

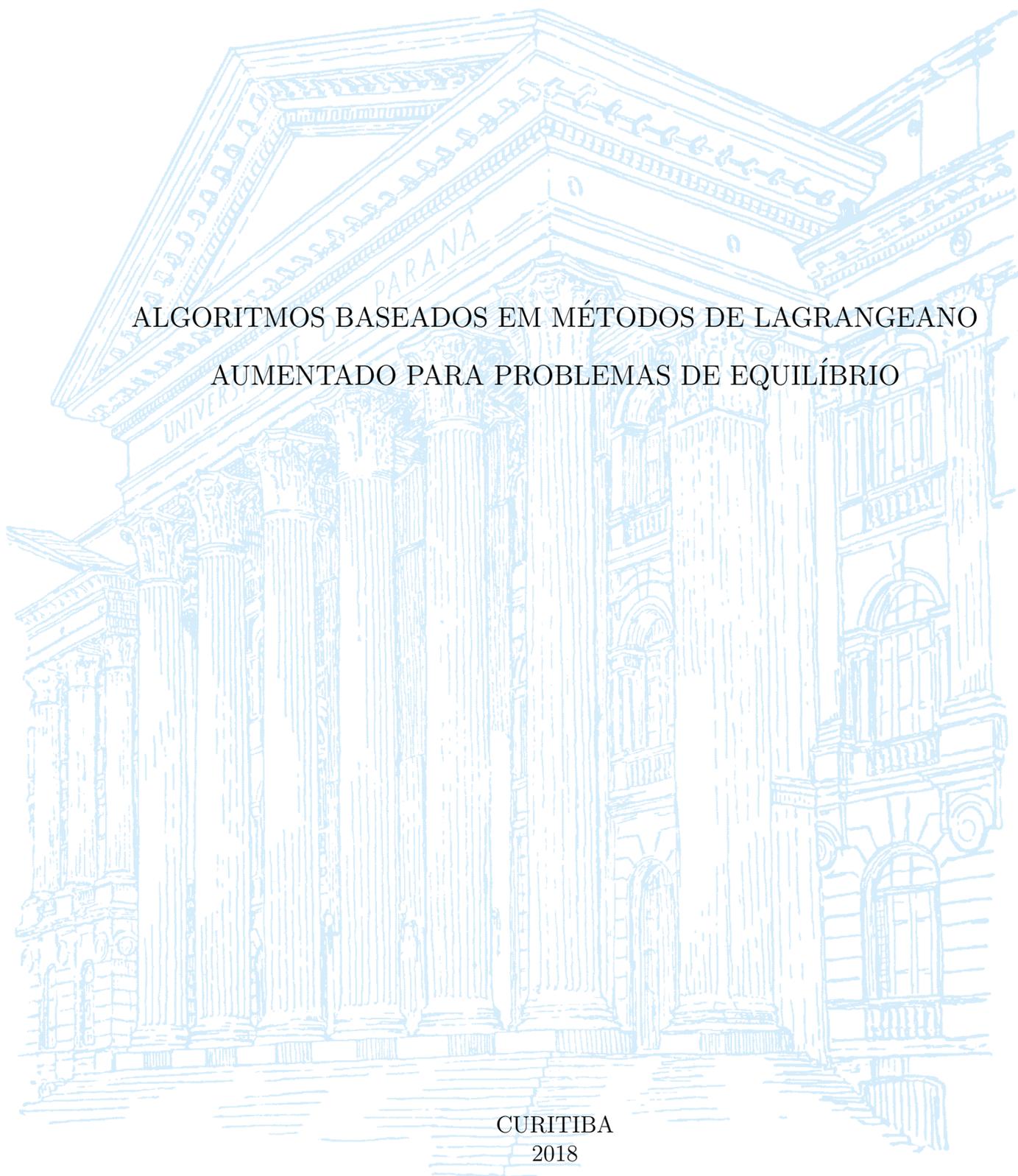
UNIVERSIDADE FEDERAL DO PARANÁ

ELVIS MANUEL RODRIGUEZ TORREALBA

ALGORITMOS BASEADOS EM MÉTODOS DE LAGRANGEANO
AUMENTADO PARA PROBLEMAS DE EQUILÍBRIO

CURITIBA

2018



ELVIS MANUEL RODRIGUEZ TORREALBA

ALGORITMOS BASEADOS EM MÉTODOS DE LAGRANGEANO
AUMENTADO PARA PROBLEMAS DE EQUILÍBRIO

Tese apresentada como requisito parcial à
obtenção do grau de Doutor em Matemática,
no Curso de Pós-Graduação em Matemática,
Setor de Ciências Exatas, da Universidade
Federal do Paraná.

Orientador: Prof. Dr. Luiz Carlos Matioli.

CURITIBA
2018

FICHA CATALOGRÁFICA ELABORADA PELO SISTEMA DE BIBLIOTECAS/UFPR
BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

T689a Torrealba, Elvis Manuel Rodriguez
Algoritmos baseados em métodos de Lagrangeano Aumentado para problemas de equilíbrio / Elvis Manuel Rodriguez Torrealba. – Curitiba, 2018.
90 p. : il. color. ; 30 cm.

Tese - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Matemática, 2018.

Orientador: Luiz Carlos Matioli.

1. Lagrangeano aumentado. 2. Ponto proximal. 3. Problemas de equilíbrio de Nash. 4. Método de Newton. 5. Método de projeção do subgradiente. I. Universidade Federal do Paraná. II. Matioli, Luiz Carlos. III. Título.

CDD: 518.1

Bibliotecária: Romilda Santos - CRB-9/1214



TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **ELVIS MANUEL RODRIGUEZ TORREALBA** intitulada: **ALGORITMOS BASEADOS EM MÉTODOS DE LAGRANGEANO AUMENTADO PARA PROBLEMAS DE EQUILÍBRIO**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 16 de Março de 2018.

LUIZ CARLOS MATIOLI
Presidente da Banca Examinadora

SOLANGE REGINA DOS SANTOS
Avaliador Externo

MAEL SACHINE
Avaliador Interno

LUCELINA BATISTA DOS SANTOS
Avaliador Interno

ROMULO ALBERTO CASTILLO CARDENAS
Avaliador Externo

PAULO SERGIO MARQUES DOS SANTOS
Avaliador Externo

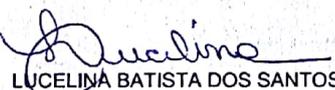


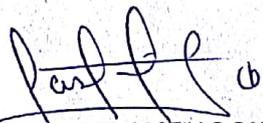
**ATA DE SESSÃO PÚBLICA DE DEFESA DE DOUTORADO PARA A OBTENÇÃO DO
GRAU DE DOUTOR EM MATEMÁTICA**

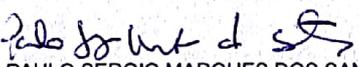
No dia dezesseis de março de dois mil e dezoito às 14:00 horas, na sala de Reuniões, Centro Politécnico da UFPR - Edifício da Administração - 3º Andar Rua Cel. Francisco H. dos Santos, 100 - Jardim das Américas, foram instalados os trabalhos de arguição do doutorando **ELVIS MANUEL RODRIGUEZ TORREALBA** para a Defesa Pública de sua tese intitulada **ALGORITMOS BASEADOS EM MÉTODOS DE LAGRANGEANO AUMENTADO PARA PROBLEMAS DE EQUILÍBRIO**. A Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná, foi constituída pelos seguintes Membros: LUIZ CARLOS MATIOLI (UFPR), LUCELINA BATISTA DOS SANTOS (UFPR), ROMULO ALBERTO CASTILLO CARDENAS (UFSC), PAULO SERGIO MARQUES DOS SANTOS (UFPI), SOLANGE REGINA DOS SANTOS (UNESPAR/PR), MAEL SACHINE (UFPR). Dando início à sessão, a presidência passou a palavra ao discente, para que o mesmo expusesse seu trabalho aos presentes. Em seguida, a presidência passou a palavra a cada um dos Examinadores, para suas respectivas arguições. O aluno respondeu a cada um dos arguidores. A presidência retomou a palavra para suas considerações finais. A Banca Examinadora, então, reuniu-se e, após a discussão de suas avaliações, decidiu-se pela aprovação do aluno. O doutorando foi convidado a ingressar novamente na sala, bem como os demais assistentes, após o que a presidência fez a leitura do Parecer da Banca Examinadora. A aprovação no rito de defesa deverá ser homologada pelo Colegiado do programa, mediante o atendimento de todas as indicações e correções solicitadas pela banca dentro dos prazos regimentais do programa. A outorga do título de doutor está condicionada ao atendimento de todos os requisitos e prazos determinados no regimento do Programa de Pós-Graduação. Nada mais havendo a tratar a presidência deu por encerrada a sessão, da qual eu, LUIZ CARLOS MATIOLI, lavrei a presente ata, que vai assinada por mim e pelos membros da Comissão Examinadora.

Curitiba, 16 de Março de 2018.


LUIZ CARLOS MATIOLI
Presidente da Banca Examinadora


LUCELINA BATISTA DOS SANTOS
Avaliador Interno


ROMULO ALBERTO CASTILLO CARDENAS
Avaliador Externo


PAULO SERGIO MARQUES DOS SANTOS
Avaliador Externo


SOLANGE REGINA DOS SANTOS
Avaliador Externo


MAEL SACHINE
Avaliador Interno

AGRADECIMENTOS

A Deus pela vida, saúde, capacitação e força para continuar.

Aos meus pais Irma e Rafael, aos meus irmãos Rafael, Esneider e Mercedes pela compreensão e apoio incondicional.

A minha esposa Aura, meu amor, pelo carinho, pelos mimos e pelos abraços e suas palavras nos momentos difíceis.

Ao resto da minha família que tanto me incentivaram e apoiaram ao longo deste caminho, em especial, aqueles que perdi fisicamente porém que estarão sempre no meu coração.

Ao meu orientador Luiz Carlos Matioli pela dedicação, tempo, conselhos e esforço neste trabalho.

Aos professores Rómulo Castillo e Hugo Lara pelo incentivo e pela ajuda para vir ao Brasil, Carlos Durán pela ajuda no Brasil e aos professores do departamento de matemática da UFPR pelo tempo e pela ajuda, e pelos conhecimentos transmitidos ao longo da minha formação.

Aos membros da banca, professores Paulo Santos, Solange Santos, Lucelina Batista, Rómulo Castillo e Mael Sachine por aceitarem o convite e pelas correções para melhorar o presente trabalho.

Aos meus amigos da Venezuela, do Brasil e aos não brasileiros que conheci no Brasil ao longo deste período. Em especial, Leonardo, Dion, Diego, Priscila, Marcos, Ariel, Patricia, Oranys, Andres e Yuli pela ajuda sempre que precisei.

Ao programa de Pós-Graduação em Matemática da UFPR pela oportunidade e formação de excelência.

E finalmente, mas não menos importante, à CAPES pelo apoio financeiro.

“ Huid del país donde uno solo ejerce todos los poderes: es un país de esclavos.”

Simón Bolívar

RESUMO

Métodos de Lagrangeano aumentado têm se mostrado bastante eficientes para resolver problemas de programação matemática, ou seja, problemas de minimização (ou maximização) com restrições. Estes métodos são iterativos e o objetivo central é, a cada iteração, transformar um problema restrito, através da penalização, em uma sequência de subproblemas irrestritos que são resolvidos usando métodos de programação não linear.

No ano de 2010, foi introduzida uma metodologia que usa a essência do método de Lagrangeano aumentado, mas no contexto do problemas de equilíbrio, os quais são mais gerais que os problemas de programação matemática. A ideia deste método é utilizar a penalidade quadrática clássica, para transformar problemas de equilíbrio restritos em uma sequência iterativa formada por subproblemas de equilíbrio irrestritos.

Neste trabalho, apresentamos uma extensão do método de Lagrangeano aumentado para problemas de equilíbrio usando penalidades com derivadas de ordem superior, diferentemente do caso quadrático que só possui derivada de primeira ordem. Além disso propomos algoritmos para resolver os subproblemas de equilíbrio irrestritos, gerados pelo método de Lagrangeano aumentado, utilizando o método de Newton e algoritmos baseados no método de projeção do subgradiente. Todos os algoritmos desenvolvidos aqui foram implementados em Matlab, e aplicados para resolver problemas conhecidos na literatura para testar sua eficiência e comparar o desempenho entre eles.

Palavras-chave: Lagrangeano aumentado; Ponto proximal; problemas de equilíbrio; problemas de equilíbrio de Nash; Método de Newton; Método de projeção do subgradiente.

ABSTRACT

Augmented Lagrangian methods have been shown to be very efficient in solving problems of mathematical programming, that is, problems of minimization (or maximization) with constraints. These methods are iterative and the main objective is, at each iteration, to transform a constrained problem through penalty into a sequence of unrestricted subproblems that are solved using nonlinear programming methods.

In the year 2010, a new methodology was introduced that uses the essence of the Augmented Lagrangian method, but in the context of the equilibrium problems, which are more general than the problems of mathematical programming. The idea of this method is to use the classical quadratic penalty to transform constrained equilibrium problems into an iterative sequence formed by unrestricted equilibrium subproblems.

In this work, we present an Augmented Lagrangian method extension for equilibrium problems using penalties with higher order derivatives, differently the quadratic case that has only first order derivatives. In addition, we propose algorithms to solve the unrestricted equilibrium subproblems, generated by the Augmented Lagrangian method, using the Newton method and algorithms based on the subgradient projection method. All algorithms developed here were implemented in Matlab, and applied to solve problems known in the literature to test their efficiency and to compare the performance between them.

Keywords: Augmented Lagrangian methods; Proximal point methods; Equilibrium problems; Nash Equilibrium problems; Newton methods; Subgradient projection methods.

LISTA DE SÍMBOLOS

\mathbb{R} : Conjunto dos números reais.

\mathbb{R}^n : Espaço vetorial de todas as n -uplas de números reais.

\mathbb{R}_+^n : Espaço vetorial de todas as n -uplas de números reais não negativos.

$\langle \cdot, \cdot \rangle$: Produto interno em \mathbb{R}^n .

$B(a, r)$: Bola aberta de centro em a e raio r .

$\overline{B(a, r)}$: Fecho da bola aberta de centro em a e raio r .

$(x_k)_{k \in \mathbb{N}}$: Sequência.

$\partial f(x)$: Subdiferencial da função f em x .

∂D : Fronteira do Conjunto D .

Jf : Jacobiana da função f .

Sumário

Introdução	11
1 Preliminares	14
1.1 O problema de equilíbrio	14
1.2 Distâncias de Bregman e suas propriedades	17
2 Métodos de ponto proximal exatos e inexatos para problemas de equilíbrio	22
3 Lagrangeano aumentado para problemas de equilíbrio	38
3.1 O Método de Lagrangeano aumentado para problemas de equilíbrio	38
3.2 Análise de convergência	44
4 Implementação dos algoritmos propostos	56
4.1 Método de Newton para resolver o Passo 2 de IALPM	56
4.2 Método de Projeção do Subgradiente para resolver o Passo 2 de IALPM . . .	62
4.2.1 Método de Projeção do Subgradiente	62
5 Testes Numéricos	74
Conclusão	85
Trabalhos futuros	87
Referências Bibliográficas	88

Introdução

Historicamente, o problema de equilíbrio foi introduzido em [10, 11] com o nome de problema de desigualdade de Ky Fan e retomado formalmente por Blum e Oettli em [3], cujo artigo apresentam o problema na forma como é conhecido até hoje. Mundialmente, este problema tem recebido muita atenção por parte de pesquisadores. Por um lado, devido à aplicações em áreas como Física, Química, Engenharia e Economia. Por outro lado, pela sua formulação matemática, que apresenta uma estrutura simples e unificada que inclui diversos problemas como casos particulares, dentre os quais podemos mencionar: problema de minimização convexa, problemas de ponto fixo, problemas de complementaridade, problemas de equilíbrio de Nash, problemas de desigualdade variacional e problemas de minimização vetorial (ver, [1, 3, 15, 20]). Muitos autores têm estudado condições para encontrar soluções do problema de equilíbrio em dimensão finita e infinita, por exemplo [1, 14, 15], e estas condições são muito ricas no caso de dimensão finita, pois permitem construir diversos algoritmos, dentre os quais destacamos [1, 15, 16, 18].

A classe de métodos de Lagrangeano aumentado tem se mostrado muito eficiente para os problemas de programação não linear [2]. Recentemente, em [16] foi introduzido um método de Lagrangeano aumentado para resolver problemas de equilíbrio em dimensão finita, para o caso em que o conjunto viável é formado por restrições de desigualdade convexas. Este método foi definido utilizando a função de penalidade quadrática, introduzida por Powell-Hestenes-Rockafellar em [12, 27, 28], para assim generalizar o método de Lagrangeano aumentado clássico para o problema de otimização restrita. Na estrutura do método de Lagrangeano aumentado para problemas de equilíbrio apresentado em [16], os autores mostram uma conexão com um método de ponto proximal para problemas de equilíbrio definido em [15], o qual, no caso de dimensão finita, permite mostrar a convergência de uma sequência gerada por este algoritmo. Para realizar implementações é preciso utilizar algoritmos externos que permitam resolver os subproblemas de equilíbrio irrestritos gerados internamente pela metodologia de Lagrangeano aumentado. Em [16] foi mostrado que uma forma de resolver estes subproblemas equivale a encontrar raízes de um sistema não linear e não diferenciável. Em [25], os autores apresentam uma abordagem para tornar o sistema diferenciável e, através desta, introduzem dois algoritmos implementáveis, aplicando-os a problemas já conhecidos da literatura, dentre os quais [24, 29, 30].

