eta, \eta]$. Juntando o primeiro e terceiro termo, tem-se:

$$\begin{aligned} &\gamma_k I_{n \times n} + \sum_{i=1}^m \left[\frac{1}{2\eta} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) \frac{\nabla h_i(x)}{\gamma_k} + \frac{\nabla h_i(x)}{2\gamma_k} \right] \nabla h_i(x)^T = \\ &\gamma_k I_{n \times n} + \sum_{i=1}^m \left[\frac{1}{2\eta} \left(\lambda_i^k + \frac{h_i(x)}{\gamma_k} \right) \frac{\nabla h_i(x)}{\gamma_k} \nabla h_i(x)^T + \frac{\nabla h_i(x)}{2\gamma_k} \nabla h_i(x)^T \right]. \end{aligned}$$

Caso essa matriz não seja positiva definida aumenta-se o valor de γ_k e todos os elementos, das duas matrizes na soma, do último termo, ficam divididos por γ_k e, portanto, diminuem. Por outro lado, o primeiro termo faz aumentar a diagonal dessas matrizes, então existirá um $\bar{k} > 0$ tal que para todo $k \geq \bar{k}$ a soma dessa última expressão

se torna uma matriz diagonalmente dominante, pois x está fixado. Portanto, essa matriz se torna positiva definida e a prova está completa. ■

A seguir, será apresentado o algoritmo para o Problema de Equilíbrio combinado com o Método de Newton para resolver o subproblema gerado. O algoritmo é composto por dois laços principais que serão denominados de externo e interno. O laço interno é o que resolve $G(x) = 0$, utilizando o Método de Newton descrito nos parágrafos precedentes. Já o laço externo é o que faz a atualização dos demais parâmetros: multiplicadores de Lagrange dados na relação (5.9), teste de convergência e contador de iterações.

Algoritmo 6.1 PROBLEMA DE EQUILÍBRIO COMBINADO COM MÉTODO DE NEWTON PARA RESOLVER O SUBPROBLEMA

Dados $x^0 \in \mathbb{R}^n$, $\lambda^0 > 0$, $\eta > 0$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ e $\gamma^0 > 0$

$k = 0$ e $aux=1$,

Enquanto $aux=1$

Faça $\ell = 0$, e

$$z^\ell = x^k,$$

Avalie $G(z^\ell)$, dada em (6.5) e (6.7)

Enquanto $\|G(z^\ell)\| > \varepsilon_1$

Avalie $JG(z^\ell)$ ($JG(z^\ell)$ é a Jacobiana da $G(x)$ em z^ℓ)

Determine d^ℓ , (direção de Newton) solução de $JG(z^\ell)d^\ell = -G(z^\ell)$

Faça $z^{\ell+1} = z^\ell + d^\ell$,

$$\ell = \ell + 1, e$$

Avalie $G(z^\ell)$, dada em (6.5) e (6.7)

Fim enquanto

Faça $x^{k+1} = z^{\ell+1}$, e

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h(x^{k+1})}{\gamma^k} \right\}, \quad i = 1, \dots, m,$$

Se $\|(x^{k+1}, \lambda^{k+1}) - (x^k, \lambda^k)\| \leq \varepsilon_2$

$$aux=0,$$

Fim se

Faça $k = k + 1$,

Fim enquanto

Observe que foram inicializados dois parâmetros para a parada do algoritmo ε_1 e ε_2 , os quais recebem valores diferentes. O parâmetro ε_1 é utilizado como critério de parada para o subproblema interno, ou seja, para o Método de Newton. Já ε_2 é utilizado como critério de parada do algoritmo, ou seja, quando um ponto de equilíbrio é encontrado, o qual é mais rigoroso que ε_1 , pois foi constatado que resolver o subproblema por Newton, com uma precisão muito boa somente tornava o algoritmo mais caro, gastando mais iterações.

Foram utilizados dois contadores de iterações diferentes. Para o subproblema interno ℓ é o número de iterações que Newton utiliza para determinar a solução de $G(x) = 0$ com precisão ε_1 . Por outro lado, k é o número de iterações para o problema externo com precisão ε_2 . Desta forma, o total de iterações que o algoritmo utiliza para resolver o problema é dado pela soma das iterações internas e externas ($k + \ell$).

Vale destacar que $\lambda \geq 0$ é o multiplicador de Lagrange associado às restrições de desigualdade e $\gamma > 0$ é o parâmetro de regularização (no dual) ou o parâmetro de penalidade (no primal). Sobre a atualização do parâmetro $\gamma > 0$, é feita uma descrição na seção de implementação numérica.

No próximo teorema, será provado que a direção de Newton, gerada pelo Algoritmo 6.1, está bem definida, no sentido que sempre será possível resolver o sistema que determina esta direção.

Teorema 6.4 *Considere a Hipótese 6.2 satisfeita e as funções F_x e $h_i(x)$, para $i = 1, \dots, m$, definidas no Problema de Equilíbrio, convexas. Então, a cada iteração ℓ do algoritmo 6.1 a direção de Newton $d^\ell = -(JG(z^\ell))^{-1}G(z^\ell)$ está bem definida.*

Prova: *A prova segue da Proposição 6.3 anterior, bastando notar que o valor de z^ℓ , no cálculo de $d^\ell = -(JG(z^\ell))^{-1}G(z^\ell)$, está fixado em todas as iterações ℓ e, desde que, γ_k pode ser escolhido, contanto que seja finito, então a matriz $JG(z^\ell)$ é positiva definida e, portanto, é invertível. O sistema*

$$JG(z^\ell)d^\ell = -G(z^\ell),$$

tem solução. ■

6.3 MÉTODO SUBGRADIENTE PARA RESOLUÇÃO DO SUBPROBLEMA GERADO PELO ALGORITMO DE EQUILÍBRIO

Nesta seção, será apresentada a segunda metodologia para a resolução do Problema de Equilíbrio de Nash. A metodologia consiste no uso de Métodos de Subgradiente para resolver o subproblema gerado pelo Algoritmo de Equilíbrio Geral linearizado 5.7. Este subproblema é o sistema não linear e não diferenciável $G(x) = 0$ sendo $G(x)$ dada em (6.1). Para usar os Métodos subgradientes este sistema será transformado em um problema de minimização da forma:

$$\min\{f(x) : x \in \mathbb{R}^n\},$$

sendo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por:

$$f(x) = \frac{1}{2}\|G(x)\|^2, \quad (6.10)$$

em que $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é dada por (6.1).

Considere $g : \mathbb{R}^n \rightarrow \mathbb{R}$ a função definida por:

$$g(x) = \frac{1}{2}\|x\|^2. \quad (6.11)$$

Assim, a função $f(x)$ dada em (6.10) pode ser reescrita como a função composta:

$$f(x) = g(G(x)). \quad (6.12)$$

Como pode ser observado $g(x)$, dada em (6.11), é diferenciável e estritamente convexa. Mas o mesmo não acontece, em geral, com a função $f(x)$ dada em (6.12), pois $G(x)$ pode não ser convexa e é não diferenciável em pontos do seu domínio. Desta forma, não é possível aplicar diretamente os Métodos Subgradientes, descritos no Capítulo 3, desta tese. Por isso, será utilizado o mesmo procedimento da Seção 6.1 em que foi suavizado o termo não diferenciável.

Vale destacar que foi tentado utilizar resultados de subgradientes estendidos para funções não convexas apresentados por CLARKE (1983), e amplamente estudados por ROCKAFELLAR (1985). O que mais aproxima, do nosso problema, é o re-

sultado apresentado por estes dois autores para o cálculo do subgradiente de uma função composta da forma (6.12). Sob hipóteses, não muito restritivas, estes mostram que $\partial f(\bar{u}) \subset \nabla g(G(\bar{u}))\partial G(\bar{u})$ para um certo $\bar{u} \in \mathbb{R}^n$. Mas, isso não é o suficiente para o nosso caso, pois um elemento de $\nabla g(G(\bar{u}))\partial G(\bar{u})$ não é necessariamente um elemento de $\partial f(\bar{u})$.

Uma vez que não foi obtido sucesso com o cálculo de um subgradiente de $f(x)$ em (6.12) diretamente, apenas usando propriedades e informações das funções envolvidas, resolveu-se utilizar um procedimento parecido com o da Seção 6.1. Ou seja, suavizando a parte da $G(x)$ em (6.1), que é não diferenciável para que esta se torne diferenciável e então imitar o Método do Gradiente para calcular um subgradiente. Neste caso, em qualquer ponto $x \in \mathbb{R}^n$ ter-se-á em (6.12) :

$$\partial f(x) = \nabla f(x) = \nabla G(x)^T G(x).$$

Um passo do Método Gradiente para minimizar uma função do tipo (6.12) parte de um ponto já conhecido $x^k \in \mathbb{R}^n$ e determina o próximo ponto utilizando:

$$x^{k+1} = x^k - \alpha_k \nabla f(x^k).$$

sendo α_k o tamanho do passo, o qual é determinado por algum método de busca.

Os Métodos Subgradientes, utilizados nesta seção, foram descritos no Capítulo 3 e são duas as principais diferenças destes com o Método do Gradiente, a saber: (a) as direções geradas pelo Método do Gradiente são de descida, ao passo que as dos Métodos Subgradientes não são e (b) o cálculo do tamanho do passo (ver Seção 3.13.2), para a determinação do tamanho do passo nos Métodos Subgradientes.

É conveniente destacar um fato importante sobre esta metodologia. Como foi dito antes, a função $G(x)$, agora suavizada, é diferenciável, mas persiste o fato de que não é garantida sua convexidade. Foi mostrado na Proposição 6.3, que a Jacobiana de $G(x)$ é positiva definida em quase na totalidade de seu domínio, e na parte em que poderia não ser, é possível remediar, aumentando o parâmetro γ_k , dado na expressão de $G(x)$, o qual é um dos poucos que se tem liberdade de escolha, contanto que este seja finito. Desde que, uma das hipóteses dos Métodos Subgradientes é a convexidade da função que o define, não se pode garantir que os algoritmos irão convergir sempre.

Para um melhor acompanhamento do desempenho dos Métodos Subgradientes, será implementado, também, o Método do Gradiente com busca Barzilai Borwein (BARZILAI; BORWEIN, 1988) (ver Seção 3.6). Um fato bem conhecido é que o Método do

Gradiente possui uma velocidade de convergência baixa, o mesmo deverá acontecer com os Métodos Subgradientes mais básicos. Isto pode ser verificado nos testes numéricos realizados e apresentados no próximo capítulo.

A seguir, são apresentados os algoritmos para o Problema de Equilíbrio combinados com os Métodos Gradiente e Subgradientes para resolver o subproblema, descritos nos parágrafos precedentes. Como no Algoritmo de Newton 6.1, estes serão compostos por dois laços principais denominados de externo e interno. O laço interno é o que resolve $G(x) = 0$ utilizando, respectivamente, os Métodos Gradiente e Subgradientes. Já o laço externo é o que faz a atualização dos demais parâmetros: multiplicadores de Lagrange dados na relação (5.9), teste de convergência e contador de iterações.

Sob a Hipótese 6.2, o algoritmo para o Problema de Equilíbrio combinado com o Método do Gradiente para resolver o subproblema é dado por:

Algoritmo 6.2 *PROBLEMA DE EQUILÍBRIO COMBINADO COM GRADIENTE PARA RESOLVER O SUBPROBLEMA*

Dados $x^0 \in \mathbb{R}^n$, $\lambda^0 > 0$, $\alpha > 0$, $\eta > 0$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ e $\gamma^0 > 0$

$k = 0$ e $aux=1$,

$\alpha_0 = 1$ (tamanho do passo da primeira iteração).

Enquanto $aux=1$

Faça: $\ell = 0$, e

$$z^\ell = x^k,$$

Avalie: $G(z^\ell)$, dada em (6.5) e (6.7) e

$$JG(z^\ell) \quad (JG(z^\ell) \text{ é a Jacobiana da } G(x) \text{ em } z^\ell)$$

Enquanto $\|G(z^\ell)\| > \varepsilon_1$

Determine $\nabla f(z^\ell) = -JG(z^\ell)G(z^\ell)$, (gradiente de $f(x)$)

em z^ℓ , com f dada em (6.12))

$$\text{Faça } z^{\ell+1} = z^\ell + \alpha_\ell \nabla f(z^\ell),$$

Avalie: $G(z^{\ell+1})$, dada em (6.5) e (6.7)

$$JG(z^{\ell+1}) \quad (JG(z^{\ell+1}) \text{ é a Jacobiana da } G(x) \text{ em } z^{\ell+1})$$

Determine $\nabla f(z^{\ell+1}) = -JG(z^{\ell+1})G(z^{\ell+1})$, (gradiente

de $f(x)$ em $z^{\ell+1}$)

Calcule: $s_\ell = z^{\ell+1} - z^\ell$ e

$$y_\ell = \nabla f(z^{\ell+1}) - \nabla f(z^\ell)$$

Faça $\alpha_\ell = \frac{s_\ell^T s_\ell}{s_\ell^T y_\ell}$, (tamanho do passo por Barzilai Borwein)

Faça $\ell = \ell + 1$,

Fim enquanto

Faça: $x^{k+1} = z^{\ell+1}$, e

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h(x^{k+1})}{\gamma^k} \right\}, \quad i = 1, \dots, m,$$

Se $\|(x^{k+1}, \lambda^{k+1}) - (x^k, \lambda^k)\| \leq \varepsilon_2$

$aux=0$,

Fim se

Faça $k = k + 1$,

Fim enquanto

A seguir, serão apresentados os algoritmos para resolver o Problema de Equilíbrio de Nash com os Métodos de Subgradientes, apresentados no Capítulo 3, para resolver o subproblema $G(x) = 0$ sendo $G(x)$ dada em (6.7).