Utilizando uma abordagem similar à utilizada em [16], combinada com a teoria do

método de Lagrangeano aumentado para problemas de otimização, apresentamos ao longo desta tese, duas propostas: a primeira consiste em apresentar de maneira unificada uma metodologia que permite definir os métodos de Lagrangeano aumentado com penalidade quadrática utilizada em [16], juntamente com a nova versão deste método utilizando a penalidade exponencial a qual permite obter métodos diferentes para abordar problemas de equilíbrio. A segunda proposta é desenvolver novos algoritmos baseados em [15, 16] e combinados com as ideias propostas em [18, 25]. Para isso, apresentamos inicialmente os métodos de Lagrangeano Aumentado inexato o qual é uma variação do método original proposto em [16]. Em seguida mostraremos que, como em [16], este método também gera, internamente, subproblemas de equilíbrio irrestritos que posteriormente serão combinados com três métodos diferentes para assim torná-lo implementável. O primeiro método está baseado em [25] e consiste em suavizar um sistema não linear e não diferenciável para então usar o método de Newton, e assim, encontrar soluções dos subproblemas gerados. O segundo e o terceiro algoritmos que estamos propondo são baseados em [18] e cada iteração consiste de duas etapas: uma maximização inexata de uma função contínua em um conjunto convexo interceptado com uma bola centrada na origem e uma projeção aproximada em um hiperplano, cujo custo computacional é insignificante. Como estes algoritmos serão aplicados para resolver os subproblemas irrestritos gerados pelo método de Lagrangeano aumentado, a tarefa computacionalmente pesada para o segundo e terceiro algoritmos se reduz a uma maximização de uma função contínua em uma bola, que é relativamente fácil, em comparação com problemas de maximização com restrições adicionais do tipo $g(x) \leq 0$, que é o caso se o mesmo algoritmo definido em [18] é aplicado para o problema original, pois este método não considera a regularização realizada pelo método de Lagrangeano aumentado. A principal diferença deste algoritmo com respeito ao primeiro é que este tem a capacidade de resolver problemas de equilíbrio não diferenciáveis.

Reforçamos esta proposta implementando estes algoritmos em Matlab. Para os testes elegemos problemas quadráticos, problemas de equilíbrio de Nash generalizado e um problema não diferenciável. Todos estes problemas são amplamente discutidos no contexto computacional, por exemplo em [9, 24, 25, 29, 30]. Salientamos que nesta tese iremos usar o termo programação não linear e problema de otimização como sinônimos, pois na literatura um problema de programação não linear é comumente chamado de problema de otimização. Por outro lado, como o problema de equilíbrio generaliza o problema de otimização (máximizsar ou minimizar uma função), então usaremos a expressão otimização ordinária quando referirmos a um problema de otimização clássica.

Em resumo, as principais contribuições que encontramos ao longo desta tese são:

- Generalizamos os métodos de ponto proximal inexatos propostos em [15], para uma classe mais ampla de funções.
- Generalizamos o conceito de Lagrangeano aumentado para problemas de equilíbrio para penalidades diferenciáveis.

- Propomos novos algoritmos implementáveis utilizando diferentes penalidades.
- Realizamos testes numéricos de cada um dos algoritmos propostos, utilizando diversos problemas da literatura.

A organização do trabalho será da seguinte forma: No Capítulo 1, vamos descrever os elementos necessários para abordagem do problema de equilíbrio. No Capítulo 2, apresentamos os métodos de ponto proximal definidos em [4, 15], assim como algumas diferenças e semelhanças entre eles. No Capítulo 3, apresentamos os algoritmos de Lagrangeano aumentado com penalidade quadrática, baseado em [16], e estendemos os resultados para o caso da penalidade exponencial. No Capítulo 4, vamos propor algoritmos implementáveis. Para isto, vamos dividir o capítulo em duas seções: na primeira utilizamos o método de Newton para resolver um sistema não linear, e na segunda consideramos dois algoritmos baseados nos métodos de projeção do subgradiente. Todos os algoritmos deste capítulo serão utilizados para resolver o subproblema gerado pelo método de Lagrangeano aumentado. No Capítulo 5, apresentamos simulações numéricas para estes algoritmos. Finalizamos este trabalho com algumas conclusões e recomendações para trabalhos futuros.

Capítulo 1

Preliminares

O objetivo deste capítulo é descrever os conceitos necessários para abordar os métodos de ponto proximal e Lagrangeano aumentado para problemas de equilíbrio. Começamos formulando o problema de equilíbrio geral no qual salientamos, de forma breve, a importância do seu estudo e em seguida, consideramos alguns conceitos básicos, notações e hipóteses que serão utilizados ao longo do texto.

1.1 O problema de equilíbrio

O problema de equilíbrio que abordaremos neste trabalho é o mais comum da literatura, por exemplo [1, 4, 14, 15, 16, 29] e está descrito da seguinte forma:

Definição 1.1. *Sejam C um subconjunto não vazio, convexo e fechado de \mathbb{R}^n e $f : C \times C \rightarrow \mathbb{R}$ uma função que satisfaz as seguintes propriedades*

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

P2: $f(x, \cdot) : C \rightarrow \mathbb{R}$ é convexa e semicontínua inferior para todo $x \in C$,

P3: $f(\cdot, y) : C \rightarrow \mathbb{R}$ é semicontínua superior para todo $y \in C$.

O problema de equilíbrio, denotado por $EP(f, C)$, consiste em

$$\text{Encontrar } x^* \in C \text{ tal que } f(x^*, y) \geq 0, \text{ para todo } y \in C, \quad (1.1)$$

ou equivalentemente

$$\text{Encontrar } x^* \in \underset{y \in C}{\operatorname{argmin}} f(x^*, y).$$

Esta equivalência será provada no próximo capítulo, na Proposição 2.16. As funções que satisfazem a propriedade P1 são chamadas de bifunção de equilíbrio [1]. Um ponto x^* que cumpra a condição (1.1) é chamado de solução de $EP(f, C)$. O conjunto formado pelas soluções de $EP(f, C)$ será denotado por $S(f, C)$.

Para fixar ideias e notações referentes a este problema, consideremos o seguinte exemplo.

Exemplo 1.2. Seja $C = [1, 2] \subset \mathbb{R}$. Defina $f : C \times C \rightarrow \mathbb{R}$ como

$$f(x, y) = y^2 - x^2.$$

Então o conjunto solução do problema de equilíbrio é $S(f, C) = \{1\}$.

De fato, note que $f(x, y) = y^2 - x^2 = (y - x)(y + x)$. Como $1 \leq x \leq 2$ e $1 \leq y \leq 2$ então temos que $f(x, y) \geq 0$ se e somente se $y \geq x$. Portanto, $y \geq x$ para todo $y \in C$ se e somente se $x = 1$, ou seja, $S(f, C) = \{1\}$. \square

Geometricamente, resolver o problema de equilíbrio significa encontrar um $x^* \in C$ tal que o conjunto de nível $\{y \in \mathbb{R}^n : f(x^*, y) \geq 0\}$ contenha completamente o conjunto C , ou equivalentemente, dado $x^* \in C$ queremos encontrar uma curva $h(y) = f(x^*, y)$ tal que sua imagem restrita ao conjunto C , esteja acima (não abaixo) do hiperplano $Z = 0$.

Na Figura 1.1 ilustramos geometricamente o Exemplo 1.2. Observe que para $x^* = 1$, a curva $h(y) = f(1, y) = y^2 - 1$ restrita ao conjunto C , está acima do plano $Z = 0$, para qualquer $y \in C$ o que reafirma que $S(f, C) = \{1\}$.

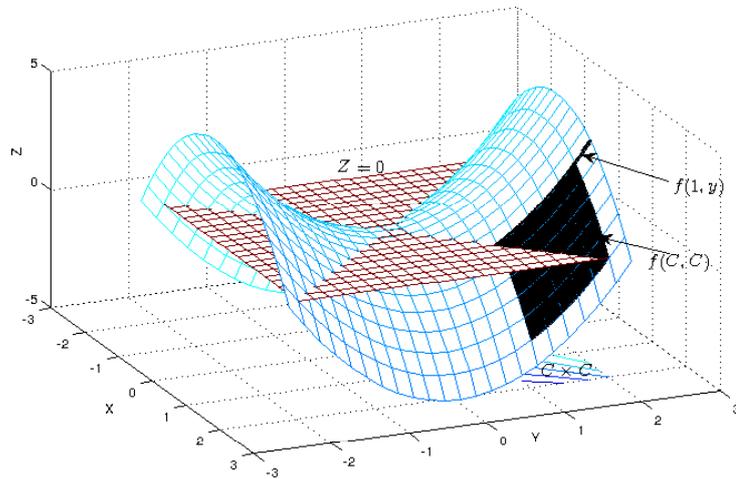


Figura 1.1: Curva $h(y) = f(1, y) = y^2 - 1$.

Um dos principais motivos pelos quais o problema de equilíbrio é estudado se deve ao fato de que ao resolver (1.1), na verdade estamos resolvendo diversos problemas como casos particulares. Para mostrar esta afirmação vamos ilustrar alguns problemas

bem conhecidos, que podem ser reescritos no formato do problema de equilíbrio dado na Definição 1.1.

- **Problemas de otimização convexa:** um problema de otimização convexa consiste em encontrar um minimizador de uma função $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa, definida sobre um conjunto convexo, não vazio e fechado $C \subset \mathbb{R}^n$. Este problema pode ser resolvido como um problema de equilíbrio tomando a bifunção

$$f(x, y) = \varphi(y) - \varphi(x).$$

- **Problemas de ponto de sela:** dados dois conjuntos convexos e fechados $C_1 \subset \mathbb{R}^{n_1}$ e $C_2 \subset \mathbb{R}^{n_2}$ e uma função a valores reais $L : C_1 \times C_2 \rightarrow \mathbb{R}$, um ponto $x^* = (x_1^*, x_2^*) \in C_1 \times C_2$ é um ponto de sela para L se

$$L(x_1^*, y_2) \leq L(x_1^*, x_2^*) \leq L(y_1, x_2^*),$$

para qualquer $y = (y_1, y_2) \in C_1 \times C_2$. Este problema é equivalente a resolver $\text{EP}(f, C)$, se consideramos $C = C_1 \times C_2$ e a bifunção f como

$$f((x_1, x_2), (y_1, y_2)) = L(y_1, x_2) - L(x_1, y_2).$$

- **Problemas de otimização de Pareto:** considere o conjunto de índices $I = \{1, \dots, m\}$ e $C \subset \mathbb{R}^n$ um conjunto fechado. Defina a função vetorial $\psi = (\psi_1, \dots, \psi_m)$, em que $\psi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ para cada $i \in I$. Um ponto $x^* \in C$ é minimizador global Pareto fraco de ψ se existe $i \in I$ tal que $\psi_i(y) - \psi_i(x^*) \geq 0$, para todo $y \in C$. Assim, encontrar um minimizador global Pareto fraco é equivalente a resolver $\text{EP}(f, C)$ para f definida como

$$f(x, y) = \max_{i \in I} [\psi_i(y) - \psi_i(x)].$$

- **Complementaridade e sistemas de equações:** dado um cone convexo e fechado $C \subset \mathbb{R}^n$ e uma aplicação $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, o problema de complementaridade consiste em determinar um ponto $x^* \in C$ tal que $\langle F(x^*), v \rangle \geq 0$ para todo $v \in C$, ou seja, $F(x^*) \in C^*$, em que C^* é o cone dual de C . Este problema pode ser resolvido considerando $f(x, y) = \langle F(x), y - x \rangle$. Como um caso particular, note que se $C = \mathbb{R}^n$, temos que a solução é dada pelos pontos $x^* \in \mathbb{R}^n$ tais que $F(x^*) = 0$. Assim resolver sistemas de equações também é um caso particular do problema de equilíbrio.
- **Desigualdade variacional:** dado um conjunto $C \subset \mathbb{R}^n$ convexo e fechado e uma aplicação $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. O problema de desigualdade variacional consiste em determinar um ponto $x^* \in C$ tal que $\langle F(x^*), y - x^* \rangle \geq 0$, para todo $y \in C$. Portanto, no contexto do problema de equilíbrio basta considerar $f(x, y) = \langle F(x), y - x \rangle$.

- **Desigualdade variacional com aplicações ponto conjunto:** considere o conjunto convexo e fechado $C \subset \mathbb{R}^n$ e uma aplicação $T : C \rightarrow \mathbb{R}^n$ com a propriedade que dado $x \in C$, o conjunto imagem $Tx \subset \mathbb{R}^n$ é não vazio, convexo e compacto. Assim, o problema de desigualdade variacional com aplicações ponto conjunto consiste em resolver o problema

Encontrar $x^* \in C$ e $\xi^* \in T(x^*)$, tal que $\langle \xi^*, y - x^* \rangle \geq 0$, para todo $y \in C$.

Logo, para escrever este problema como um problema de equilíbrio, basta considerar

$$f(x, y) = \max_{\xi \in T(x)} \langle \xi, y - x \rangle.$$

- **Problema de ponto fixo:** considere o conjunto fechado $C \subset \mathbb{R}^n$. Um ponto fixo $x^* \in C$ de uma aplicação $F : C \rightarrow C$ é tal que $F(x^*) = x^*$. Assim, encontrar um ponto com esta característica consiste em resolver $\text{EP}(f, C)$, com

$$f(x, y) = \langle x - F(x), y - x \rangle.$$

Como acabamos de descrever, problemas importantes estudados em diferentes contextos da matemática, pertencem à classe de problemas de equilíbrio que queremos resolver. Tendo isto em vista, na sequência vamos lembrar o conceito de distância de Bregman, assim como algumas propriedades importantes que utilizaremos ao longo do texto para definir os métodos de ponto proximal e Lagrangeano aumentado para problemas de equilíbrio.

1.2 Distâncias de Bregman e suas propriedades

Seja $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função estritamente convexa, própria e semi-contínua inferior. Ao longo do texto denotaremos o domínio de ψ por \mathcal{D} e o interior deste domínio por $\text{int}\mathcal{D}$. Além disso, exigiremos que $\text{int}\mathcal{D} \neq \emptyset$ e que ψ admite derivada direcional em $\text{int}\mathcal{D}$. Com estas considerações, definimos a distância de Bregman com respeito à função ψ como a função $D_\psi : \mathcal{D} \times \text{int}\mathcal{D} \rightarrow \mathbb{R}$ dada por

$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y \rangle, \quad (1.2)$$

em que $\nabla\psi(\cdot)$ denota o gradiente da função ψ e $\langle \cdot, \cdot \rangle$ o produto interno usual em \mathbb{R}^n .

A seguinte proposição estabelece algumas propriedades básicas satisfeitas pela função D_ψ definida em (1.2). A demonstração será omitida pois decorre diretamente da definição.

Proposição 1.3. *Sejam ψ e D_ψ definidas acima, no início da seção, então*

- (i) $D_\psi(x, y) \geq 0$ e $D_\psi(x, y) = 0$ se, e somente $x = y$, para todo $x \in \mathcal{D}$ e $y \in \text{int}\mathcal{D}$.
- (ii) $D_\psi(\cdot, y)$ é estritamente convexa, para todo $y \in \text{int}\mathcal{D}$.
- (iii) Dado $\alpha \in \mathbb{R}$, os conjuntos de nível à esquerda $\Gamma(x, y) = \{x \in \mathcal{D} : D(x, y) \leq \alpha\}$ são limitados, para todo $y \in \text{int}\mathcal{D}$.
- (iv) Para quaisquer $x \in \mathcal{D}$ e $y, z \in \text{int}\mathcal{D}$, tem-se que $D_\psi(x, y) - D_\psi(x, z) - D_\psi(z, y) = \langle \nabla\psi(y) - \nabla\psi(z), z - x \rangle$.
- (v) Para quaisquer $x, y \in \text{int}\mathcal{D}$, $\nabla_x D(x, y) = \nabla\psi(x) - \nabla\psi(y)$.