Primeiramente, será apresentado o algoritmo básico, no sentido que utiliza procedimentos simples para atualizar o tamanho do passo (ver Subseção 3.13.2). Este algoritmo é muito parecido com o algoritmo do Gradiente. Os demais algoritmos implementados foram: Ponto ótimo (quando o valor ótimo de $f(x)$ é conhecido - descrito em detalhes na Subseção 3.13.5), Método de Dilatação do Espaço na Direção do Subgradiente no Ponto Atual (ver Subseção 3.13.8) e Método de Elipsóides (descrito na Subseção 3.13.9). Para tanto será considerada a Hipótese 6.2.

Algoritmo 6.3 PROBLEMA DE EQUILÍBRIO COMBINADO COM O MÉTODO SUBGRADIENTE PARA RESOLVER O SUBPROBLEMA

Dados $x^0 \in \mathbb{R}^n$, $\lambda^0 > 0$, $\alpha > 0$, $a > 0$ (coeficiente de dilatação), $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ e $\gamma^0 > 0$

$k = 0$ e $aux=1$,

Enquanto $aux=1$

Faça: $\ell = 0$, e

$$z^\ell = x^k,$$

Avalie $G(z^\ell)$, dada em (6.5) e (6.7)

Enquanto $\|G(z^\ell)\| > \varepsilon_1$

Avalie $JG(z^\ell)$ ($JG(z^\ell)$ é a Jacobiana da $G(x)$ em z^ℓ)

Determine $\partial f(z^\ell) = -JG(z^\ell)^T G(z^\ell)$, (um subgradiente de $f(x)$)

em z^ℓ , com f dada em (6.12))

Escolher α_ℓ apropriado (conforme Seção 3.11.2)

Faça: $d_\ell = \frac{\partial f(z^\ell)}{\|\partial f(z^\ell)\|}$, e

$$z^{\ell+1} = z^\ell + \alpha_\ell d_\ell,$$

Avalie $G(z^{\ell+1})$, dada em (6.5) e (6.7)

Faça $\ell = \ell + 1$,

Fim enquanto

Faça: $x^{k+1} = z^{\ell+1}$, e

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h(x^{k+1})}{\gamma^k} \right\}, \quad i = 1, \dots, m,$$

Se $\|(x^{k+1}, \lambda^{k+1}) - (x^k, \lambda^k)\| \leq \varepsilon_2$

aux=0,

Fim se

Faça $k = k + 1$,

Fim enquanto

Observe que a diferença básica entre os Algoritmos 6.2 e 6.3 está na determinação da direção e no comprimento do passo.

Para o algoritmo de Subgradiente Ponto Ótimo, ou seja, quando f^* é conhecido, o esquema geral é o mesmo do Algoritmo 6.3. O que muda é o tamanho do passo que, neste caso, na iteração ℓ do subproblema é dado por $\alpha_\ell = \frac{f(z^\ell) - f^*}{\|\partial f(z^\ell)\|^2}$, sendo f dada em (6.12), $\partial f(z^\ell)$ um subgradiente de f em z^ℓ e f^* o valor ótimo de f . Para o nosso problema, este valor é $f^* = 0$.

Para o algoritmo de Dilatação do Espaço na Direção do Subgradiente no Ponto Atual (ver Algoritmo 3.17) é considerado o coeficiente de dilatação $a < 1$, com a alteração no cálculo do tamanho do passo, dado por $\alpha_\ell = \rho \frac{f(z^\ell) - m}{\|B_\ell^T s^k\|^2}$, e na direção $d_\ell = \frac{B_\ell^T s^k}{\|B_\ell^T s^k\|}$, sendo B_ℓ atualizada pela relação (3.39).

Por fim, o Algoritmo Elipsóides, ou seja, do Método Elipsóides, apresenta o mesmo esquema geral do Algoritmo 6.3, mudando a maneira de calcular a direção para $d_\ell = \frac{B_\ell^T s^k}{\|B_\ell^T s^k\|}$ e o tamanho do passo para $\alpha_\ell = \frac{R}{n+1} \frac{\beta^k}{\|B_\ell^T s^k\|}$, com $\beta = \frac{n}{\sqrt{n^2-1}}$. O coeficiente de dilatação para este método é determinado por $a_\ell = a = \sqrt{\frac{n-1}{n+1}}$ ($\forall k$), onde n é a dimensão do espaço (ver Algoritmo 3.18).

7 RESULTADOS NUMÉRICOS

Neste capítulo, serão apresentados alguns exemplos de aplicações encontrados na literatura da área, para servir de comparação aos algoritmos implementados, baseados na metodologia proposta. Foram implementados os Algoritmos 6.1, 6.2 e o Algoritmo 6.3 combinado com os métodos: Método Subgradiente Básico (ver Algoritmo 3.15), Ponto Ótimo (quando f^* é conhecido) (ver a Subseção 3.13.5), Método de Elipsoides (ver Algoritmo 3.18) e Método Dilatação do Espaço na Direção do Subgradiente no Ponto Atual (ver Algoritmo 3.17) para resolver o subproblema. O primeiro exemplo será apresentado de forma detalhada conforme a metodologia proposta. Já para os demais serão apresentados o problema e os resultados numéricos. Os algoritmos para solução dos exemplos apresentados a seguir foram implementados no programa Matlab, versão 6.0.

Nas implementações dos exemplos a seguir foi considerado $\gamma^k = \frac{1}{250}$ para o algoritmo do Problema de Equilíbrio combinado com os Métodos de Newton e Dilatação de Espaço na Direção do Subgradiente no Ponto Atual. Para os demais métodos utilizou-se $\gamma^k = 1$. O coeficiente de dilatação para o Método de Dilatação de Espaço na Direção do Subgradiente no Ponto Atual $a = 0, 1$ e para o Método Elipsoides $a = 1$.

Para todos os métodos implementados foi considerado $\eta = 0, 1$, o critério de parada para o subproblema interno $\|G(z^\ell)\| < \varepsilon_1$, onde $\varepsilon_1 = 10^{-4}$ e o critério de parada para o Problema de Equilíbrio $\|(x^{k+1}, \lambda^{k+1}) - (x^k, \lambda^k)\| \leq \varepsilon_2$, onde $\varepsilon_2 = 10^{-4}$. Embora na seção 6.2 foi afirmado que ε_1 e ε_2 recebem valores diferentes, nas implementações foi considerado $\varepsilon_1 = \varepsilon_2$. A bifunção utilizada para os Exemplos 7.1, 7.2 e 7.4 é dada por:

$$f(x, y) = \sum_{j=1}^N [\theta_j(y^j, x^{-j}) - \theta_j(x^j, x^{-j})]. \quad (7.1)$$

7.1 PROBLEMAS MENORES

São considerados problemas menores os exemplos com o número de jogadores menor que dez.

Exemplo 7.1 *POLUIÇÃO DE UMA BACIA HIDROGRÁFICA (River Basin Pollution)* (KRAWCZYK; URYASEV, 1999)

Neste exemplo, considere três jogadores $j = 1, 2, 3$ localizados ao longo de um rio. Cada jogador está envolvido em uma atividade econômica (produção de pasta de papel, por exemplo) em um nível x^j escolhido, mas os jogadores devem atender às condições ambientais estabelecidas pela autoridade local.

Os poluentes podem ser expelidos para o rio, onde se dispersam. Duas estações de monitoramento $l = 1, 2$ estão localizadas ao longo do rio, em que as autoridades estabeleceram níveis máximos de concentração de poluentes.

A seguir, é dada a tabela com as constantes de poluição de uma bacia hidrográfica, em que c_{1j} e c_{2j} são os coeficientes de poluição, u_{jl} os coeficientes de decaimento e transporte e e_j o coeficiente de emissão.

Jogador j	c_{1j}	c_{2j}	e_j	u_{j1}	u_{j2}
1	0.10	0.01	0.50	6,5	4,583
2	0.12	0.05	0.25	5,0	6,250
3	0.15	0.01	0.75	5,5	3,750

TABELA 1: CONSTANTES DE POLUIÇÃO

A receita para o jogador j é:

$$R_j(x) = [d_1 - d_2(x_1 + x_2 + x_3)]x_j,$$

em que as constantes econômicas d_1 e d_2 determinam a lei de demanda inversa e são dadas por 3.0 e 0.01 respectivamente, com custo:

$$F_j(x) = (c_{1j} + c_{2j}x_j)x_j.$$

Assim, o lucro líquido do jogador j , ou função objetivo do jogador j , é dado por:

$$\theta_j(x) = [d_1 - d_2(x_1 + x_2 + x_3) - c_{1j} - c_{2j}x_j]x_j = d_1x_j - d_2x_j(x_1 + x_2 + x_3) - c_{1j}x_j - c_{2j}x_j^2. \quad (7.2)$$

As restrições de emissão de poluentes impostas pelas autoridades são:

$$q_l(x) = \sum_{j=1}^3 u_{jl}e_jx_j \leq K_l, \quad l = 1, 2, \quad (7.3)$$

em que $K_l = 100$ para $l = 1, 2$.

Desta forma, o modelo matemático para o problema de poluição de uma bacia hidrográfica é dado por:

$$\begin{aligned} & \min \theta_j(x); \\ & \text{s.a } q_l(x) \leq 0, \quad j=1,2,3 \quad \text{e } l=1,2. \end{aligned}$$

Substituindo o valores nas equações (7.2) e (7.3), tem-se:

$$\begin{aligned} \theta_j(x) &= c_{2j}x_j^2 + d_2x_j(x_1 + x_2 + x_3) - b_jx_j = \\ &= c_{2j}x_j^2 + 0,01x_j(x_1 + x_2 + x_3) - b_jx_j \quad (j = 1, 2, 3), \end{aligned}$$

com $b_j = d_1 - c_{1j} = (2, 90; 2, 88; 2, 85)$ e as restrições dadas por:

$$q_l(x) = \begin{cases} 3, 25x_1 + 1, 25x_2 + 4, 125x_3 \leq 100 \\ 2, 291x_1 + 1, 5625x_2 + 2, 8125x_3 \leq 100. \end{cases}$$

Para obter o sistema $G(x) = 0$ dado por (6.1), são consideradas as seguintes etapas:

1ª Etapa: Escrever a bifunção (7.1). Neste caso:

$$\begin{aligned} f(x, y) &= \{[c_{21}y_1^2 + 0,01y_1(y_1 + x_2 + x_3) - b_1y_1] - [c_{21}x_1^2 + 0,01x_1(x_1 + x_2 + x_3) - b_1x_1]\} + \\ & \{[c_{22}y_2^2 + 0,01y_2(x_1 + y_2 + x_3) - b_2y_2] - [c_{22}x_2^2 + 0,01x_2(x_1 + x_2 + x_3) - b_2x_2]\} + \{[c_{23}y_3^2 + \\ & 0,01y_3(x_1 + x_2 + y_3) - b_3y_3] - [c_{23}x_3^2 + 0,01x_3(x_1 + x_2 + x_3) - b_3x_3]\}. \end{aligned}$$

2ª Etapa: Determinar o gradiente da bifunção $f(x, y)$ em relação a y , conforme segue abaixo:

$$\nabla F_x(y) = \nabla_y f(x, y) = [2c_{21}y_1 + 0, 02y_1 + 0, 01x_2 + 0, 01x_3 - b_1; 2c_{22}y_2 + 0, 01x_1 + 0, 02y_2 + 0, 01x_3 - b_2; 2c_{23}y_3 + 0, 01x_1 + 0, 01x_2 + 0, 02y_3 - b_3]^T.$$

3ª Etapa: Calcular $\nabla F_x(x)$, substituindo y por x em $\nabla F_x(y)$, ou seja:

$$\nabla F_x(x) = [2c_{21}x_1 + 0, 02x_1 + 0, 01x_2 + 0, 01x_3 - b_1; 2c_{22}x_2 + 0, 01x_1 + 0, 02x_2 + 0, 01x_3 - b_2; 2c_{23}x_3 + 0, 01x_1 + 0, 01x_2 + 0, 02x_3 - b_3]^T.$$

Portanto, com os resultados obtidos nas etapas acima, pode-se escrever o sistema $G(x)$ dado em (6.1). Após obter $G(x) = 0$ é realizado uma aproximação quadrática no intervalo $[-\eta, \eta]$ com o objetivo de suavizar o sistema, conforme (6.7).

Na implementação deste problema, foi considerado para todos os métodos $x^0 = [5, 9, 3]^T$, $\lambda^0 = [1, \frac{3}{2}]^T$, $\gamma^0 = 2$ e o número máximo de iterações no laço interno que resolve $G(x)$ igual a 1000.

Na tabela a seguir, pode ser verificado o resultado obtido em cada método utilizado. Para facilitar a visualização por meio de tabela, será abreviado o título do Método Dilatação de Espaço na Direção do Subgradiente no Ponto Atual para Método de Dilatação de Espaço e o título Problema de Equilíbrio Combinado com Método de Newton para Resolver o Subproblema para Método de Newton, esta mesma abreviação será utilizada nos demais resultados apresentados neste capítulo.