Note que $D_\psi(x, y)$ não necessariamente é uma distância (pois as propriedades de simetria e a desigualdade triangular podem não ser satisfeitas) porém como ψ é estritamente convexa então as propriedades $D_\psi(x, y) \geq 0$ e $D_\psi(x, y) = 0$ se, e só se $x = y$ são satisfeitas de forma imediata. No caso especial em que escolhermos $\psi(x) = \frac{1}{2}\|x\|^2$, com $S = \mathbb{R}^n$ obtemos a distância de Bregman $D_\psi(x, y) = \frac{1}{2}\|x - y\|^2$, a qual é a distância clássica. Outra exemplo muito utilizado na literatura é a distância induzida pela função Kullback-Leibler $\psi(x) = \sum_{i=1}^n x_i \log(x_i) - x_i + 1$, em que $S = \{x \in \mathbb{R}^n : x_i > 0, i = 1, \dots, m\}$, e para os pontos na fronteira, usamos a convenção de que $0 \log(0) = 0$. A distância de Bregman induzida por esta função é dada por $D_\psi(x, y) = \sum_{i=1}^n \left(x_i \log\left(\frac{x_i}{y_i}\right) + y_i - x_i + 1 \right)$. Outras caracterizações e exemplos de funções de Bregman podem ser encontradas em [6, 7].

Além das propriedades dadas na Proposição 1.3, vamos estabelecer outras hipóteses sobre a função ψ que define D_ψ .

H1: Dado $\alpha \in \mathbb{R}$, os conjuntos de nível à direita $\Gamma(x, y) = \{y \in \text{int}\mathcal{D} : D(x, y) \leq \alpha\}$ são limitados, para todo $x \in \mathcal{D}$.

Para definir a segunda hipótese precisaremos lembrar o conceito de convexidade total introduzido em [5]. Daqui em diante, denotaremos por \mathbb{R}_+^n e $\mathbb{R}_{++}^n = \mathbb{R}_+^n \setminus \{0\}$ os ortantes não negativo e positivo, respectivamente, do espaço vetorial \mathbb{R}^n .

Definição 1.4. [Convexidade total] Definimos o módulo de convexidade de ψ , como sendo a função $\nu_\psi : \text{int}\mathcal{D} \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ dada por

$$\nu_\psi(z, t) = \inf_{x \in \mathcal{D}} \{D_\psi(x, z) : \|x - z\| = t\}$$

e dizemos que uma função ψ é dita totalmente convexa em $\text{int}\mathcal{D}$ se e somente se $\nu_\psi(z, t) > 0$ para todo $z \in \text{int}\mathcal{D}$ e $t > 0$.

Este conceito é estudado em [5] e uma das propriedades mais importantes em dimensão finita é estabelecida na seguinte proposição.

Proposição 1.5. Seja $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função diferenciável, convexa, própria e semicontínua inferior. Então vale que:

(i) Se $c \in [1, +\infty)$ e $t \geq 0$, então $\nu_\psi(x, ct) \geq c\nu_\psi(x, t)$.

(ii) Para quaisquer $s, t \in [0, +\infty)$, $\nu_\psi(x, s+t) \geq \nu_\psi(x, s) + \nu_\psi(x, t)$.

(iii) Se ψ é totalmente convexa então, é estritamente convexa.

(vi) Seja \mathcal{D} fechado. Se ψ é contínua e estritamente convexa em \mathcal{D} então, ψ é totalmente convexa.

Demonstração. Ver págs. 18 e 27 de [5]. □

Esta última proposição estabelece que uma função ψ é totalmente convexa equivale a dizer que ψ é estritamente convexa quando \mathcal{D} é fechado. Com isto claro, definimos a segunda hipótese, sobre conjuntos limitados.

H2: ψ é totalmente convexa em conjuntos limitados, isto é, $\inf_{x \in C} \nu_\psi(x, t) > 0$, para todo conjunto C limitado no $\text{int}\mathcal{D}$ e $t > 0$.

Com esta hipótese podemos obter o seguinte resultado mostrado em [15].

Proposição 1.6. Se ψ satisfaz H2, então para $v, \tilde{y} \in \mathbb{R}^n$ tal que $v \neq 0$, $x \in H^+ = \{y \in \mathbb{R}^n : \langle v, y - \tilde{y} \rangle \geq 0\}$ e $\bar{x} \in H^- = \{y \in \mathbb{R}^n : \langle v, y - \tilde{y} \rangle \leq 0\}$ vale que $D_\psi(\bar{x}, x) \geq D_\psi(\bar{x}, z) + D_\psi(z, x)$ em que z é o único minimizador de $D_\psi(\cdot, \bar{x})$ em $H = \{y \in \mathbb{R}^n : \langle v, y - \tilde{y} \rangle = 0\}$.

Demonstração. Ver Proposição 2.4 de [15]. □

Proposição 1.7. Suponha que ψ satisfaz H2. Sejam $(x^j)_{j \in \mathbb{N}}, (y^j)_{j \in \mathbb{N}}$ seqüências em \mathbb{R}^n tais que pelo menos uma é limitada. Se $\lim_{j \rightarrow \infty} D_\psi(y^j, x^j) = 0$ então $\lim_{j \rightarrow \infty} \|x^j - y^j\| = 0$.

Demonstração. Veja Proposição 2.6 de [15]. □

Antes de definir a próxima hipótese, consideremos a seguinte caracterização das funções uniformemente contínuas.

Proposição 1.8. $h : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é uniformemente contínua se e somente se para todo par de seqüências $(x_n)_{n \in \mathbb{N}}, (y_n)_{n \in \mathbb{N}} \subset A$, com $\lim_{n \rightarrow \infty} \|x_n - y_n\| = 0$ vale que $\lim_{n \rightarrow \infty} |h(x_n) - h(y_n)| = 0$.

Demonstração. Suponha inicialmente que h é uniformemente contínua, ou seja, para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que $\forall x, y \in A$ com $\|x - y\| < \delta$ implica $|h(x) - h(y)| < \varepsilon$. Logo, como $(x_n)_{n \in \mathbb{N}}, (y_n)_{n \in \mathbb{N}} \subset A$ tais que $\lim_{n \rightarrow \infty} \|x_n - y_n\| = 0$, então para $\delta > 0$, existe $n_0 \in \mathbb{N}$ tal que $\|x_n - y_n\| < \delta$ para todo $n \geq n_0$ que pela continuidade uniforme implica que $|h(x_n) - h(y_n)| < \varepsilon$, para $n \geq n_0$, ou seja, $\lim_{n \rightarrow \infty} |h(x_n) - h(y_n)| = 0$. Reciprocamente suponha por absurdo que h não é uniformemente contínua. Assim, existe $\varepsilon > 0$ tal que para todo $n \in \mathbb{N}$, podemos encontrar $x_n, y_n \in A$ tais que $\|x_n - y_n\| < \frac{1}{n}$ e $|h(x_n) - h(y_n)| \geq \varepsilon$, ou seja, $\lim_{n \rightarrow \infty} \|x_n - y_n\| = 0$ e pela hipótese $0 = \lim_{n \rightarrow \infty} |h(x_n) - h(y_n)| \geq \varepsilon$ o qual é absurdo. □

Como uma consequência desta última proposição, consideramos a seguinte hipótese que requer que $\nabla\psi$ seja uniformemente contínua em conjuntos limitados em $\text{int}\mathcal{D}$.

H3: Se $(x^k)_{k \in \mathbb{N}} \subset \text{int}\mathcal{D}$ e $(y^k)_{k \in \mathbb{N}} \subset \text{int}\mathcal{D}$ são sequências limitadas tais que $\lim_{k \rightarrow \infty} \|x^k - y^k\| = 0$, então $\lim_{k \rightarrow \infty} (\nabla\psi(x^k) - \nabla\psi(y^k)) = 0$.

Estas hipóteses (H1-H3), foram consideradas de maneira independente tanto em [15] como em [4], para definir algoritmos de ponto proximal que resolvem o problema de equilíbrio (1.1). Mencionamos que em [15] foram definidos três algoritmos de ponto proximal, um algoritmo exato e dois algoritmos inexatos. Para obter uma boa definição dos algoritmos inexatos é preciso adicionar a seguinte hipótese

H4: $\nabla\psi$ é sobrejetora em $\text{int}\mathcal{D}$.

Algumas consequências que temos destas hipóteses são estabelecidas nas seguintes proposições.

Proposição 1.9. *Se ψ satisfaz H3, então ψ e $\nabla\psi$ são limitadas em conjuntos limitados.*

Demonstração. Ver Proposição 4 de [17].

□

Proposição 1.10. *Suponha que ψ satisfaz H1-H2 e $C \subset \mathcal{D}$ é convexo, fechado e não vazio. Se $\bar{x} \in \text{int}\mathcal{D}$, então o problema,*

$$\begin{aligned} \min \quad & D_\psi(x, \bar{x}) \\ \text{s.a} \quad & x \in C, \end{aligned} \tag{1.3}$$

admite solução única. Em particular, para $C \subset \text{int}\mathcal{D}$ temos que \hat{x} é uma solução de (1.3) se e somente se $\hat{x} \in C$ e $\langle \nabla\psi(\bar{x}) - \nabla\psi(\hat{x}), y - \hat{x} \rangle \leq 0$, para todo $y \in C$.

Demonstração. Veja [5] pág. 70.

□

Para finalizar este capítulo vamos analisar a seguinte hipótese de coercividade introduzida em [15].

H5: $\lim_{\|x\| \rightarrow \infty} [\psi(x) - \rho \|x - z\|] = +\infty$ para todo $z \in C$ fixo e $\rho \geq 0$.

Esta hipótese pode ser caracterizada da seguinte forma.

Proposição 1.11. *ψ satisfaz H5 se, e somente se,*

- i) $\lim_{\|x\| \rightarrow \infty} \psi(x) = +\infty$,
- ii) $\lim_{\|x\| \rightarrow \infty} \frac{\psi(x)}{\|x - z\|} = +\infty$.

Demonstração. Note que (i) é satisfeita triivialmente quando $\rho = 0$. Agora, suponha por absurdo que $\lim_{\|x\| \rightarrow \infty} \psi(x) < \infty$. Logo, como $\rho > 0$ então temos que $\lim_{\|x\| \rightarrow \infty} [\psi(x) - \rho \|x - z\|] =$

$-\infty$, o que é absurdo. Para o segundo item, suponha novamente por absurdo que existe $L > 0$ tal que $\lim_{\|x\| \rightarrow \infty} \frac{\psi(x)}{\|x - z\|} \leq L < +\infty$. Observe que se consideramos $\rho = L + 1$ então

$$\begin{aligned} \lim_{\|x\| \rightarrow \infty} [\psi(x) - \rho \|x - z\|] &= \lim_{\|x\| \rightarrow \infty} \|x - z\| \left[\frac{\psi(x)}{\|x - z\|} - \rho \right] \\ &\leq \lim_{\|x\| \rightarrow \infty} \|x - z\| [L - \rho] \\ &= -\infty, \end{aligned} \tag{1.4}$$

o que contradiz a hipótese H5. Para a recíproca, basta combinar (i) e (ii) com a igualdade (1.4). \square

No próximo capítulo, descrevemos os algoritmos de ponto proximal. Veremos a influência que cada uma das propriedades H1-H5 têm na definição dos algoritmos, assim como também na sua análise de convergência. Além disso, introduziremos uma hipótese dada em [4] que posteriormente adaptaremos para uma versão inexata a qual nos permitirá enfraquecer a hipótese H5, usada nos algoritmos inexatos tratados em [15], mantendo os mesmos resultados de convergência.

Capítulo 2

Métodos de ponto proximal exatos e inexatos para problemas de equilíbrio

Neste capítulo, apresentaremos os algoritmos de ponto proximal definidos em [4, 15]. Apesar de abordarmos a versão simplificada no espaço euclidiano, enfatizamos que a discussão realizada aqui pode ser feita em espaços de Banach. A finalidade é usar estes algoritmos como uma ferramenta auxiliar para mostrar que existe uma equivalência destes com os algoritmos de Lagrangeano aumentado que definiremos no próximo capítulo. Começamos definindo algumas hipóteses adicionais às P1-P3 dadas para bifunção f do problema (1.1), as quais serão usadas para garantir a boa definição e convergência dos algoritmos.

P4 : Existe $\theta \geq 0$ tal que $f(x, y) + f(y, x) \leq \theta \langle \nabla \psi(x) - \nabla \psi(y), x - y \rangle$ para todo $x, y \in C$.

P4* : Sempre que $f(x, y) \geq 0$ com $x, y \in C$, vale que $f(y, x) \leq 0$.

P4' : Para todo $x^1, x^2, \dots, x^\ell \in C$ e $t_1, t_2, \dots, t_\ell \in \mathbb{R}_+$ tal que $\sum_{i=1}^{\ell} t_i = 1$, vale que

$$\sum_{i=1}^{\ell} t_i f \left(x^i, \sum_{k=1}^{\ell} t_k x^k \right) \leq 0.$$

P5 : Para toda sequência $(x^j)_{j \in \mathbb{N}} \subset C$ satisfazendo $\lim_{j \rightarrow \infty} \|x^j\| = +\infty$, existe $u \in C$ e $j_0 \in \mathbb{N}$ tal que $f(x^j, u) \leq 0$ para todo $j \geq j_0$.

Estas hipóteses são das mais comuns encontradas na literatura para abordar a existência de soluções do problema de equilíbrio. Algumas referências onde podemos encontrar estas hipóteses são [1, 4, 14, 15, 16] e nestes artigos, os autores chamaram algumas destas propriedades como segue.

Definição 2.1. *Seja $EP(f, C)$ como na Definição 1.1. A função f é dita*

- θ -submonótona se $P4$ é satisfeita.
- Pseudomonótona se $P4^*$ é satisfeita.

A hipótese clássica que podemos encontrar na literatura é dada no caso particular em que $\theta = 0$, na hipótese $P4$. Neste caso, note que a expressão se reduz à desigualdade $f(x, y) + f(y, x) \leq 0$, para todo $x, y \in C$ e tal bifunção é conhecida como monótona. Outro caso particular para a hipótese $P4$ é dado quando $\psi = \frac{1}{2}\|\cdot\|^2$, pois neste caso, a hipótese $P4$ é reescrita como $f(x, y) + f(y, x) \leq \theta\|x - y\|^2$ para todo $x, y \in C$, a qual é usada em [16]. Logo, a propriedade de monotonicidade é mais forte que a hipótese $P4$. Vejamos com alguns exemplos dados em [14, 16, 20], algumas relações entre estas propriedades.

Proposição 2.2. *Suponha que f satisfaz as hipóteses $P1$ - $P4$ com $\theta = 0$, isto é, f é monótona. Então f é pseudomonótona.*

Demonstração. Sejam $x, y \in C$ tais que $f(x, y) \geq 0$. Como f é monótona segue que $f(x, y) + f(y, x) \leq 0$ e portanto, $f(y, x) \leq -f(x, y) \leq 0$, como queríamos. \square

A recíproca desta proposição não é necessariamente verdadeira. Vejamos o seguinte exemplo discutido em [20].

Exemplo 2.3. Seja $C = [\frac{1}{2}, 1] \subset \mathbb{R}$ e defina $f : C \times C \rightarrow \mathbb{R}$ como

$$f(x, y) = x(x - y).$$

Então f satisfaz $P1$ - $P4$ e $P4^*$ porém não é monótona.

De fato, evidentemente, f satisfaz $P1$ - $P3$. Por outro lado, para $x, y \in C$ temos que $f(x, y) + f(y, x) = (x - y)^2$, implicando que f é 1-monótona, isto é, f cumpre $P4$ com $\theta \neq 0$ e, portanto, f não é monótona. Finalmente, também temos que f satisfaz $P4^*$ pois, se $f(x, y) \geq 0$ para $x, y \in C$, então como $x \geq \frac{1}{2}$ segue que $x \geq y$. Assim, como $y \geq \frac{1}{2}$ então $f(y, x) = y(y - x) \leq 0$. \square

Em relação à hipótese $P4'$, note que claramente esta hipótese é mais fraca que monotonicidade. Por outro lado, estudemos a recíproca com o seguinte exemplo dado em [16].

Exemplo 2.4. Seja $C \subset \mathbb{R}^n$ convexo, fechado e não vazio. Defina $f : C \times C \rightarrow \mathbb{R}$ como

$$f(x, y) = -\alpha \|x\|^2 + \beta \|y\|^2 + (\alpha - \beta)\langle x, y \rangle,$$

com $\beta > \alpha > 0$. Então f satisfaz $P1$ - $P4$ e $P4'$ mas não é monótona.