Na Tabela 2, a seguir, a coluna 1 identifica o nome dos algoritmos utilizados, a coluna 2, a solução aproximada x^* do Exemplo 7.1, a coluna 3 representa a norma do sistema $G(x)$ na última iteração interna, a coluna 4, o número de iterações externas e na coluna 5, o total de iterações realizadas pelo algoritmo.

Métodos	x^*	$\ G(x)\ $	Ite. Ext.	Total Iter.
Método de Cauchy	$[21.2088; 16.6970; 2.4446]^T$	9.9644×10^{-5}	277	257358
Subgradiente Básico	$[21.1487; 16.0276; 2.7229]^T$	9.8729×10^{-5}	155	66922
Ponto Ótimo	$[21.1397; 16.0279; 2.7300]^T$	9.9665×10^{-5}	274	73146
Método de Elipsoides	$[21.1390; 16.0280; 2.7305]^T$	8.8532×10^{-5}	267	39742
Dilatação do Espaço	$[21.1428; 16.0279; 2.7276]^T$	8.4908×10^{-5}	06	55
Newton	$[21.1430; 16.0279; 2.7274]^T$	8.4132×10^{-5}	06	19

TABELA 2: RESULTADO NUMÉRICO DO EXEMPLO 7.1

Este mesmo exemplo foi apresentado por SANTOS e SCHEIMBERG (2011), ao aplicar o Método Subgradiente Projetado Inexato (IPSM), que obteve a solução ótima

com 7 iterações, sendo $x^* = [21.149; 16.028; 2.722]^T$ e também utilizado por MATIOLI, SOSA e YUAN (2012), que introduzem um algoritmo baseado em um sistema de projeção sucessiva, em que a solução ótima é encontrada com 9 iterações.

A metodologia proposta neste exemplo apresentou-se eficiente, porém alguns Métodos Subgradientes conforme a Tabela 2, exigiram um número considerável de iterações internas e externas. Os Métodos Dilatação do Espaço e o Método de Newton apresentaram uma solução ótima com menor número de iterações internas e externas.

A seguir, será apresentado o resultado da implementação do Exemplo 2.1, extraído do artigo de FACCHINEI e KANZOW (2010a), p.178.

Exemplo 7.2 *Considere um jogo com dois jogadores, $N = 2$ com $n_1 = 1$ e $n_2 = 1$, assim, cada jogador controla uma variável; por simplicidade define-se $x_1^1 = x^1$ e $x_1^2 = x^2$. Assumindo que os problemas dos jogadores são:*

$$\min_{x^1} (x^1 - 1)^2; \quad s.a \quad x^1 + x^2 \leq 1 \qquad \min_{x^2} (x^2 - \frac{1}{2})^2; \quad s.a \quad x^1 + x^2 \leq 1.$$

Pelo enunciado do problema, pode ser verificado que possui somente uma restrição.

Para o Exemplo 7.2, foi considerado $x^0 = [0, 0]^T$, $\lambda^0 = 1$, $\gamma^0 = 2$ e o número máximo de iterações no laço interno que resolve $G(x) = 0$ igual a 1000.

Na Tabela 3, a seguir, a coluna 1 identifica o nome dos algoritmos utilizados, a coluna 2, a solução aproximada x^* do problema 7.2, a coluna 3, representa a norma do sistema $G(x)$, na última iteração interna, a coluna 4, o número de iterações externas, e na coluna 5, o total de iterações realizadas pelo algoritmo.

Métodos	x^*	$\ G(x)\ $	Iter. Ext.	Total Iter.
Método de Cauchy	$[0.7500; 0.2500]^T$	9.9644×10^{-5}	12	72
Subgradiente Básico	$[0.7500; 0.2500]^T$	1.1028×10^{-5}	48	13402
Ponto Ótimo	$[0.7500; 0.2500]^T$	5.8968×10^{-5}	11	144
Elipsoides	$[0.7500; 0.2500]^T$	7.0114×10^{-5}	17	1116
Dilatação do Espaço	$[0.7500; 0.2500]^T$	1.4356×10^{-6}	4	17
Newton	$[0.7500; 0.2500]^T$	1.1069×10^{-14}	4	10

TABELA 3: RESULTADO NUMÉRICO DO EXEMPLO 7.2

Neste exemplo, pode ser verificado que os algoritmos mais eficientes, ou seja, que resolveram em um número menor de iterações foram: Método de Dilatação do Espaço e o Método de Newton.

Exemplo 7.3 *Considere os dois problemas de equilíbrio apresentados em (TRAN; DUNG; NGUYEN, 2006), onde:*

$$X = \{x \in \mathbb{R}^5 : \sum_{i=1}^5 x_i \geq -1, \quad -5 \leq x_i \leq 5, \quad i = 1, \dots, 5\},$$

e a bifunção dada por:

$$f(x, y) = \langle Px + Qy + q, y - x \rangle.$$

As matrizes P , Q e o vetor q são definidos por:

$$P_1 = \begin{bmatrix} 3.1 & 2 & 0 & 0 & 0 \\ 2 & 3.6 & 0 & 0 & 0 \\ 0 & 0 & 3.5 & 2 & 0 \\ 0 & 0 & 2 & 3.3 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 3.1 & 2 & 0 & 0 & 0 \\ 2 & 3.6 & 0 & 0 & 0 \\ 0 & 0 & 3.5 & 2 & 0 \\ 0 & 0 & 2 & 3.3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$Q = \begin{bmatrix} 1.6 & 1 & 0 & 0 & 0 \\ 1 & 1.6 & 0 & 0 & 0 \\ 0 & 0 & 1.5 & 1 & 0 \\ 0 & 0 & 1 & 1.5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

e

$$q = [1, -2, -1, 2, -1]^T.$$

Na solução do Exemplo 7.3 as restrições foram consideradas da seguinte forma:

$$h_i(x) = \{-1 - \sum x_i \leq 0; x_i - 5 \leq 0; -x_i - 5 \leq 0\}, \quad (i = 1, \dots, 5).$$

Para o exemplo foi utilizado $\gamma^0 = 2$, $\lambda^0 = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1]^T$, $x^0 = [1, 3, 1, 1, 2]^T$, e o número máximo de iterações no laço interno que resolve $G(x) = 0$ igual a 1000.

A seguir, nas Tabelas 4 e 5 são apresentadas as soluções aproximadas do Exemplo 7.3 considerando P_1 e P_2 respectivamente. A coluna 1 identifica o nome dos algoritmos utilizados, a coluna 2, a solução aproximada x^* , a coluna 3, representa a norma do sistema $G(x) = 0$, na última iteração interna, a coluna 4, o número de iterações externas, e na coluna 5, o total de iterações realizadas pelo algoritmo.

Métodos	x^*	$\ G(x)\ $	Iter. Ext.	Total Iter.
Cauchy	$x^* = [-0.7254; 0.8031; 0.7200; -0.8667; 0.2500]^T$	9.3292×10^{-5}	10	147
Subgradiente Básico*	$x^* = [-0.7264; 0.8020; 0.7201; -0.8666; 0.2500]^T$	0.0113	11	7712
Ponto Ótimo	$x^* = [-0.7254; 0.8031; 0.7200; -0.8666; 0.2500]^T$	7.6990×10^{-5}	10	409
Elipsóides	$x^* = [-0.7254; 0.8031; 0.7199; -0.8666; 0.2499]^T$	8.6207×10^{-5}	18	7618
Dilatação do Espaço	$x^* = [-0.7254; 0.8031; 0.7200; -0.8667; 0.2500]^T$	2.8998×10^{-5}	04	32
Newton	$x^* = [-0.7254; 0.8031; 0.7200; -0.8667; 0.2500]^T$	5.6816×10^{-6}	04	12

TABELA 4: Solução Ótima para P1

Métodos	x^*	$\ G(x)\ $	Iter. Ext.	Total Iter.
Cauchy	$x^* = [-0.7254; 0.8031; 0.7200; -0.8667; 0.2000]^T$	2.0855×10^{-5}	11	152
Subgradiente Básico*	$x^* = [-0.7264; 0.8020; 0.7201; -0.8666; 0.2000]^T$	6.6589×10^{-5}	20	14021
Ponto Ótimo	$x^* = [-0.7254; 0.8031; 0.7200; -0.8666; 0.2000]^T$	5.6867×10^{-5}	10	399
Elipsóides	$x^* = [-0.7254; 0.8031; 0.7200; -0.8666; 0.2000]^T$	9.4885×10^{-5}	32	13428
Dilatação do Espaço	$x^* = [-0.7254; 0.8031; 0.7200; -0.8667; 0.2000]^T$	5.0912×10^{-5}	04	33
Newton	$x^* = [-0.7254; 0.8031; 0.7200; -0.8667; 0.2000]^T$	5.4367×10^{-6}	04	12

TABELA 5: Solução Ótima para P2

* No Método Subgradiente Básico, o número máximo de iterações internas é dada por ($\ell = 700$).

Para solução desse Exemplo TRAN, DUNG e NGUYEN (2006) utilizaram dois algoritmos Extragradiante (2a e 2b). O primeiro problema foi resolvido em 10 iterações para o algoritmo 2a, o segundo problema foi resolvido em 10 iterações no algoritmo 2a e em 21 iterações no algoritmo 2b. Este mesmo problema foi apresentado por SANTOS e SCHEIMBERG (2011), ao aplicar o Método Subgradiente Projetado Inexato (IPSM) que obteve a solução ótima em 10 iterações.

7.2 PROBLEMAS MAIORES

São considerados problemas maiores os exemplos que envolvem um número igual ou superior a dez jogadores. Nesta classe de problemas os métodos que apresentaram-se eficientes foram Dilatação de Espaço na Direção do Subgradiente no Ponto Atual e Newton com Aproximação Quadrática do Termo não Diferenciável, os demais métodos não convergiram para solução aproximada.

Exemplo 7.4 *Baseado no problema POLUIÇÃO DE UMA BACIA HIDROGRÁFICA, foi implementado um novo problema em que a função objetivo do jogador j é dado por:*

$$\theta_j = u_j x_j^2 + \beta x_j (\sum_{k=1}^n x_k) - v_j x_j, \quad j = 1, \dots, n$$

e a restrição de emissão dos poluentes é dada por:

$$A_{m \times n} x_j \leq b_{m \times 1},$$

onde $\beta = 0.01$, m é o número de restrições e n o número de jogadores. A matriz $A_{m \times n}$, o vetor $b_{m \times 1}$ e os vetores u_j, v_j foram gerados randomicamente.

O problema então consiste em:

$$\begin{aligned} \min \theta_j &= u_j x_j^2 + \beta x_j (\sum_{k=1}^n x_k) - v_j x_j, \\ \text{s.a. } A_{m \times n} x_j - b_{m \times 1} &\leq 0, \quad \text{com } j = 1, \dots, n. \end{aligned}$$

Para resolução do Exemplo 7.4 foi considerado $\gamma^0 = 2$, o ponto inicial $x^0 = [1, 1, \dots, 1]^T$, os multiplicadores de Lagrange λ^0 , as matrizes $A_{m \times n}$ e $b_{m \times 1}$ e os vetores u e v são gerados randomicamente.

Na Tabela 6, a seguir, na coluna 1 está identificado o nome dos algoritmos utilizados, as colunas 2 e 3 representam respectivamente, o número de jogadores e de restrições,

a coluna 4 a norma do sistema $G(x)$ na última iteração interna, na coluna 5 o número de iterações externas, e na coluna 5, o total de iterações realizadas.

Métodos	n	m	$\ G(x)\ $	nº Iter. ext.	Total de Iter.
Método Dilatação de Espaço	10	05	6.7408×10^{-5}	03	41
Método de Newton	10	05	3.2976×10^{-15}	03	10
Método Dilatação do Espaço	30	20	8.2773×10^{-5}	04	293
Método de Newton	30	20	3.3823×10^{-9}	04	18
Método Dilatação do Espaço	50	10	5.9428×10^{-5}	05	417
Método de Newton	50	10	7.9643×10^{-10}	05	19
Método Dilatação de Espaço	100	10	9.4900×10^{-5}	05	934
Método Newton	100	10	1.9619×10^{-5}	05	17
Metodo Dilatação do Espaço	200	20	9.3370×10^{-5}	11	8240
Método Newton	200	20	9.3042×10^{-9}	06	37

TABELA 6: RESULTADO NUMÉRICO DO EXEMPLO 7.4

Exemplo 7.5 *Considere o problema:*

$$X = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i \geq -1, \quad -5 \leq x_i \leq 5, \quad i = 1, \dots, n\},$$

e a bifunção dada por:

$$f(x, y) = \langle Px + Qy + q, y - x \rangle.$$

As matrizes $P_{n \times n}$, $Q_{n \times n}$ e o vetor $q_{n \times 1}$ são gerados de maneira randômica, mas com a mesma estrutura do Exemplo 7.1 e as restrições do problema são da seguinte forma:

$$X = \{x \in \mathbb{R}^n : -\sum_{i=1}^n x_i - 1 \leq 0; x_i - 5 \leq 0 \quad e \quad -x_i - 5 \leq 0\},$$

logo o número de restrições passa a ser $m = 2n + 1$.