De fato, as propriedades P1-P3 são satisfeitas trivialmente. Para P4, basta ver que $f(x, y) + f(y, x) = (\beta - \alpha) \|x - y\|^2$ e como por hipótese $\beta > \alpha > 0$ então f satisfaz P4 com constante de monotonicidade $\theta = (\beta - \alpha) > 0$ e assim f não pode ser monótona. Finalmente para P4', consideremos $x^1, \dots, x^q \in C$ e $t_1, \dots, t_q \geq 0$ tal que $\sum_{\ell=1}^q t_\ell = 1$. Logo,

$$\begin{aligned}
\sum_{\ell=1}^q t_\ell f(x^\ell, \sum_{\ell=1}^q t_\ell x^\ell) &= \sum_{\ell=1}^q t_\ell \left[-\alpha \|x^\ell\|^2 + \beta \left\| \sum_{\ell=1}^q t_\ell x^\ell \right\|^2 + (\alpha - \beta) \langle x^\ell, \sum_{\ell=1}^q t_\ell x^\ell \rangle \right] \\
&= -\alpha \sum_{\ell=1}^q t_\ell \|x^\ell\|^2 + \beta \sum_{\ell=1}^q t_\ell \left\| \sum_{\ell=1}^q t_\ell x^\ell \right\|^2 + (\alpha - \beta) \sum_{\ell=1}^q t_\ell \langle x^\ell, \sum_{\ell=1}^q t_\ell x^\ell \rangle \\
&= -\alpha \sum_{\ell=1}^q t_\ell \|x^\ell\|^2 + \beta \left\| \sum_{\ell=1}^q t_\ell x^\ell \right\|^2 \sum_{\ell=1}^q t_\ell + (\alpha - \beta) \langle \sum_{\ell=1}^q t_\ell x^\ell, \sum_{\ell=1}^q t_\ell x^\ell \rangle \\
&= -\alpha \sum_{\ell=1}^q t_\ell \|x^\ell\|^2 + \beta \left\| \sum_{\ell=1}^q t_\ell x^\ell \right\|^2 + (\alpha - \beta) \left\| \sum_{\ell=1}^q t_\ell x^\ell \right\|^2 \\
&= \alpha \left[\left\| \sum_{\ell=1}^q t_\ell x^\ell \right\|^2 - \sum_{\ell=1}^q t_\ell \|x^\ell\|^2 \right] \\
&\leq 0.
\end{aligned}$$

□

Salientamos que existem outras propriedades diferentes das hipóteses P4, P4*, P4', P5 que consideramos aqui. Em [14], os autores realizaram um estudo destas e outras condições sobre f utilizadas extensivamente na literatura para mostrar existência de soluções para o problema de equilíbrio.

Voltando ao objetivo do capítulo de resolver o problema de equilíbrio (1.1) utilizando algoritmos de ponto proximal exatos ou inexatos, vamos introduzir, como em [15], o conceito de regularização de uma bifunção f .

Definição 2.5. *Sejam $\gamma > 0$, $\bar{x} \in C \subset \text{int}\mathcal{D}$ e $e \in \mathbb{R}^n$. A regularização de f , com perturbação e , a qual denotaremos como $\tilde{f}^e : C \times C \rightarrow \mathbb{R}$ é definida como:*

$$\tilde{f}^e(x, y) = f(x, y) + \gamma \langle \nabla \psi(x) - \nabla \psi(\bar{x}), y - x \rangle - \langle e, y - x \rangle. \quad (2.1)$$

O seguinte resultado é uma adaptação da Proposição 3.1 dada em [4], a qual estabelece algumas propriedades básicas para a função \tilde{f}^e .

Proposição 2.6. *Seja $\bar{x} \in C \subset \text{int}\mathcal{D}$. Suponha que f satisfaz as condições P1-P4 com $\theta < \gamma$, em que θ é a constante de submonotonicidade não necessariamente nula. Então a função $\tilde{f}^e : C \times C \rightarrow \mathbb{R}$ dada como em (2.1), satisfaz P1-P4 com $\theta = 0$ (isto é, \tilde{f}^e é monótona). Além disso, se*

P6: Para cada $(x^j)_{j \in \mathbb{N}} \subset C$ tal que $\lim_{j \rightarrow \infty} \|x^j\| = \infty$, satisfaz que

$$\liminf_{j \rightarrow \infty} f(\bar{x}, x^j) + (\gamma - \theta) \langle \nabla \psi(\bar{x}) - \nabla \psi(x^j), \bar{x} - x^j \rangle + \langle e, \bar{x} - x^j \rangle > 0$$

então \tilde{f}^e satisfaz P5.

Demonstração. \tilde{f}^e satisfaz P1 por definição. Agora, se f satisfaz P2 e $y \rightarrow \langle \nabla \psi(x) - \nabla \psi(\bar{x}) - e, y - x \rangle$ é contínua e convexa então, \tilde{f}^e também satisfaz P2. Analogamente, observe que a função $x \rightarrow \langle \nabla \psi(x) - \nabla \psi(\bar{x}) - e, y - x \rangle$ é contínua e como f cumpre P3 então \tilde{f}^e também. Se f satisfaz P4, então

$$\begin{aligned} \tilde{f}^e(x, y) + \tilde{f}^e(y, x) &= f(x, y) + f(y, x) - \gamma \langle \nabla \psi(y) - \nabla \psi(x), y - x \rangle \\ &\leq (\theta - \gamma) \langle \nabla \psi(y) - \nabla \psi(x), y - x \rangle \\ &\leq 0. \end{aligned}$$

Ou seja, \tilde{f}^e é monótona. Para finalizar, mostremos P5. Tome $(x^j)_{j \in \mathbb{N}} \subset C$ tal que $\|x^j\| \rightarrow \infty$. Utilizando novamente o fato que f satisfaz P4 temos que para cada $j \in \mathbb{N}$

$$\begin{aligned} \tilde{f}^e(x^j, \bar{x}) &= f(x^j, \bar{x}) + \gamma \langle \nabla \psi(x^j) - \nabla \psi(\bar{x}), \bar{x} - x^j \rangle - \langle e, \bar{x} - x^j \rangle \\ &= f(x^j, \bar{x}) - \gamma \langle \nabla \psi(\bar{x}) - \nabla \psi(x^j), \bar{x} - x^j \rangle - \langle e, \bar{x} - x^j \rangle \\ &\leq -f(\bar{x}, x^j) - (\gamma - \theta) \langle \nabla \psi(\bar{x}) - \nabla \psi(x^j), \bar{x} - x^j \rangle - \langle e, \bar{x} - x^j \rangle \\ &= -\left(f(\bar{x}, x^j) + (\gamma - \theta) \langle \nabla \psi(\bar{x}) - \nabla \psi(x^j), \bar{x} - x^j \rangle + \langle e, \bar{x} - x^j \rangle \right). \end{aligned} \quad (2.2)$$

Logo, como por hipótese f cumpre P6, então temos que a expressão em parênteses em (2.2) é não negativa para k suficientemente grande, em outras palavras, existe $j_0 \in \mathbb{N}$ tal que $\tilde{f}^e(x^j, \bar{x}) \leq 0$ para todo $j \geq j_0$. □

Como consequência, da proposição anterior, veremos que também é possível obter a seguinte proposição a qual estende os resultados dados em [4], para o caso em que a função de regularização é perturbada, e por fim estabelecer as soluções do problema regularizado $EP(\tilde{f}^e, C)$.

Proposição 2.7. *Seja $\bar{x} \in C \subset \text{int}\mathcal{D}$. Suponha que f satisfaz as condições P1-P4 com $\theta < \gamma$. Então as seguintes afirmações são válidas:*

- i) *Se a hipótese P6 é válida então $EP(\tilde{f}^e, C)$ admite pelo menos uma solução.*
- ii) *Se ψ é estritamente convexa então $EP(\tilde{f}^e, C)$ admite no máximo uma solução.*
- iii) *Se a hipótese P6 é válida e ψ é estritamente convexa então $EP(\tilde{f}^e, C)$ tem uma única solução.*

Demonstração. Pela proposição anterior, sabemos que se \tilde{f}^e satisfaz P6, então \tilde{f}^e satisfaz P5. Em consequência pelo Teorema 4.3 de [14], temos que $EP(\tilde{f}^e, C)$ admite solução, o

que mostra (i). Agora suponha que $\text{EP}(\tilde{f}^e, C)$ admite duas soluções x' e x'' . Novamente, pela proposição anterior

$$\begin{aligned} 0 &\leq \tilde{f}^e(x', x'') + \tilde{f}^e(x'', x') \\ &= f(x', x'') + f(x'', x') - \gamma \langle \nabla \psi(x'') - \nabla \psi(x'), x'' - x' \rangle \\ &\leq (\theta - \gamma) \langle \nabla \psi(x'') - \nabla \psi(x'), x'' - x' \rangle \\ &\leq 0, \end{aligned}$$

em que a última desigualdade é verdadeira pois por hipótese $\gamma > \theta$. Portanto, $x'' = x'$ pois ψ é estritamente convexa, mostrando assim, (ii). Finalmente (iii) é válida, pois é uma combinação dos itens (i) e (ii). □

Esta última proposição evidencia a importância da hipótese P6, pois esta é uma condição suficiente para garantir soluções do problema de equilíbrio regularizado com perturbação e . Destacamos que em [15], um resultado análogo foi mostrado considerando a hipótese H5, definida no capítulo anterior. No próximo resultado, veremos que a hipótese P6 é mais fraca que hipótese H5.

Proposição 2.8. *Seja $\bar{x} \in C \subset \text{int}\mathcal{D}$. Suponha que f satisfaz P1-P4. Se ψ admite derivada direcional diferenciável em C e satisfaz H5, então P6 é satisfeita e conseqüentemente, $S(\tilde{f}^e, C) \neq \emptyset$.*

Demonstração. Com efeito, seja $(x^j)_{j \in \mathbb{N}} \subset C$ tal que $\|x^j\| \rightarrow \infty$. Defina $\Delta_j = \langle \nabla \psi(\bar{x}) - \nabla \psi(x^j), \bar{x} - x^j \rangle$. Tomando $\rho \geq 0$, $z \in C$ e levando em consideração a convexidade de ψ temos que

$$\begin{aligned} \Delta_j &= \langle \nabla \psi(\bar{x}) - \nabla \psi(x^j), \bar{x} - x^j \rangle \\ &= \langle \nabla \psi(\bar{x}), \bar{x} - x^j \rangle - \langle \nabla \psi(x^j), \bar{x} - x^j \rangle \\ &\geq \langle \nabla \psi(\bar{x}), \bar{x} - x^j \rangle + [\psi(x^j) - \psi(\bar{x})] \\ &= \|x^j - z\| \left[\frac{\langle \nabla \psi(\bar{x}), \bar{x} - x^j \rangle}{\|x^j - z\|} + \frac{\psi(x^j)}{\|x^j - z\|} - \frac{\psi(\bar{x})}{\|x^j - z\|} \right] \\ &= \|x^j - z\| g(x^j), \end{aligned} \tag{2.3}$$

$$\text{em que } g(x^j) = \frac{\langle \nabla \psi(\bar{x}), \bar{x} - x^j \rangle}{\|x^j - z\|} + \frac{\psi(x^j)}{\|x^j - z\|} - \frac{\psi(\bar{x})}{\|x^j - z\|}.$$

Por outro lado, utilizando o fato que o subdiferencial de $f(\bar{x}, \cdot)$ existe, então para

$z \in C$ existe $v \in \partial f(\bar{x}, z)$ tal que

$$\begin{aligned} f(\bar{x}, x^j) + \langle e, \bar{x} - x^j \rangle &\geq f(\bar{x}, z) + \langle v, x^j - z \rangle + \langle e, \bar{x} - x^j \rangle \\ &= f(\bar{x}, z) - \langle v, z \rangle + \langle e, \bar{x} \rangle + \langle v - e, x^j \rangle \\ &\geq A - B \|x^j\| \\ &\geq A - B (\|x^j - z\| + \|z\|), \end{aligned} \quad (2.4)$$

em que, $A = f(\bar{x}, z) - \langle v, z \rangle + \langle e, \bar{x} \rangle$ e $B = \|v - e\|$. Assim, combinando as desigualdades (2.3) e (2.4) obtemos que

$$\begin{aligned} \liminf_{j \rightarrow \infty} [f(\bar{x}, x^j) + (\gamma - \theta)\Delta_j] &\geq \liminf_{j \rightarrow \infty} A - B (\|x^j - z\| + \|z\|) + (\gamma - \theta) \|x^j - z\| g(x^j) \\ &= \liminf_{j \rightarrow \infty} A + (\gamma - \theta) \|x^j - z\| \left[g(x^j) - \frac{B}{\gamma - \theta} \left(1 + \frac{\|z\|}{\|x^j - z\|} \right) \right]. \end{aligned} \quad (2.5)$$

Para que a propriedade P6 seja satisfeita, basta observar que a expressão em colchete na desigualdade (2.5) é positiva para j suficientemente grande. Isto evidentemente é satisfeito pois para j suficientemente grande, o termo $1 + \frac{\|z\|}{\|x^j - z\|}$ é limitado e $g(x^j) \rightarrow \infty$ devido a que ψ cumpre H5. □

Esta última proposição estabelece que a classe de funções que satisfazem H5, concideradas em [15], são automaticamente de funções que cumprem a propriedade P6. Um exemplo claro disto, são as funções 1-coercivas, isto é, $\lim_{\|x\| \rightarrow \infty} \frac{\psi(x)}{\|x\|} = +\infty$, entre as quais temos a função quadrática $\psi(x) = \frac{1}{2} \|x\|^2$ e Kullback–Leibler $\psi(x) = x \log x - x + 1$. Assim, conhecendo uma classe de funções para as quais garantimos solução do problema regularizado com perturbação e , uma pergunta natural é saber se existem funções que satisfazem P6 mas não cumprem H5. Tal questão é respondida na seguinte observação.

Observação 2.9. *A condição P6 é estritamente mais fraca que H5 pois, se consideramos $\psi : \mathbb{R} \rightarrow \mathbb{R}$ dada por $\psi(x) = x - \log x + 1$ e $C \subset \mathcal{D} = (0, +\infty)$ temos que a condição P6 é satisfeita desde que $\gamma - \theta$ seja suficientemente grande mas H5 não é satisfeita. Com efeito, seja $(x^j) \subset \mathcal{D}$, tal que $x^j \rightarrow \infty$. Evidentemente, $\lim_{x^j \rightarrow \infty} \frac{\psi(x^j)}{x^j - z} < \infty$ e pela Proposição 1.11 obtemos que H5 não é satisfeita. Por outro lado, realizando um procedimento similar ao considerado em (2.4) na proposição anterior obtemos:*

$$\begin{aligned} f(\bar{x}, x^j) + (\gamma - \theta)\Delta_j &\geq A - Bx^j + (\gamma - \theta) [(\psi'(\bar{x}) - \psi'(x^j))(\bar{x} - x^j)] \\ &= A - Bx^j + (\gamma - \theta) \frac{(\bar{x} - x^j)^2}{\bar{x}x^j} \\ &= A + x^j \left[\frac{(\gamma - \theta)}{\bar{x}} \left(1 - \frac{\bar{x}}{x^j} \right)^2 - B \right]. \end{aligned} \quad (2.6)$$

Em que $A = f(\bar{x}, z) - \langle v, z \rangle + \langle e, \bar{x} \rangle$ e $B = \|v - e\|$. Note que o sinal da expressão em colchete (2.6) é positivo quando $\gamma - \theta$ é suficiente grande. Logo, tomando limite inferior nesta última desigualdade obtemos que a propriedade P6 é satisfeita.

A exposição anterior foi realizada para garantir uma boa definição dos algoritmos que descreveremos a seguir, os quais foram introduzidos em [15]. Assumiremos daqui em diante, que $C \subset \text{int}D$ e que existe uma função ψ que satisfaz as hipóteses H1-H4 e P6.

Algoritmo IPPBPM: Método de ponto proximal inexato com projeção para $\text{EP}(f, C)$.

Passo 1: Considere a sequência limitada $\gamma_j \subset \mathbb{R}_{++}$, uma tolerância para o erro $\sigma \in (0, 1)$ e um ponto inicial $x^0 \in C$.

Passo 2: Dado o ponto x^j , encontre o par $(\tilde{x}^j, e^j) \in \mathbb{R}^n \times \mathbb{R}^n$ tal que \tilde{x}^j resolve o problema $\text{EP}(\tilde{f}_j^e, C)$, com \tilde{f}_j^e dada por

$$\tilde{f}_j^e(x, y) = f(x, y) + \gamma_j \langle \nabla\psi(x) - \nabla\psi(x^j), y - x \rangle - \langle e^j, y - x \rangle, \quad (2.7)$$

e e^j satisfazendo

$$\|e^j\| \leq \begin{cases} \sigma \gamma^j D_\psi(\tilde{x}^j, x^j) & \text{se } \|x^j - \tilde{x}^j\| < 1 \\ \sigma \gamma^j \nu_\psi(x^j, 1) & \text{se } \|x^j - \tilde{x}^j\| \geq 1. \end{cases} \quad (2.8)$$

Passo 3: Defina $v^j = \gamma^j [\nabla\psi(x^j) - \nabla\psi(\tilde{x}^j)] + e^j$. Se $v^j = 0$ ou $x^j = \tilde{x}^j$ então pare. Caso contrário, considere o hiperplano $H_j = \{x \in D \mid \langle v^j, x - \tilde{x}^j \rangle = 0\}$ e tome

$$x^{j+1} = \arg \min_{x \in H_j} D_\psi(x, x^j). \quad (2.9)$$

Algoritmo IPPEM: Método de ponto proximal inexato com passo extragradiente para $\text{EP}(f, C)$.

Passo 1: Considere a sequência limitada $\gamma_j \subset \mathbb{R}_{++}$, uma tolerância para o erro $\sigma \in (0, 1)$ e um ponto inicial $x^0 \in C$.

Passo 2: Dado o ponto x^j , encontre o par $(\tilde{x}^j, e^j) \in \mathbb{R}^n \times \mathbb{R}^n$ tal que $\tilde{x}^j \in C$ resolve o problema $\text{EP}(\tilde{f}_j^e, C)$, com \tilde{f}_j^e dada por

$$\tilde{f}_j^e(x, y) = f(x, y) + \gamma_j \langle \nabla\psi(x) - \nabla\psi(x^j), y - x \rangle - \langle e^j, y - x \rangle, \quad (2.10)$$

e e^j satisfaz

$$D_\psi(\tilde{x}^j, (\nabla\psi)^{-1}[\nabla\psi(\tilde{x}^j) - \gamma_j^{-1}e^j]) \leq \sigma D_\psi(\tilde{x}^j, x^j), \quad (2.11)$$

Passo 3: Se $\tilde{x}^j = x^j$ então pare. Caso contrário, tome

$$x^{j+1} = (\nabla\psi)^{-1}[\nabla\psi(\tilde{x}^j) - \gamma_j^{-1}e^j]. \quad (2.12)$$

As diferenças entre os Algoritmos IPPBPM e IPPEM são evidentes nos passos 2 e 3. A ideia no Passo 2 é praticamente a mesma para ambos algoritmos, diferenciando-se apenas na condição que e^j deve cumprir. Agora, o Passo 3 em cada um dos algoritmos se divide em duas etapas, a primeira consiste em verificar se o critério de parada é atingido e caso isto não ocorra, cada algoritmo utiliza uma forma específica para atualizar a próxima iteração. Note que estes passos estão bem definidos para ψ satisfazendo as hipóteses H1-H4 e pela Proposição 1.10.

Por outro lado, as diferenças presentes nos algoritmos IPPBPM e IPPEM são evidentemente eliminadas quando $e^j = 0$, para todo $j \in \mathbb{N}$, pois neste caso vemos que o Passo 2 de cada algoritmo se reduz a encontrar uma solução do subproblema de equilíbrio $EP(\tilde{f}_j, C)$, com \tilde{f}_j dada como em (2.10), com $e^j = 0$. Analogamente, no Passo 3, podemos observar que se $e^j = 0$ para todo $j \in \mathbb{N}$, então de (2.9) e (2.12) obtemos que a próxima iteração é simplificada para a igualdade $x^{j+1} = \tilde{x}^j$, em que \tilde{x}^j é a solução obtida previamente do Passo 2. Com estes argumentos podemos definir o método de ponto proximal Exato.

Algoritmo EPPM (Método de Ponto proximal Exato para $EP(f, C)$)

Passo 1: Considere a sequência limitada $\gamma_j \subset \mathbb{R}_{++}$ e um ponto inicial $x^0 \in C$.

Passo 2: Dado o ponto x^j , encontre $\tilde{x}^j \in \mathbb{R}^n$ tal que \tilde{x}^j resolve o problema $EP(\tilde{f}_j, C)$, em que \tilde{f}_j é definida como

$$\tilde{f}_j(x, y) = f(x, y) + \gamma_j \langle \nabla \psi(x) - \nabla \psi(x^j), y - x \rangle. \quad (2.13)$$

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

$$x^{j+1} = \tilde{x}^j.$$

A análise de convergência para cada algoritmo é feita com detalhes em [15], considerando algumas hipóteses diferentes às consideradas aqui. Para uma maior completude do texto, vamos realizar uma análise similar, porém, considerando a hipótese P6.

Proposição 2.10. *Considere o problema $EP(f, C)$ e suponha que uma dentre as hipóteses P_4^* e P_4' é satisfeita. Se P2 é válida, então $f(y, x^*) \leq 0$ para todo $y \in C$ e $x^* \in S(f, C)$.*

Demonstração. Ver Proposição 5.2 de [15]. □

Lema 2.11. *Considere ψ satisfazendo H2. Se $(x^j)_{j \in \mathbb{N}}$, $(\tilde{x}^j)_{j \in \mathbb{N}}$ e $(\gamma_j)_{j \in \mathbb{N}}$ são geradas pelo algoritmo IPPBPM então para $\sigma \in (0, 1)$ vale que*

$$\|e^j\| \|x^j - \tilde{x}^j\| \leq \sigma \gamma_j D_\psi(\tilde{x}^j, x^j) \leq \gamma_j D_\psi(\tilde{x}^j, x^j), \quad (2.14)$$

para todo $j \in \mathbb{N}$.

Demonstração. Suponha inicialmente que $\|x^j - \tilde{x}^j\| < 1$. Logo por (2.8) obtemos que $\|e^j\| \|x^j - \tilde{x}^j\| \leq \|e^j\| \leq \sigma \gamma_j D_\psi(\tilde{x}^j, x^j)$. Agora, se $\|x^j - \tilde{x}^j\| \geq 1$, então pela definição de convexidade total e a Proposição 1.5, obtemos que:

$$\sigma \gamma_j D_\psi(\tilde{x}^j, x^j) \geq \sigma \gamma_j \nu_\psi(x^j, \|\tilde{x}^j - x^j\|) \geq \sigma \gamma_j \|\tilde{x}^j - x^j\| \nu_\psi(x^j, 1) \geq \|x^j - \tilde{x}^j\| \|e^j\|.$$

O que mostra a primeira desigualdade em (2.14). Para a segunda desigualdade, basta considerar que $\sigma \in (0, 1)$. □

Proposição 2.12. *Suponha que f satisfaz as hipóteses P1-P4 e uma dentre as hipóteses P4* e P4'. Tome $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfazendo H1-H3 e uma sequência exógena (externa ao método) $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma})$, em que θ é a constante de submonotonicidade de P4. Se $(x^j)_{j \in \mathbb{N}}$ é uma sequência gerada pelo algoritmo IPPBPM e $EP(f, C)$ admite solução, então*

i) *Para todo $x^* \in S(f, C)$, $D_\psi(x^*, x^j)$ é não crescente e convergente.*

ii) *$(x^j)_{j \in \mathbb{N}}$ é limitada.*

iii) *$(x^{j+1} - x^j)_{j \in \mathbb{N}}$ converge para zero.*

iv) *$(\gamma_j^{-1} e^j)_{j \in \mathbb{N}}$ é limitada.*

v) *$(x^j - \tilde{x}^j)_{j \in \mathbb{N}}$ converge para zero.*

vi) *$(v^j)_{j \in \mathbb{N}}$ converge para zero.*

Demonstração. (i) Seja $x^* \in S(f, C)$. Note que, se $\tilde{x}^j \in S(f_j^e, C)$, então para qualquer $y \in C$ obtemos que

$$\begin{aligned} f_j^e(\tilde{x}^j, y) \geq 0 &\Leftrightarrow f(\tilde{x}^j, y) + \langle \gamma_j [\nabla \psi(\tilde{x}^j) - \nabla \psi(x^j)] - e^j, y - x^j \rangle \geq 0 \\ &\Leftrightarrow f(\tilde{x}^j, y) \geq \langle \gamma_j [\nabla \psi(x^j) - \nabla \psi(\tilde{x}^j)] + e^j, y - x^j \rangle \\ &\Leftrightarrow f(\tilde{x}^j, y) \geq \langle v^j, y - x^j \rangle. \end{aligned} \tag{2.15}$$

Em particular tomando $y = x^*$ na desigualdade anterior, obtemos que $f(\tilde{x}^j, x^*) \geq \langle v^j, x^* - x^j \rangle$. Agora, como f satisfaz P4* ou P4' então da Proposição 2.10 obtemos que $0 \geq f(\tilde{x}^j, x^*) \geq \langle v^j, x^* - x^j \rangle$, isto é $x^* \in H_j^-$, com H_j^- dado como na Proposição 1.6. Pela definição de v^j e Proposição 1.3 (iv) obtemos

$$\begin{aligned} \langle v^j, x^j - \tilde{x}^j \rangle &= \gamma_j \langle \nabla \psi(x^j) - \nabla \psi(\tilde{x}^j), x^j - \tilde{x}^j \rangle + \langle e^j, x^j - \tilde{x}^j \rangle \\ &= \gamma_j (D_\psi(\tilde{x}^j, x^j) + D_\psi(x^j, \tilde{x}^j)) + \langle e^j, x^j - \tilde{x}^j \rangle \\ &\geq \gamma_j D_\psi(x^j, \tilde{x}^j) + (\gamma_j D_\psi(\tilde{x}^j, x^j) - \|e^j\| \|x^j - \tilde{x}^j\|) \\ &\geq \gamma_j D_\psi(x^j, \tilde{x}^j), \end{aligned} \tag{2.16}$$

sendo a última desigualdade válida pelo Lema 2.11. Note que de (2.16) concluímos que $\langle v^j, x^j - \tilde{x}^j \rangle \geq 0$, ou seja, $x^j \in H_j^+$ com H_j^+ dado como na Proposição 1.6. Assim, escolhendo $\tilde{y} = \tilde{x}^j$, $x = x^j$, $v = v^j$, $\bar{x} = x^*$ e $z = x^{j+1}$ na Proposição 1.6 obtemos que

$$D_\psi(x^*, x^j) \geq D_\psi(x^*, x^{j+1}) + D_\psi(x^{j+1}, x^j) \geq D_\psi(x^*, x^{j+1}) \quad (2.17)$$

o que mostra que $(D_\psi(x^*, x^j))_{j \in \mathbb{N}}$ é uma sequência monótona, limitada e, portanto, convergente.

(ii) Do item anterior, note que $(x^j)_{j \in \mathbb{N}} \subset \{y \in \mathbb{R}^n : D_\psi(x^*, y) \leq D_\psi(x^*, x^0)\}$ e portanto, pela hipótese H1 obtemos que $(x^j)_{j \in \mathbb{N}}$ é limitada.

(iii) De (2.17) obtemos que

$$D_\psi(x^*, x^j) - D_\psi(x^*, x^{j+1}) \geq D_\psi(x^{j+1}, x^j) \geq 0. \quad (2.18)$$

Como por (i), $(D_\psi(x^*, x^j))_{j \in \mathbb{N}}$ converge então

$$\lim_{j \rightarrow \infty} D_\psi(x^{j+1}, x^j) = 0 \quad (2.19)$$

Em consequência, pela Proposição 1.7 concluímos que (iii) é válido.

(iv) Seja $A = \{y \in \mathbb{R}^n : \|x^j - y\| \leq 1, \text{ para algum } j\}$. Note que para $y \in A$, $\|y\| \leq \|y - x^j\| + \|x^j\| \leq 1 + \|x^j\|$ e como $(x^j)_{j \in \mathbb{N}}$ é limitada então o conjunto A é limitado. Logo, pela Proposição 1.9 concluímos que D_g é limitada em $A \times A$, digamos por $\xi > 0$. Do Passo 2 do algoritmo IPPBPM sabemos se $\|x^j - \tilde{x}^j\| \leq 1$ então $x^j, \tilde{x}^j \in A$ e portanto, $\|\gamma_j^{-1} e^j\| \leq D_\psi(\tilde{x}^j, x^j) \leq \xi$. Agora se $\|x^j - \tilde{x}^j\| > 1$, tomando $y = x^j + \frac{x^j - \tilde{x}^j}{\|x^j - \tilde{x}^j\|}$, obtemos que $\|y - x^j\| = 1$ e portanto, $y \in A$. Por conseguinte, $\|\gamma_j^{-1} e^j\| \leq \nu_\psi(x^j, 1) \leq D_\psi(y, x^j) \leq \xi$, mostrando que em qualquer caso $(\gamma_j^{-1} e^j)_{j \in \mathbb{N}}$ é limitada.

(v) Note inicialmente que:

$$0 \leq |\langle \gamma_j^{-1} e^j, x^{j+1} - x^j \rangle| \leq \|\gamma_j^{-1} e^j\| \|x^{j+1} - x^j\|. \quad (2.20)$$

Logo, de (iii) e (iv) segue que

$$\lim_{j \rightarrow \infty} \langle \gamma_j^{-1} e^j, x^{j+1} - x^j \rangle = 0. \quad (2.21)$$

Por outro lado, pela definição de distância de Bregman e utilizando o fato que

$x^{j+1} \in H_j$, com H_j dado como na Proposição 1.6 obtemos que

$$\begin{aligned} D_\psi(x^{j+1}, x^j) - D_\psi(x^{j+1}, \tilde{x}^j) - D_\psi(\tilde{x}^j, x^j) &= \langle \nabla\psi(x^j) - \nabla\psi(\tilde{x}^j), \tilde{x}^j - x^{j+1} \rangle \\ &= \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^{j+1} \rangle - \langle \gamma_j^{-1} e^j, \tilde{x}^j - x^{j+1} \rangle \\ &= \langle \gamma_j^{-1} e^j, x^{j+1} - \tilde{x}^j \rangle \\ &= \langle \gamma_j^{-1} e^j, x^{j+1} - x^j \rangle + \langle \gamma_j^{-1} e^j, x^j - \tilde{x}^j \rangle. \end{aligned}$$

Logo, obtemos que

$$\begin{aligned} D_\psi(x^{j+1}, x^j) - D_\psi(x^{j+1}, \tilde{x}^j) &= \langle \gamma_j^{-1} e^j, x^{j+1} - x^j \rangle + D_\psi(\tilde{x}^j, x^j) + \langle \gamma_j^{-1} e^j, x^j - \tilde{x}^j \rangle \\ &\geq \langle \gamma_j^{-1} e^j, x^{j+1} - x^j \rangle + \gamma_j^{-1} [\gamma_j D_\psi(\tilde{x}^j, x^j) - \|e^j\| \|x^j - \tilde{x}^j\|] \\ &\geq \langle \gamma_j^{-1} e^j, x^{j+1} - x^j \rangle, \end{aligned} \quad (2.22)$$

em que as últimas desigualdades são obtidas utilizando a desigualdade de Cauchy-Schwarz e o Lema 2.11. Agora, tomando limite na desigualdade (2.22) e levando em conta (2.19) e (2.21), concluímos que

$$\lim_{j \rightarrow \infty} D_\psi(x^{j+1}, \tilde{x}^j) = 0. \quad (2.23)$$

Logo, pela Proposição 1.7 concluímos que $\|x^{j+1} - \tilde{x}^j\|$, converge para zero e finalmente pela desigualdade triangular obtemos que

$$\lim_{j \rightarrow \infty} \|x^j - \tilde{x}^j\| \leq \lim_{j \rightarrow \infty} \|x^j - x^{j+1}\| + \|x^{j+1} - \tilde{x}^j\| = 0, \quad (2.24)$$

o que mostra (v).

(vi) Como $D_\psi(x^{j+1}, x^j)$, $D_\psi(x^{j+1}, \tilde{x}^j)$ e $\langle \gamma_j^{-1} e^j, x^{j+1} - x^j \rangle$ convergem para zero, então da igualdade

$$D_\psi(x^{j+1}, x^j) - D_\psi(x^{j+1}, \tilde{x}^j) - D_\psi(\tilde{x}^j, x^j) = \langle \nabla\psi(x^j) - \nabla\psi(\tilde{x}^j), \tilde{x}^j - x^{j+1} \rangle,$$

concluímos que $D_\psi(\tilde{x}^j, x^j)$ também converge para zero. Em consequência, para j suficientemente grande obtemos que $\|x^j - \tilde{x}^j\| < 1$ e por (2.8), concluímos que $(e^j)_{j \in \mathbb{N}}$ converge para zero.

Por outro lado, vimos no item anterior que $(x^j - \tilde{x}^j)_{j \in \mathbb{N}}$ converge para zero e da hipótese H3, concluímos que $\nabla\psi(x^j) - \nabla\psi(\tilde{x}^j)$ também converge para zero. Em consequência, pela definição de v^j no Passo 3 de IPPBPM, obtemos que $(v^j)_{j \in \mathbb{N}}$ converge para zero.

□

De forma análoga pode ser mostrado para IPPEM as seguintes proposições.

Proposição 2.13. *Considere $EP(f, C)$ com f satisfazendo P1-P3, P6 e uma dentre as hipóteses $P4^*$ e $P4'$. Tome $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfazendo H1-H4. Sejam*

$(x^j)_{j \in \mathbb{N}}, (\tilde{x}^j)_{j \in \mathbb{N}}, (\gamma_j)_{j \in \mathbb{N}}$ e σ como no Algoritmo IPPEM. Então para qualquer $x^* \in S(f, C)$,

$$D_\psi(x^*, x^{j+1}) \leq D_\psi(x^*, x^j) - \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle - (1 - \sigma) D_\psi(\tilde{x}^j, x^j) \leq D_\psi(x^*, x^j), \quad (2.25)$$

em que v^j é dado como no algoritmo IPPBPM.