A seguir são apresentados os resultados na implementação do Exemplo 7.5. Na Tabela 7, a seguir, na coluna 1 está identificado o nome dos algoritmos utilizados, as colunas 2 e 3 representam respectivamente, o número de jogadores e de restrições, a coluna 4 a norma do sistema $G(x)$ na última iteração interna, na coluna 5 o número de iterações externas, e na coluna 5, o total de iterações realizadas.

Métodos	n	m	$\ G(x)\ $	nº Iter. ext.	Total de Iter.
Método Dilatação de Espaço	10	21	8.6566×10^{-5}	04	71
Método de Newton	10	21	1.5983×10^{-13}	04	15
Método Dilatação do Espaço	20	41	9.9301×10^{-5}	04	193
Método de Newton	20	41	1.6585×10^{-13}	04	17
Método Dilatação de Espaço	50	101	6.6665×10^{-5}	06	612
Método de Newton	50	101	6.5028×10^{-13}	5	25
Método Dilatação do Espaço	100	201	9.7635×10^{-5}	05	874
Método de Newton	100	201	8.2030×10^{-6}	5	19
Método Dilatação de Espaço	150	301	8.7668×10^{-5}	04	1193
Método de Newton	150	301	3.5226×10^{-12}	04	17

TABELA 7: RESULTADO NUMÉRICO DO EXEMPLO 7.5

Conclusão sobre os testes numéricos:

Pelos testes realizados e os resultados mostrados, nesta seção, demonstram que a metodologia proposta e os algoritmos implementados neste trabalho tiveram um excelente desempenho, no sentido de convergirem para solução ótima dos exemplos apresentados.

Para problemas envolvendo um número menor que dez jogadores todos os algoritmos 6.2 e o Algoritmo 6.3 combinado com os algoritmos Subgradiente Básico, Elipsoides e Ponto Ótimo convergiram para solução do ótima dos exemplos, porém exigiram um número considerável de iterações internas e externas. Já o Algoritmo 6.1 e o Algoritmo 6.3, combinado com Método de Dilatação do Espaço na Direção do Subgradiente no Ponto Atual, resolveram os problemas menores com o mesmo número de iterações externas, porém o Método de Newton resolveu com menos iterações internas mostrando-se mais eficiente.

Para os problemas maiores, ou seja, problemas envolvendo um número igual ou superior a dez jogadores, os Algoritmos 6.1 e 6.3 combinado Dilatação do Espaço na Direção do Subgradiente no Ponto Atual, para solução do subproblema, convergiram para a solução ótima, sendo o Algoritmo 6.1 o mais eficiente ao resolver os exemplos em um número menor de iterações externas e internas. Já os demais algoritmos não convergiram para a solução ótima dos Exemplos 7.4 e 7.5.

Como pode ser visto pelos exemplos de aplicações do Problema de Equilíbrio de Nash, independente do número de jogadores os algoritmos Problema de Equilíbrio Combinado com Newton, para resolver o subproblema, converge para a solução ótima num menor número de iterações.

8 CONSIDERAÇÕES FINAIS

O objetivo desta tese foi propor algoritmos para resolução do Problema de Equilíbrio de Nash, pois existem mais pesquisas sobre a existência de soluções, do que boas implementações (FACCHINEI; KANZOW, 2010a). Baseado no artigo, apresentado por IUSEM e NASRI (2007a), em que foi introduzido o Método de Lagrangeano Aumentado, para a resolução de um Problema de Equilíbrio geral, mais precisamente no algoritmo IA-LEM e no resultado (65) deste artigo, como o Problema de Equilíbrio de Nash é um caso particular do Problema de Equilíbrio, são apresentados nesta pesquisa dois novos métodos para a solução do Problema de Equilíbrio de Nash, sendo eles: Método de Newton e Método Subgradiente para resolução do subproblema gerado pelo algoritmo de equilíbrio.

A metodologia proposta foi resolver o Problema de Equilíbrio de Nash baseado na resolução do sistema não linear $G(x) = 0$ apresentado em IUSEM e NASRI (2007a). A primeira dificuldade encontrada foi que o último termo do lado direito da equação (6.1) é contínuo, mas não diferenciável. Para isto, o primeiro passo foi a suavização de $G(x)$ por meio da aproximação quadrática do termo não diferenciável. Desta forma $G(x)$ dada em (6.7) passa a ser diferenciável. Após a suavização de $G(x)$, foram aplicados os Métodos de Newton e Subgradientes para solução do sistema $G(x) = 0$.

Nos métodos propostos, os algoritmos consistem em resolver o Problema de Equilíbrio em dois laços principais denominados de externos e internos. No laço interno, resolve-se o sistema $G(x) = 0$, por isto, os algoritmos são uma combinação do Problema de Equilíbrio com Método de Newton (ver Algoritmo 6.1), e com Métodos Subgradientes (ver Algoritmo 6.3) para solução do subproblema. Os algoritmos foram implementados e aplicados em problemas já conhecidos na literatura e mostraram-se eficientes ao convergirem para solução dos problemas. Dentre eles pode ser verificado que o Problema de Equilíbrio combinado com o Método Dilatação do Espaço na Direção do Subgradiente no Ponto Atual (ver Algoritmo 3.17), e combinado com o Método de Newton (ver Algoritmo 6.1), resolveram os problemas em número menor de iterações externas e internas. Estes métodos também mostraram ser mais eficientes ao resolverem os problemas maio-

res, ou seja, com o número de jogadores igual ou superior a dez. Dentre os algoritmos implementados, o que se mostrou mais instável foi o algoritmo para solução do Problema de Equilíbrio combinado com o Método Subgradiente Básico (ver Algoritmo 3.15), como pode ser visto nos resultados numéricos apresentados.

A metodologia proposta mostrou-se promissora, como pode ser visto nos testes numéricos apresentados no Capítulo 7. Portanto dois novos algoritmos são oferecidos como opções a mais para os interessados na solução do Problema de Equilíbrio de Nash (NEP).

Como sugestão para continuação desta pesquisa, propõe-se:

- Usar outros métodos para resolver o subproblema interno $G(x) = 0$, por exemplo o Método BFGS;
- Usar, no Método Lagrangeano Aumentado, penalidades que possuem derivadas de ordem maior, tal que, as funções que formam o sistema $G(x) = 0$ sejam diferenciáveis.