Demonstração. Consideremos o vetor auxiliar

$$v^j = \gamma_j [\nabla \psi(x^j) - \nabla \psi(\tilde{x}^j)] + e^j. \quad (2.26)$$

Agora, do Passo 3 do algoritmo IPPEM obtemos que $\nabla \psi(x^{j+1}) = \nabla \psi(\tilde{x}^j) - \gamma_j^{-1} e^j$ e assim

$$\gamma_j^{-1} e^j = \nabla \psi(\tilde{x}^j) - \nabla \psi(x^{j+1}). \quad (2.27)$$

Logo, combinando (2.26) e (2.27) obtemos que

$$\gamma_j^{-1} v^j + \nabla \psi(x^{j+1}) - \nabla \psi(x^j) = 0. \quad (2.28)$$

Por outro lado, note que de (2.12) e (2.11) vale que $D_\psi(\tilde{x}^j, x^{j+1}) \leq \sigma D_\psi(\tilde{x}^j, x^j)$. Assim, para $x^* \in S(f, C)$

$$\begin{aligned} D_\psi(x^*, x^{j+1}) &= D_\psi(x^*, x^j) + \langle \nabla \psi(x^j) - \nabla \psi(x^{j+1}), x^* - \tilde{x}^j \rangle + D_\psi(\tilde{x}^j, x^{j+1}) - D_\psi(\tilde{x}^j, x^j) \\ &= D_\psi(x^*, x^j) + \langle \gamma_j^{-1} v^j, x^* - \tilde{x}^j \rangle + D_\psi(\tilde{x}^j, x^{j+1}) - D_\psi(\tilde{x}^j, x^j) \\ &\leq D_\psi(x^*, x^j) - \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle - (1 - \sigma) D_\psi(\tilde{x}^j, x^j) \\ &\leq D_\psi(x^*, x^j) - \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle, \end{aligned} \quad (2.29)$$

em que usamos, (2.28), $\sigma \in (0, 1)$ e $D_\psi(\cdot, \cdot) \geq 0$. Agora como $\tilde{x}^j \in S(f_j^e, C)$ então $f_j^e(\tilde{x}^j, y) \geq 0$, para todo $y \in C$. Em particular, se $y = x^*$ então $f(\tilde{x}^j, x^*) \geq \langle v^j, x^* - \tilde{x}^j \rangle$ e pela Proposição 2.10 concluímos que $\langle v^j, x^* - \tilde{x}^j \rangle \leq 0$. Utilizando esta última desigualdade em (2.29) concluímos que $D_\psi(x^*, x^{j+1}) \leq D_\psi(x^*, x^j) - \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle - (1 - \sigma) D_\psi(\tilde{x}^j, x^j) \leq D_\psi(x^*, x^j)$. □

Proposição 2.14. *Considere $EP(f, C)$ com f satisfazendo P1-P4, P6 e uma dentre as hipóteses $P4^*$ e $P4'$. Tome $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfazendo H1-H4 e uma sequência exógena (externa ao método) $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma})$, em que θ é constante de submonotonicidade em P4. Seja $(x^j)_{j \in \mathbb{N}}$ uma sequência gerada pelo Algoritmo IPPEM. Se $EP(f, C)$ tem solução, então*

i) Para todo $x^* \in S(f, C)$, $(D_\psi(x^*, x^j))_{j \in \mathbb{N}}$ é não crescente e convergente.

ii) $(x^j)_{j \in \mathbb{N}}$ é limitada.

iii) $\sum_{j=0}^{\infty} \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle < \infty$, em que v^j é dado como no algoritmo IPPBMP.

iv) $\sum_{j=0}^{\infty} D_\psi(\tilde{x}^j, x^j) < \infty$.

v) $\sum_{j=0}^{\infty} D_\psi(\tilde{x}^j, x^{j+1}) < \infty$.

vi) $(\tilde{x}^j - x^j)_{j \in \mathbb{N}}$ converge para zero e conseqüentemente $(\tilde{x}^j)_{j \in \mathbb{N}}$ é limitada.

vii) $(x^{j+1} - x^j)_{j \in \mathbb{N}}$ converge para zero.

viii) $(v^j)_{j \in \mathbb{N}}$ converge para zero.

Demonstração. Seja $x^* \in S(f, C)$. Note inicialmente que pela Proposição 2.13, a sequência $(D_\psi(x^*, x^j))_{j \in \mathbb{N}}$ é não crescente e não negativa e conseqüentemente convergente o que mostra (i). Agora, como $(x^j)_{j \in \mathbb{N}}$ está contida no conjunto de nível $D_\psi(x^*, \cdot)$, então a hipótese H1 implica que esta sequência é limitada, ou seja, vale (ii). Mostremos (iii)-(v). Para isto, basta observar que da proposição anterior obtemos que

$$0 \leq \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle + (1 - \sigma) D_\psi(\tilde{x}^j, x^j) \leq D_\psi(x^*, x^j) - D_\psi(x^*, x^{j+1})$$

para todo $j \in \mathbb{N}$. Agora, como

$$\sum_{j=0}^{\infty} \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle + (1 - \sigma) D_\psi(\tilde{x}^j, x^j) \leq \sum_{j=1}^{\infty} D_\psi(x^*, x^j) - D_\psi(x^*, x^{j+1}),$$

são séries de números reais e como a série $\sum_{j=1}^{\infty} D_\psi(x^*, x^j) - D_\psi(x^*, x^{j+1})$ é convergente segue que

$$\sum_{j=0}^{\infty} \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle + (1 - \sigma) D_\psi(\tilde{x}^j, x^j) \tag{2.30}$$

converge. Além disso, note que cada termo de (2.30) é não negativo portanto as séries $\sum_{j=0}^{\infty} \gamma_j^{-1} \langle v^j, \tilde{x}^j - x^* \rangle$ e $\sum_{j=0}^{\infty} D_\psi(\tilde{x}^j, x^j)$ convergem, mostrando (iii) e (iv). Note que (v) segue pelo fato de que $D_\psi(x^*, x^{j+1}) \leq D_\psi(x^*, x^j)$, para todo $j \in \mathbb{N}$, então $\sum_{j=0}^{\infty} D_\psi(x^*, x^{j+1})$ converge. Da mesma maneira, (vi) e (vii) são válidas quando utilizamos $\lim_{j \rightarrow \infty} D_\psi(\tilde{x}^j, x^j) = \lim_{j \rightarrow \infty} D_\psi(\tilde{x}^j, x^{j+1}) = 0$, na Proposição 1.7. Finalmente, para mostrar (viii), note que por (2.28) temos que $\gamma_j v^j = \nabla \psi(x^j) - \nabla \psi(\tilde{x}^j)$ para todo $j \in \mathbb{N}$ e como $(x^j - \tilde{x}^j)_{j \in \mathbb{N}}$ converge para zero então de H3 obtemos que $\nabla \psi(x^j) - \nabla \psi(\tilde{x}^j)$ converge para zero. Em conseqüência, $(\gamma_j v^j)_{j \in \mathbb{N}}$ converge para zero e como $(\gamma_j)_{j \in \mathbb{N}}$ é limitada então $(v^j)_{j \in \mathbb{N}}$ converge para zero como desejávamos. \square

Com os resultados expostos anteriormente podemos considerar o seguinte teorema de convergência.

Teorema 2.15. *Considere $EP(f, C)$. Suponha que f satisfaz P1-P4 e P6, e que uma dentre as hipóteses $P4^*$ e $P4'$ é válida. Tome $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfazendo H1-H4 e uma seqüência exógena (externa ao método) $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma}]$, em que $\bar{\gamma} > 0$ e θ é constante de submonotonicidade de $P4$. Seja $(x^j)_{j \in \mathbb{N}}$ a seqüência gerada pelo algoritmo IPPBPM (ou IPPEM). Se $EP(f, C)$ tem solução, então*

i) A seqüência $(\tilde{x}^j)_{j \in \mathbb{N}}$ resolve assintoticamente $EP(f, C)$, isto é, $\liminf_{j \rightarrow \infty} f(x^j, y) \geq 0$, para todo $y \in C$.

ii) Todo ponto de acumulação da seqüência $(x^j)_{j \in \mathbb{N}}$ é solução de $EP(f, C)$.

Demonstração. Veja Teoremas 5.5 e 5.8 de [15]. □

Para finalizar a análise de convergência destes algoritmos, vamos estabelecer alguns resultados adicionais que usaremos no próximo capítulo. Considere a seguinte notação. Para cada $x \in C$, $F_x : C \rightarrow \mathbb{R} \cup \{+\infty\}$ é dado como

$$F_x(y) = f(x, y). \quad (2.31)$$

Note que para cada $x \in C$, a função F_x é convexa, devido à propriedade P2. Usaremos esta notação no próximo resultado o qual relaciona as soluções de um problema de equilíbrio com um problema de otimização.

Proposição 2.16. *Considere o problema $EP(f, C)$. As seguintes afirmações são equivalentes:*

a) $x^ \in S(f, C)$*

b) x^ minimiza a função F_{x^*} no conjunto C , com F_x dada como em (2.31).*

Demonstração. Primeiro suponha que $x^* \in S(f, C)$. Pela propriedade P1 e (2.31) vale que

$$F_{x^*}(y) = f(x^*, y) \geq 0 = f(x^*, x^*) = F_{x^*}(x^*),$$

para todo $y \in C$. Reciprocamente, se (b) vale, então usando novamente P1 e (2.31) obtemos

$$f(x^*, y) = F_{x^*}(y) \geq F_{x^*}(x^*) = f(x^*, x^*) = 0,$$

para todo $y \in C$. □

Entre as funções ψ que satisfazem as propriedades H1-H4 e P6 podemos encontrar funções de zona coerciva, isto é,

H6: Se $(y^k)_{k \in \mathbb{N}} \subset \text{int}\mathcal{D}$ e $\lim_{k \rightarrow \infty} y^k = y \in \partial\mathcal{D}$, em que $\partial\mathcal{D}$ denota a fronteira do domínio de \mathcal{D} , então $\lim_{k \rightarrow \infty} \langle \nabla \psi(y^k), x - y^k \rangle = -\infty$ para todo $x \in \text{int}\mathcal{D}$.

Na literatura, as funções que satisfazem esta propriedade são chamadas de coercivas na fronteira. Este fato é importante pois a seguinte proposição estabelece que os pontos gerados pelos algoritmos de ponto proximal com ψ satisfazendo H6, estão contidos no interior do domínio de ψ .

Proposição 2.17. *Considere $EP(f, \mathcal{D})$ em que \mathcal{D} é um domínio convexo. Sejam f satisfazendo P1-P4 e P6 e também uma das hipóteses $P4^*$ ou $P4'$, $\psi : \mathcal{D} \rightarrow \mathbb{R}$ satisfazendo H1-H2 e H6 e uma sequência exógena (externa ao método) $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma})$. Se $(\hat{x}^j)_{j \in \mathbb{N}}$ é uma sequência gerada pelo Passo 2 do algoritmo IPPEM (ou IPPBMP) então $(\hat{x}^j)_{j \in \mathbb{N}} \subset \text{int}\mathcal{D}$.*

Demonstração. Sabemos da Proposição 2.16, que \hat{x}^j é o único elemento em $x \in \mathcal{D}$ tal que

$$e^j + \gamma_j \nabla \psi(x^j) \in \partial(F_x + \gamma_j \psi)(x). \quad (2.32)$$

Assim, mostraremos que $\partial(F + \lambda \psi) = \emptyset$ para todo $x \in \partial \mathcal{D}$ o que por (2.32) implica que x não pode ser um elemento da fronteira de \mathcal{D} . Considere x na fronteira de \mathcal{D} e suponha que existe $\xi \in \partial(F + \lambda \psi)$. Tome $z \in \text{int}\mathcal{D}$ e defina a sequência $(y^\ell)_{\ell \in \mathbb{N}}$ dada por

$$y^\ell = (1 - \varepsilon_\ell)x + \varepsilon_\ell z, \quad (2.33)$$

em que $\varepsilon_\ell \rightarrow 0$. Note que pela convexidade de \mathcal{D} , $(y^\ell)_{\ell \in \mathbb{N}} \subset \text{int}\mathcal{D}$ tal que $y^\ell \rightarrow x \in \partial \mathcal{D}$. Por outro lado, note que

$$\begin{aligned} \varepsilon_\ell \langle \xi, z - x \rangle &= \langle \xi, y^\ell - x \rangle \\ &\leq F_{y^\ell}(y^\ell) + \gamma_j \psi(y^\ell) - (F_{y^\ell}(x) + \gamma_j \psi(x)) \\ &= F_{y^\ell}(y^\ell) - F_{y^\ell}(x) + \gamma_j (\psi(y^\ell) - \psi(x)) \\ &\leq F_{y^\ell}(y^\ell) - F_{y^\ell}(x) + \gamma_j \langle \nabla \psi(y^\ell), y^\ell - x \rangle \\ &\leq (1 - \varepsilon_\ell) F_{y^\ell}(x) + \varepsilon_\ell F_{y^\ell}(z) - F_{y^\ell}(x) + \gamma_j \langle \nabla \psi(y^\ell), y^\ell - x \rangle \\ &= \varepsilon_\ell (F_{y^\ell}(z) - F_{y^\ell}(x)) + \gamma_j \langle \nabla \psi(y^\ell), \varepsilon_\ell (z - x) \rangle \\ &= \varepsilon_\ell (F_{y^\ell}(z) - F_{y^\ell}(x)) + \gamma_j \frac{\varepsilon_\ell}{1 - \varepsilon_\ell} \langle \nabla \psi(y^\ell), z - y^\ell \rangle, \end{aligned} \quad (2.34)$$

em que utilizamos (2.33), a definição de subdiferenciabilidade de $F_{y^\ell} + \gamma_j \psi$, a convexidade de ψ e $F_{y^\ell}(\cdot)$ para obter as desigualdades acima. Agora, observe que de (2.34) obtemos que:

$$\frac{1 - \varepsilon_\ell}{\gamma_j} (F_{y^\ell}(x) - F_{y^\ell}(z) + \langle \xi, (z - x) \rangle) \leq \langle \nabla \psi(y^\ell), z - y^\ell \rangle.$$

Assim, tomando limite na desigualdade acima obtemos que, o lado esquerdo tem limite finito enquanto que o lado direito da desigualdade tende $-\infty$, pois ψ cumpre a hipótese H6, o qual é absurdo. Este absurdo implica que $x \notin \partial \mathcal{D}$.

□

Assim, se consideramos as hipóteses H1-H4 e P6 em todo o domínio da ψ , a Proposição 2.17 garante que isto pode ser feito, desde que as hipóteses estejam bem definidas. No próximo capítulo, apresentaremos métodos de Lagrangeano aumentado para problemas de equilíbrio. Como no caso de otimização, estes métodos buscam resolver o problema $EP(f, C)$ através de uma sequência de subproblemas de equilíbrio irrestritos. Também, vamos estabelecer a relação entre os algoritmos de ponto proximal e os algoritmos de Lagrangeano aumentado para resolver problemas de equilíbrio.

Capítulo 3

Lagrangeano aumentado para problemas de equilíbrio

Este capítulo e o próximo contém as principais contribuições desta tese. Inicialmente, definimos uma extensão dos métodos de Lagrangeano aumentado para problemas de equilíbrio introduzidos em [16], em que além da penalidade quadrática consideramos a função de penalidade exponencial e, no próximo capítulo, apresentaremos implementações dos métodos definidos aqui. A ideia central que apresentamos é obter um teorema de equivalência entre os algoritmos de ponto proximal definidos no capítulo anterior e os algoritmos de Lagrangeano aumentado que definimos ao longo deste capítulo. As funções que definem os algoritmos de ponto proximal são dadas por: $\psi^1(x) = \frac{1}{2} \|x\|^2$ e $\psi^2(x) = \sum_{i=1}^n (x_i \log x_i - x_i + 1)$. Note que o domínio \mathcal{D} destas funções são diferentes, pois, $\mathcal{D} = \mathbb{R}^n$ para ψ^1 e $\mathcal{D} = \mathbb{R}_+^n$ para ψ^2 , em que $0 \log(0) = 0$.

3.1 O Método de Lagrangeano aumentado para problemas de equilíbrio

Para definir o método de Lagrangeano aumentado para $EP(f, C)$, vamos assumir que o domínio de f , pode ser estendido ao domínio $\mathbb{R}^n \times \mathbb{R}^n$ satisfazendo as propriedades P1-P4, e P4'. Adicionalmente, o conjunto C em $EP(f, C)$, será considerado como

$$C = \{x \in \mathbb{R}^n : g_i(x) \leq 0, \forall i \in \{1, 2, \dots, m\}\}, \quad (3.1)$$

em que $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ($1 \leq i \leq m$) são funções convexas. Além disso, sobre este conjunto C , vamos considerar a condição de qualificação de Slater, definida em [16].

Definição 3.1. Dizemos que o conjunto C satisfaz a condição de qualificação de Slater, se I é o conjunto de índices i (possivelmente vazio) tal que as funções g_i são afins, então existe $w \in \mathbb{R}^n$ tal que $g_i(w) \leq 0$ para todo $i \in I$ e $g_i(w) < 0$ para $i \notin I$.