Referências

- BARZILAI, J.; BORWEIN, J. M. Two point step size gradient method. **IMA Journal of Numerical Analysis**, v. 8, p. 141–148, 1988.
- BOYD, S.; MUTAPCIC, A. Subgradient methods. **Notes for EE364b**, Winter 2006-07, abril 2008.
- CHUQUIPOMA, J. A. D. Inequações variacionais teoria e aplicações. **ERMAC**, p. 36–64, nov. 2010.
- CLARKE, F. H. Optimization and nonsmooth analysis. **Wiley-Interscience**, New York, 1983.
- EUSTÁQUIO, R. G. **Condições de Otimalidade e de Qualificação para Problemas de Programação não Linear**. 126 p. Dissertação (Mestrado) — Setores de Ciências Exatas e Tecnologia, Universidade Federal do Paraná, Curitiba, 2007.
- FACCHINEI, F.; KANZOW, C. Penalty methods for the solution of generalized Nash equilibrium problems. **SIAM J. OPTIM**, Society for Industrial and Applied Mathematics, v. 20, n° 5, p. 2228–2253, 2010.
- FACCHINEI, F.; KANZOW, C. Generalized Nash equilibrium problems. **Springer**, Ann Oper Res 175, p. 177–211, 2010a.
- FRIEDLANDER, A. **Elementos de Programação Não-linear**. Campinas - SP: UNICAMP, 1994.
- GONZAGA, C. C. **Um Curso de Programação não Linear**. Florianópolis - SC: UFSC, 2004.
- HESTENES, M. Multiplier and gradient methods. **Journal Optimization Theory and Applications**, v. 4, p. 303–320, 1969.
- HEUSINGER, A. **Numerical Methods for the Solution of the Generalized Nash Equilibrium Problems**. Tese (Tese de PhD) — University of Würzburg, ago. 2009.
- HEUSINGER, A.; KANZOW, C. SC^1 optimization reformulations of the generalized Nash equilibrium problem. **Optimization Methods and Software**, Würzburg, v. 23, n° 6, p. 953–973, out. 2008.
- HEUSINGER, A.; KANZOW, C. Optimization reformulations of the generalized Nash equilibrium using Nikaido-Isoda-type functions. **Comput Optim Appl**, Würzburg, v. 43, p. 353–377, 2009.
- IUSEM, A. N.; NASRI, M. Augmented lagrangian methods for equilibrium problems. 2007a.

IUSEM, A. N.; NASRI, M. Inexact proximal point methods for equilibrium problems in Banach spaces. **Numerical Functional Analysis and Optimization**, v. 28, n. 11 e 12, p. 1279–1308, 2007b.

IUSEM, A. N.; SOSA, W. A proximal point method for equilibrium problems in Hilbert spaces. p. 17, jun. 2007.

JOHN. **John Forbes Nash Junior**: Site. 2011. Disponível em: [http : //pt.wikipedia.org/wiki/JohnForbesNash](http://pt.wikipedia.org/wiki/JohnForbesNash)>. Acesso em: 26 abril 2011.

JUSSIANI, L. F. **Desempenho do Método de Lagrangeano Aumentado com Penalidade Quadrática**. 96 p. Dissertação (Mestrado em Ciências) — Curso de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Curitiba, 2004.

KRAWCZYK, J.; URYASEV, S. Relaxation algorithms to find Nash equilibria whit economic applications. **Baltzer Science Publishers BV**, Environmental Modeling and Assessment 5(2000), p. 63–73, set. 1999.

MARTINEZ, J. M. **Otimização Prática Usando o Lagrangeano Aumentado**. Departamento de Matemática Aplicada: IMECC - UNICAMP, 2009.

MARTINEZ, J. M.; SANTOS, S. A. **Métodos computacionais de otimização**. Departamento de Matemática Aplicada: IMECC - UNICAMP, 1998.

MATIOLI, C. L.; SOSA, W.; YUAN, J. A numerical algorithm for solutions of a generalized Nash equilibrium problem. **Computational Optimization and Applications**, v. 52, p. 281–292, 2012.

MEDEIROS, L. A. S. **Acelerando o método de Levenberg-Marquardt para a Minimização da Soma de Quadrados de Funções com Restrições em Caixa**. Tese (Doutorado) — Universidade Estadual de Campinas, Campinas - SP, 2008.

MINOUX, M. **Mathematical Programming Theory and Algorithms**. [S.l.]: John Wiley and Sons, 1986.

NABETANI, K. **Variational Inequality Approaches to Generalized Nash Equilibrium Problems**. 26 p. Dissertação (Mestrado) — Kyoto University, Department of Applied Mathematics and Physics Graduate School of Informatics, fev. 2008.

NASH, J. F. **Non-Cooperative Games**. Tese (Doutorado) — Princeton University, Maio 1950.

NOCEDAL, J.; WRIGHT, S. J. **Numerical Optimization**. Series in Operations Research: Springer, 1999.

POLYAK, B. T. **Introduction to Optimization**. New York: [s.n.], 1987.

POWELL, M. A method for nonlinear constraints in minimizations problems in optimization. **R. Fletcher**, p. 283–298, 1969.

RIBEIRO, A. A.; KARAS, E. W. **Otimização Contínua: Aspectos Teóricos e Computacionais**. Universidade Federal do Paraná: Cengage Learning, 2013.

ROCKAFELLAR, R. T. Augmented lagrangians and applications of the proximal point algorithm in convex programming. **Mathematics of Operations Research**, v. 1,n. 2, p. 97–116, 1976a.

ROCKAFELLAR, R. T. Monotone operators and the proximal point algorithm. **Journal on Control Optimization**, v. 14,n. 5, p. 877–898, 1976b.

ROCKAFELLAR, R. T. Nonlinear analysis: Theory, methods applications. Vol.9, n.7, p. 665–698, 1985.

ROSEN, J. Existence and uniqueness of equilibrium points for concave n-person games. **Econometrica**, v. 33,n. 3, p. 520–534, julho 1965.

SANT'ANA, D. C. **Uma Nova Proposta Utilizando Métodos de Lagrangeano Aumentado com Penalidades Modernas na Resolução de Problemas de Contato**. 89 p. Dissertação (Mestrado em Ciências) — Curso de Pós-graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Curitiba, 2005.

SANTOS, P.; SCHEIMBERG, S. An inexact subgradient algorithm for equilibrium problems. **Computational Applied Mathematics**, v. 30,n. 1, p. 91–107, 2011.

SOBRAL, F. N. C. **Programação em Dois Níveis: reformulação utilizando as condições de KKT**. 89 p. Dissertação (Mestrado em Ciências) — Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo, 2008.

SOUZA, T. C. A. **Métodos Subgradientes em Otimização Convexa não Diferenciável**. 104 p. Dissertação (Mestrado em Ciências) — Universidade Federal de Juiz de Fora, Juiz de Fora, MG, 2008.

TANAKA, M. **Algoritmos de Direções Viáveis para Otimização não Diferenciável**. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, Rio de Janeiro - RJ, 2011.

TRAN, D. Q.; DUNG, M. L.; NGUYEN, V. H. Extragradient algorithms extended to equilibrium problems. **Bibliothèque Univ De Namur**, At: 11:49, março 2006.