Antes de definir os métodos de Lagrangeano aumentado, precisamos considerar a definição de função Lagrangeana associada ao problema de equilíbrio $EP(f, C)$. Este conceito foi introduzido em [16] como a bifunção $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ dada por

$$\mathcal{L}((x, \lambda), (y, \mu)) = f(x, y) + \sum_{i=1}^m \lambda_i g_i(y) - \sum_{i=1}^m \mu_i g_i(x). \quad (3.2)$$

Agora, para $i \in \{1, 2, \dots, m\}$, consideramos a função de penalidade $P_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}_{++} \rightarrow \mathbb{R} \cup \{+\infty\}$ dada por

$$P_i(x, y, \lambda_i, \gamma) = \gamma [\mathcal{P}^\psi(g_i(y), \gamma, \lambda_i) - \mathcal{P}^\psi(g_i(x), \gamma, \lambda_i)], \quad (3.3)$$

em que $\mathcal{P}^\psi : \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ denota a função de penalidade associada à função $\psi = \psi^s$ com $s = \{1, 2\}$, mais precisamente, se tomamos $\psi = \psi^1$ para definir os algoritmos de ponto proximal dados na seção anterior, associamos a penalidade quadrática $\mathcal{P}^\psi(t, \gamma, \lambda) = \frac{1}{2}(\max\{0, \lambda + \frac{t}{\gamma}\}^2 - \lambda^2)$ para definir o métodos de Lagrangeano Aumentado. Analogamente, se consideramos $\psi = \psi^2$ nos algoritmos de ponto proximal, associamos a penalidade exponencial $\mathcal{P}^\psi(s, \gamma, \lambda) = \lambda \exp(\frac{s}{\gamma}) - 1$. Este fato será esclarecido com detalhe na Observação 3.4.

Finalmente, definimos a função Lagrangeano aumentado, $\tilde{\mathcal{L}} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^n \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ para $EP(f, C)$, como sendo

$$\tilde{\mathcal{L}}(x, y, \lambda, z, \gamma) = f(x, y) + \gamma \langle x - z, y - x \rangle + \sum_{i=1}^m P_i(x, y, \lambda_i, \gamma). \quad (3.4)$$

Em seguida, introduzimos versões diferentes dos algoritmos de Lagrangeano aumentado em que a definição de cada algoritmo é dada dependendo da penalidade escolhida. Veremos que em cada método a essência do método de Lagrangeano aumentado para problemas de otimização é preservada, pois, internamente cada algoritmo gera uma sequência de subproblemas de equilíbrio irrestritos que devem ser resolvidos com ferramentas externas ao próprio método. Os dois primeiros algoritmos são versões inexatas que posteriormente veremos que se simplificam em uma versão exata. Salientamos que em [16], os autores introduziram os métodos utilizando a penalidade quadrática, assim como a sua relação com os métodos de ponto proximal. Aqui apresentamos versões destes métodos de uma forma mais geral e compacta, pois podemos considerar penalidades diferentes. Chamaremos ao longo do texto estes três algoritmos de IALEM, IALPM para nos referir aos algoritmos inexatos e EALM para algoritmo exato.

Algoritmo IALEM (Método de Lagrangeano aumentado inexato extragradiante)

Passo 1: Considere a sequência limitada $(\gamma_j)_{j \in \mathbb{N}} \subset \mathbb{R}_{++}$, uma tolerância para o erro $\sigma \in (0, 1)$ e um ponto inicial $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}_+^m$.

Passo 2: Dado o ponto (x^j, λ^j) , encontre o par $(\tilde{x}^j, e^j) \in \mathbb{R}^n \times \mathbb{R}^n$ tal que \tilde{x}^j resolve o problema EP($\tilde{\mathcal{L}}_j^e, \mathbb{R}^n$), com $\tilde{\mathcal{L}}_j^e$ definida como

$$\tilde{\mathcal{L}}_j^e(x, y) = f(x, y) + \gamma_j \langle x - x^j, y - x \rangle + \sum_{i=1}^m P_i(x, y, \lambda_i, \gamma_j) - \langle e^j, y - x \rangle, \quad (3.5)$$

em que P_i é dada por (3.3), e e^j satisfaz

$$\|e^j\| \leq \sigma \gamma_j \|\tilde{x}^j - x^j\|. \quad (3.6)$$

Passo 3: Defina λ^{j+1} como sendo

$$\lambda_i^{j+1} = \gamma_j (\mathcal{P}^\psi)'(g_i(\tilde{x}^j), \gamma_j, \lambda_i^j) \quad (3.7)$$

para cada $i \in \{1, \dots, m\}$.

Passo 4: Se $(\tilde{x}^j, \lambda^{j+1}) = (x^j, \lambda^j)$ então pare. Caso contrário, tome

$$x^{j+1} = \tilde{x}^j - \frac{1}{\gamma_j} e^j. \quad (3.8)$$

Na sequência introduzimos, outro método de Lagrangeano aumentado inexato com projeção que denotaremos ao longo do texto como IALPM, o qual é construído usando as mesmas ideias de [15, 16].

Algoritmo IALPM (Método de Lagrangeano aumentado inexato com Projeção)

Passo 1: Considere a sequência limitada $(\gamma_j)_{j \in \mathbb{N}} \subset \mathbb{R}_{++}$, uma tolerância para o erro $\sigma \in (0, 1)$ e um ponto inicial $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}_{++}^m$.

Passo 2: Dado o ponto (x^j, λ^j) , encontre o par $(\tilde{x}^j, e^j) \in \mathbb{R}^n \times \mathbb{R}^n$ tal que \tilde{x}^j resolve o problema EP($\tilde{\mathcal{L}}_j^e, \mathbb{R}^n$), com $\tilde{\mathcal{L}}_j^e$ definida como

$$\tilde{\mathcal{L}}_j^e(x, y) = f(x, y) + \gamma_j \langle x - x^j, y - x \rangle + \sum_{i=1}^m P_i(x, y, \lambda_i, \gamma_j) - \langle e^j, y - x \rangle, \quad (3.9)$$

em que P_i é dada por (3.3), e e^j satisfaz

$$\|e^j\| \leq \begin{cases} \sigma \gamma_j \|x^j - \tilde{x}^j\|^2 & \text{se } \|x^j - \tilde{x}^j\| < 1 \\ \sigma \gamma_j & \text{se } \|x^j - \tilde{x}^j\| \geq 1. \end{cases} \quad (3.10)$$

Passo 3: Defina λ^{j+1} como sendo

$$\lambda_i^{j+1} = \gamma_j (\mathcal{P}^\psi)'(g_i(\tilde{x}^j), \gamma_j, \lambda_i^j), \quad (3.11)$$

para cada $i \in \{1, \dots, m\}$.

Passo 4: Defina $v^j = \gamma_j[x^j - \tilde{x}^j] + e^j$. Se $(\tilde{x}^j, \lambda^{j+1}) = (x^j, \lambda^j)$ então pare. Caso contrário, considere o hiperplano $H_j = \{x \in \mathbb{R}^n \mid \langle v^j, x - \tilde{x}^j \rangle = 0\}$ e tome

$$x^{j+1} = \operatorname{argmin}_{x \in H_j} \frac{1}{2} \|x - x^j\|^2. \quad (3.12)$$

Note inicialmente que a regularização no passo 2 e a atualização das variáveis duais no passo 3 de IALEM e IALPM, quando consideramos a penalidade exponencial são dadas, respectivamente, por

$$\begin{aligned} \tilde{\mathcal{L}}_j^e(x, y) &= f(x, y) + \gamma_j \langle x - x^j, y - x \rangle + \sum_{i=1}^m P_i(x, y, \lambda_i, \gamma_j) - \langle e^j, y - x \rangle, \\ \lambda_i^{j+1} &= \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right), \quad (1 \leq i \leq m). \end{aligned} \quad (3.13)$$

Além disso, diferenças entre os Algoritmos IALPM e IALEM, estão presentes nos passos 2 e 4. A ideia do passo 2 é praticamente a mesma, diferindo apenas na condição que e^j deve cumprir. Este passo requer uma atenção especial pois estamos assumindo que existe uma ferramenta que resolve (3.5) e (3.6) (ou, (3.9) e (3.10)) simultaneamente. Observe que em (3.5) assim como para o problema de otimização restrita, são gerados, a cada iteração, subproblemas de equilíbrio irrestritos que devem ser resolvidos por uma ferramenta auxiliar dentro do mesmo algoritmo e devido a esta característica é que o método recebe o nome de Lagrangeano aumentado para problemas de equilíbrio. Por outro lado, o Passo 4 de cada um dos algoritmos se divide em duas etapas, a primeira consiste em verificar se o critério de parada é atingido e a segunda é considerada quando a condição de parada não é satisfeita, e neste caso, cada algoritmo utiliza uma forma específica para atualizar a próxima iteração x^{j+1} , isto é, em IALEM utilizamos uma fórmula extragradiente (3.8) e em IALPM, a solução (\tilde{x}^j, e^j) determinada no subproblema gerado no Passo 2, é usada para construir um hiperplano H_j que separa o ponto x^j do conjunto $S(f, C)$ e assim, a próxima iteração é determinada como solução do problema (3.12) o qual é uma projeção com distâncias de Bregman de x^j sobre o hiperplano H_j .

O lema seguinte estabelece que o critério de parada $(\tilde{x}^j, \lambda^{j+1}) = (x^j, \lambda^j)$ definido no Passo 4 de IALPM pode ser trocado por $(v^j, \lambda^j) = (0, \lambda^{j+1})$.

Lema 3.2. *Seja $(x^j)_{j \in \mathbb{N}}$ uma sequência gerada pelo algoritmo IALPM, então $x^j = \tilde{x}^j$ se e somente se $v^j = 0$.*

Demonstração. Suponha inicialmente que $x^j = \tilde{x}^j$. Por definição sabemos que $v^j = \gamma_j[x^j - \tilde{x}^j] + e^j = e^j$. Logo, por (3.10) $\|v^j\| = \|e^j\| \leq \sigma \gamma_j \|\tilde{x}^j - x^j\|^2 = 0$. Reciprocamente, suponha por absurdo que $v^j = 0$ porém $x^j \neq \tilde{x}^j$. Assim, pela definição de v^j no Passo 4 temos que $\|e^j\| = \gamma_j \|x^j - \tilde{x}^j\|$. Agora, pelo Passo 2 de IALPM, podem ocorrer

duas situações para e^j , a primeira destas acontece quando $\|x^j - \tilde{x}^j\| \geq 1$. Neste caso obtemos

$$\|e^j\| \leq \sigma\gamma_j < \gamma_j \leq \gamma_j \|x^j - \tilde{x}^j\| = \|e^j\|,$$

pois $\sigma \in (0, 1)$, o que é absurdo. A outra possibilidade é $0 < \|x^j - \tilde{x}^j\| < 1$, e novamente pelo Passo 2 obtemos

$$\begin{aligned} \|e^j\| &\leq \sigma\gamma_j \|x^j - \tilde{x}^j\|^2 \\ &\leq \sigma\gamma_j \|x^j - \tilde{x}^j\| \\ &< \gamma_j \|x^j - \tilde{x}^j\| \\ &= \|e^j\|, \end{aligned}$$

que também é absurdo. Assim, concluímos que se $v^j = 0$ então $x^j = \tilde{x}^j$.

□

Observação 3.3. O Passo 4 de IALPM consiste em determinar x^{j+1} , o qual é a projeção ortogonal do ponto x^j sobre o hiperplano H_j . A Proposição 1.10 garante que este passo está bem definido pois este resultado garante solução única. Logo, realizando alguns cálculos simples, podemos definir este passo com uma fórmula fechada dada por

$$x^{j+1} = x^j - \frac{1}{\gamma_j} \left(1 - \frac{1}{\|v^j\|^2} \langle e^j, v^j \rangle \right) v^j, \quad (3.14)$$

ou também

$$x^{j+1} = x^j - \eta_j v^j,$$

em que $\eta_j = \frac{1}{\gamma_j} \left(1 - \frac{1}{\|v^j\|^2} \langle e^j, v^j \rangle \right)$. Retornaremos a esta expressão no Capítulo 4, a qual será utilizada nos testes numéricos da implementação deste algoritmo.

Discutidas as diferenças presentes nos algoritmos IALEM e IALPM vejamos que estas diferenças são evidentemente eliminadas quando $e^j = 0$, para todo $j \in \mathbb{N}$. Neste caso, o Passo 2 de cada algoritmo se reduz a encontrar uma solução do subproblema de equilíbrio irrestrito $\text{EP}(\tilde{\mathcal{L}}_j, \mathbb{R}^n)$, em que $\tilde{\mathcal{L}}_j = \tilde{\mathcal{L}}_j^e$ definido como em (3.9). Analogamente, no Passo 4, podemos observar que se $e^j = 0$ para todo $j \in \mathbb{N}$, então de (3.8) e (3.12) obtemos que a próxima iteração é simplificada para a igualdade $x^{j+1} = \tilde{x}^j$, em que \tilde{x}^j é a solução obtida previamente do Passo 2. Com estes argumentos podemos definir o método de Lagrangeano aumentado Exato (EALM).

Algoritmo EALM (Método de Lagrangeano aumentado Exato para $\text{EP}(f, C)$)

Passo 1: Considere a sequência limitada $(\gamma_j)_{j \in \mathbb{N}} \subset \mathbb{R}_{++}$ e um ponto inicial $(x^0, \lambda^0) \in \mathbb{R} \times \mathbb{R}_+^m$.

Passo 2: Dado o ponto (x^j, λ^j) , encontre $\tilde{x}^j \in \mathbb{R}$ tal que \tilde{x}^j resolve o problema $\text{EP}(\tilde{\mathcal{L}}_j, \mathbb{R})$, em que $\tilde{\mathcal{L}}_j$ é definida como

$$\tilde{\mathcal{L}}_j(x, y) = f(x, y) + \gamma_j \langle x - x^j, y - x \rangle + \sum_{i=1}^m P_i(x, y, \lambda_i, \gamma_j), \quad (3.15)$$

com P_i dada como em (3.3).

Passo 3: Defina λ^{j+1} como

$$\lambda_i^{j+1} = \gamma_j (\mathcal{P}^\psi)'(g_i(\tilde{x}^j), \gamma_j, \lambda_i^j) \quad (3.16)$$

para cada $i \in \{1, \dots, m\}$.

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

$$x^{j+1} = \tilde{x}^j.$$

Definidos os algoritmos, vamos esclarecer um pouco as diferentes penalidades que usaremos para definir os diferentes métodos de Lagrangeano aumentado.

Observação 3.4. *Considere $\psi = \psi^1$ e defina os algoritmos de ponto proximal dados na seção anterior. Tomando a penalidade quadrática $\mathcal{P}^\psi(t, \gamma, \lambda) = \frac{1}{2}(\max\{0, \lambda + \frac{t}{\gamma}\}^2 - \lambda^2)$ e realizando alguns cálculos no algoritmo IALEM descrito acima, podemos notar que obtemos exatamente o método de Lagrangeano aumentado definido em [16] e conseqüentemente podemos recuperar a análise de convergência. Analogamente, note que se substituirmos esta penalidade, podemos definir os algoritmos IALPM e EALM definidos implicitamente em [16]. Tendo isto claro, mencionamos que a análise de convergência que realizamos aqui tem como foco principal, mostrar que escolhendo outra função de penalidade \mathcal{P}^ψ , a abordagem realizada em [16] pode ser estendida. Isto permitirá definir novos algoritmos de Lagrangeano aumentado que admitem derivadas de ordem superior.*

Mais precisamente, a função de penalidade que usaremos para definir o método de Lagrangeano aumentado é a penalidade exponencial dada por $\mathcal{P}^\psi(s, \gamma, \lambda) = \lambda \exp(\frac{s}{\gamma}) - 1$.

Portanto, obtemos que a função de penalidade e a atualização das variáveis duais, nos Algoritmos IALEM, IALPM e EALM são dados por:

- 1) $P_i(x, y, \lambda_i, \gamma) = \gamma \lambda_i \left[\exp\left(\frac{g_i(y)}{\gamma}\right) - \exp\left(\frac{g_i(x)}{\gamma}\right) \right],$
- 2) $(\mathcal{P}^\psi)'(g_i(\tilde{x}^j), \gamma_j, \lambda_i^j) = \frac{\lambda_i^j}{\gamma_j} \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right).$

Uma vez esclarecida a utilização destas penalidades para definir as diferentes variantes dos algoritmos de Lagrangeano aumentado, apresentamos a análise de convergência

para cada algoritmo, tomando como principal foco o caso da penalidade exponencial. A abordagem será realizada para os algoritmos inexatos (IALEM ou IALPM) pois, como vimos anteriormente o algoritmo exato EALM é um caso particular dos algoritmos inexatos.

3.2 Análise de convergência

Tendo definido os métodos de Lagrangeano aumentado, as diferenças e semelhanças, agora apresentamos uma análise de convergência das sequências geradas por estes. O objetivo principal desta análise é estabelecer a relação entre os algoritmos de ponto proximal descritos no capítulo anterior e os algoritmos de Lagrangeano aumentado definidos neste capítulo. Para tal, precisamos exibir algumas propriedades entre a função Lagrangeana descrita em (3.2) e a função f do problema $EP(f, C)$. Esclarecemos que embora os algoritmos sejam diferentes em alguns passos, tanto o roteiro como a maioria dos resultados a seguir foram propostos inicialmente para o Algoritmo IALEM em [16], considerando a penalidade quadrática, e extendidos aqui para a penalidade exponencial. O foco principal desta análise será realizada para o caso em que a penalidade considerada é a penalidade exponencial.

Começamos estabelecendo o seguinte resultado, dado em [16], no qual vemos que as propriedades P1-P4 e P4' necessárias para a convergência dos algoritmos de ponto proximal podem ser transferidas de forma natural para a função Lagrangeana (3.2).

Proposição 3.5. *Seja $EP(f, C)$, com C dado como em (3.1). Se f satisfaz P1-P4 e P4' em $\mathbb{R}^n \times \mathbb{R}^n$, então a função lagrangeana \mathcal{L} definida em (3.2), satisfaz P1-P4 e P4' em $(\mathbb{R}^n \times \mathbb{R}_+^m) \times (\mathbb{R}^n \times \mathbb{R}_+^m)$.*

Demonstração. Sejam (x, λ) e (y, μ) pontos em $\mathbb{R}^n \times \mathbb{R}_+^m$. Como f satisfaz P1-P3 e cada função g_i ($1 \leq i \leq m$) é convexa, então pela definição da função Lagrangeana dada em (3.2) podemos ver que as propriedades P1-P3 são transferidas para a função Lagrangeana \mathcal{L} . Como P4 é válida para f , com constante de monotonicidade $\theta \geq 0$, segue de (3.2) que

$$\mathcal{L}((x, \lambda), (y, \mu)) + \mathcal{L}((y, \mu), (x, \lambda)) = f(x, y) + f(y, x) \leq \theta \langle x - y, x - y \rangle = \theta \|x - y\|^2,$$

o que mostra que \mathcal{L} também satisfaz P4 com a mesma constante θ que f . Finalmente, vamos usar que f satisfaz P4' em $(\mathbb{R}^n \times \mathbb{R}_+^m) \times (\mathbb{R}^n \times \mathbb{R}_+^m)$. Para isto, considere $(x^1, \lambda^1), \dots, (x^p, \lambda^p) \in \mathbb{R}^n \times \mathbb{R}_+^m$ e $t_1, \dots, t_p \geq 0$, tais que $\sum_{i=1}^p t_i = 1$. Note primeiramente que

$$\begin{aligned} \mathcal{L} \left((x^\ell, \lambda^\ell), \left(\sum_{k=1}^p t_k x^k, \sum_{k=1}^p t_k \lambda^k \right) \right) &= f \left(x^\ell, \sum_{k=1}^p t_k x^k \right) + \sum_{i=1}^m \lambda_i^\ell g_i \left(\sum_{k=1}^p t_k x^k \right) - \sum_{i=1}^m \sum_{k=1}^p t_k \lambda_i^k g_i(x^\ell) \\ &\leq f \left(x^\ell, \sum_{k=1}^p t_k x^k \right) + \sum_{i=1}^m \sum_{k=1}^p \lambda_i^\ell t_k g_i(x^k) - \sum_{i=1}^m \sum_{k=1}^p t_k \lambda_i^k g_i(x^\ell), \end{aligned}$$

em que utilizamos o fato de que as funções g_i ($1 \leq i \leq m$) são convexas na última desigualdade. Assim, multiplicando a desigualdade acima por t_ℓ e somando os termos $1 \leq \ell \leq p$ obtemos que

$$\begin{aligned} \sum_{\ell=1}^p t_\ell \mathcal{L} \left((x^\ell, \lambda^\ell), \left(\sum_{k=1}^p t_k x^k, \sum_{k=1}^p t_k \lambda^k \right) \right) &\leq \sum_{\ell=1}^p t_\ell f \left(x^\ell, \sum_{k=1}^p t_k x^k \right) \\ &\quad + \sum_{\ell=1}^p \sum_{i=1}^m \sum_{k=1}^p t_\ell t_k \lambda_i^\ell g_i(x^k) - \sum_{\ell=1}^p \sum_{i=1}^m \sum_{k=1}^p t_\ell t_k \lambda_i^k g_i(x^\ell) \\ &= \sum_{\ell=1}^p t_\ell f \left(x^\ell, \sum_{k=1}^p t_k x^k \right) \\ &\leq 0, \end{aligned}$$

em que a última desigualdade é válida porque f satisfaz P4'.

□

Sabendo que as propriedades P1-P4 e P4' de f são transferidas para a função Lagrangeana \mathcal{L} definida em (3.2), então podemos pensar em resolver um problema de equilíbrio dado por $\text{EP}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$, com os métodos de ponto proximal IPPBPM, IPPEM e EPPM definidos no capítulo anterior e, desta forma, mostrar que a sequência gerada pelo método de ponto proximal é equivalente à sequência gerada pelo método de Lagrangeano aumentado.

Considere a função quadrática $\psi : \mathbb{R}^n \times \mathbb{R}_+^m \rightarrow \mathbb{R}$ dada por $\psi(x, \lambda) = \frac{1}{2}(\|x\|^2 + \|\lambda\|^2)$. Note que $\nabla\psi(x, \lambda) = (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_+^m$. Assim, para $\bar{z} = (\bar{x}, \bar{\lambda}) \in \mathbb{R}^n \times \mathbb{R}_+^m$ fixado, podemos definir a função de regularização dada em (2.1) adaptada à função Lagrangeana $\widehat{\mathcal{L}}^e : \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^n \times \mathbb{R}_+^m \rightarrow \mathbb{R}$ dada por

$$\widehat{\mathcal{L}}^e(z, w) = \mathcal{L}(z, w) + \gamma \langle \nabla\psi(z) - \nabla\psi(\bar{z}), w - z \rangle - \langle (e, 0), w - z \rangle. \quad (3.17)$$

Se consideramos $z = (x, \lambda)$ e $w = (y, \mu)$ na igualdade acima, obtemos

$$\begin{aligned} \widehat{\mathcal{L}}^e((x, \lambda), (y, \mu)) &= \mathcal{L}((x, \lambda), (y, \mu)) + \gamma \langle \nabla\psi((x, \lambda)) - \nabla\psi(\bar{x}, \bar{\lambda}), (y, \mu) - (x, \lambda) \rangle - \langle e, y - x \rangle \\ &= \mathcal{L}((x, \lambda), (y, \mu)) + \gamma \langle (x - \bar{x}, \lambda - \bar{\lambda}), (y - x, \mu - \lambda) \rangle - \langle e, y - x \rangle \\ &= \mathcal{L}((x, \lambda), (y, \mu)) + \gamma \langle x - \bar{x}, y - x \rangle + \gamma \langle \lambda - \bar{\lambda}, \mu - \lambda \rangle - \langle e, y - x \rangle. \end{aligned} \quad (3.18)$$

Note que esta função é a regularização quadrática definida em [16]. Neste artigo, os Autores consideram esta regularização para definir os algoritmos de ponto proximal que posteriormente relacionam com os algoritmos de Lagrangeano aumentado com penalidade quadrática. Seguindo estas ideias, para definir a função de regularização da função Lagrangeana, considere a seguinte função

$$\psi(x, \lambda) = \frac{1}{2} \|x\|^2 + \sum_{i=1}^m (\lambda_i \log \lambda_i - \lambda_i + 1). \quad (3.19)$$

Observe que a diferença entre esta função e a ψ considerada em [16], é que as variáveis primais são dadas por uma função quadrática e as variáveis duais são dadas por uma função Kullback Leibler. Desde que $\nabla\psi(x, \lambda) = (x, \log \lambda)$, a função de regularização $\widehat{\mathcal{L}}^e : \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^n \times \mathbb{R}_+^m \rightarrow \mathbb{R}$ é definida como

$$\widehat{\mathcal{L}}^e((x, \lambda), (y, \mu)) = \mathcal{L}((x, \lambda), (y, \mu)) + \gamma \langle x - \bar{x}, y - x \rangle + \gamma \langle \log \lambda - \log \bar{\lambda}, \mu - \lambda \rangle - \langle e, y - x \rangle. \quad (3.20)$$

Note que, com esta regularização, as propriedades mostradas nas Proposições 2.6 e 2.7 são válidas para $\widehat{\mathcal{L}}^e$. Como consequência, podemos reescrever os algoritmos IPPBMP e IPPEM para resolver o problema $\text{EP}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$, isto é, se assumirmos que as condições iniciais definidas no passo 1 de IPPBMP são válidas, então podemos reescrever os passos 2 e 3 como:

Passo 2: Encontre o par $\left((\hat{x}_j, \hat{\lambda}_j), (e^j, 0) \right) \in (\mathbb{R}^n \times \mathbb{R}_+^m) \times (\mathbb{R}^n \times \mathbb{R}^m)$ tal que $(\hat{x}_j, \hat{\lambda}_j)$ resolve $\text{EP}(\widehat{\mathcal{L}}_j^e, \mathbb{R}^n \times \mathbb{R}_+^m)$, em que $\widehat{\mathcal{L}}_j^e$ é a regularização para \mathcal{L} na iteração j dada por:

$$\begin{aligned} \widehat{\mathcal{L}}_j^e((x, \lambda), (y, \mu)) &= \mathcal{L}((x, \lambda), (y, \mu)) + \gamma_j \langle x - x^j, y - x \rangle \\ &\quad + \gamma_j \langle \log \lambda - \log \lambda^j, \mu - \lambda \rangle - \langle e^j, y - x \rangle \end{aligned} \quad (3.21)$$

e o vetor erro $(e^j, 0) \in \mathbb{R}^n \times \mathbb{R}^m$, deve satisfazer

$$\|(e^j, 0)\| = \|e^j\| \leq \begin{cases} \sigma \gamma_j D_\psi((\hat{x}^j, \hat{\lambda}^j), (x^j, \lambda^j)) & \text{se } \|(x^j, \lambda^j) - (\hat{x}^j, \hat{\lambda}^j)\| < 1 \\ \sigma \gamma_j \nu_\psi((x^j, \lambda^j), 1) & \text{se } \|(x^j, \lambda^j) - (\hat{x}^j, \hat{\lambda}^j)\| \geq 1. \end{cases} \quad (3.22)$$

Passo 3: Se $(x^j, \lambda^j) = (\hat{x}^j, \hat{\lambda}^j)$ então pare. Caso contrário, tome $v^j = \gamma^j [x^j - \hat{x}^j] + e^j$, defina o hiperplano $H_j = \{x \in \mathbb{R}^n \mid \langle v^j, x - \hat{x}^j \rangle = 0\}$ e atualize

$$x^{j+1} = \underset{x \in H_j}{\operatorname{argmin}} \frac{1}{2} \|x - x^j\|^2. \quad (3.23)$$

$$\lambda^{j+1} = \hat{\lambda}^j. \quad (3.24)$$

Analogamente, podemos considerar o Algoritmo IPPEM para resolver o problema $\text{EP}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$. Assim, os passos 2 e 3 são dados da seguinte forma:

Passo 2: Encontre o par $\left((\hat{x}_j, \hat{\lambda}_j), (e^j, 0) \right) \in (\mathbb{R}^n \times \mathbb{R}_+^m) \times (\mathbb{R}^n \times \mathbb{R}^m)$ tal que $(\hat{x}_j, \hat{\lambda}_j)$ resolve

EP($\widehat{\mathcal{L}}_j^e, \mathbb{R}^n \times \mathbb{R}_+^m$), em que $\widehat{\mathcal{L}}_j^e$ é a regularização para \mathcal{L} na iteração j dada por:

$$\widehat{\mathcal{L}}_j^e((x, \lambda), (y, \mu)) = \mathcal{L}((x, \lambda)(y, \mu)) + \gamma_j \langle x - x^j, y - x \rangle + \gamma_j \langle \nabla \psi(\lambda) - \nabla \psi(\lambda^j), \mu - \lambda \rangle - \langle e^j, y - x \rangle$$

e o vetor erro $(e^j, 0) \in \mathbb{R}^n \times \mathbb{R}^m$ deve satisfazer

$$\|(e^j, 0)\| = \|e\| \leq \sigma \gamma_j D_\psi((\hat{x}^j, \hat{\lambda}^j), (x^j, \lambda^j))$$

Passo 3: Se $(x^j, \lambda^j) = (\hat{x}^j, \hat{\lambda}^j)$ então pare. Caso contrário, tome

$$x^{j+1} = \hat{x}^j - \gamma_j^{-1} e^j. \quad (3.25)$$

$$\lambda^{j+1} = \hat{\lambda}^j. \quad (3.26)$$

Observação 3.6. Note que a distância de Bregman induzida por (3.19) é dada por $D_\psi((x, \lambda), (y, \mu)) = D_{\psi_1}(x, y) + D_{\psi_2}(\lambda, \mu)$, em que D_{ψ_1} e D_{ψ_2} são distâncias de Bregman induzidas por $\psi_1(x) = \frac{1}{2} \|x\|^2$ e $\psi_2(\lambda) = \sum_{i=1}^m (\lambda \log \lambda - \lambda + 1)$, respectivamente.

Definidos os algoritmos de ponto proximal, o objetivo a seguir será mostrar um teorema de equivalência, entre estes algoritmos e os algoritmos de Lagrangeano aumentado. Antes disso, vamos precisar de alguns resultados auxiliares. Começamos apresentando a definição par ótimo dada em [16], para o problema de equilíbrio EP(f, C).

Definição 3.7. Dizemos que o par $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ é um par ótimo para EP(f, C) quando

$$0 \in \partial F_{x^*}(x^*) + \sum_{i=1}^m \lambda_i^* \partial g_i(x^*), \quad (3.27)$$

$$\lambda_i^* \geq 0, \quad (1 \leq i \leq m), \quad (3.28)$$

$$g_i(x^*) \leq 0 \quad (1 \leq i \leq m), \quad (3.29)$$

$$\lambda_i^* g_i(x^*) = 0 \quad (1 \leq i \leq m), \quad (3.30)$$

em que os conjuntos $\partial F_{x^*}(x^*)$ e $\partial g_i(x^*)$ denotam o subdiferencial no ponto x^* das funções convexas F_{x^*} e $g_i(x^*)$, respectivamente.

Esta definição foi introduzida em [16] e note que coincide com as condições de Karush Kuhn Tucker (KKT) para o problema de minimizar a função F_{x^*} em um conjunto

convexo $C \subset \mathbb{R}^n$, com C dado como em (3.1). Por outro lado, note que não necessariamente um par que cumpra as condições de KKT para o problema de minimizar a função F_x em um conjunto convexo C é um par ótimo.

Na sequência apresentamos algumas proposições relacionadas ao problema de equilíbrio $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$. Apesar de as demonstrações possuírem uma certa semelhança às dadas em [16], iremos reescrever as provas para completude do texto e para uma melhor organização na exposição do conteúdos que virão na sequência.

Proposição 3.8. *Considere $EP(f, C)$ e suponha que f satisfaz P1-P3 em $\mathbb{R}^n \times \mathbb{R}^m$. Então as seguintes afirmações são equivalentes:*

$$a) (x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$$

$$b) (x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}_+^m \text{ é um par ótimo para } EP(f, C).$$

Demonstração. Suponha que $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$. Como f cumpre as propriedades P1-P3, então pela Proposição 3.5, \mathcal{L} também cumpre as mesmas propriedades. Assim, pela Proposição 2.16 temos que (x^*, λ^*) é solução do problema

$$\min \mathcal{F}_{(x^*, \lambda^*)}(x, \lambda) \tag{3.31}$$

$$\text{s.a. } (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_+^m, \tag{3.32}$$

em que $\mathcal{F}_{(x^*, \lambda^*)}(x, \lambda) = \mathcal{L}((x^*, \lambda^*), (x, \lambda))$. Como o problema acima é convexo, e as restrições são afins, então a condição de qualificação de Slater, dada na Definição 3.1, é satisfeita para este problema, ou seja, existe $\eta^* \in \mathbb{R}^m$ tal que

$$0 \in \partial F_{x^*}(x^*) + \sum_{i=1}^m \lambda_i^* \partial g_i(x^*), \tag{3.33}$$

$$g_i(x^*) + \eta_i^* = 0, \quad (1 \leq i \leq m), \tag{3.34}$$

$$\eta_i^* \geq 0, \quad (1 \leq i \leq m) \tag{3.35}$$

$$\eta_i^* \lambda_i^* = 0, \quad (1 \leq i \leq m) \tag{3.36}$$

$$\lambda_i^* \geq 0, \quad (1 \leq i \leq m). \tag{3.37}$$

Agora, observe que de (3.34), para cada índice $i \in \{1, \dots, m\}$, $\eta_i^* = -g_i(x^*)$ e substituindo em (3.35) e (3.36) obtemos que $g_i(x^*) \leq 0$ e $\lambda_i^* g_i(x^*) = 0$, respectivamente. Assim, obtemos que as equações (3.33)-(3.37) são equivalentes à

$$0 \in \partial F_{x^*}(x^*) + \sum_{i=1}^m \lambda_i^* \partial g_i(x^*), \quad (3.38)$$

$$g_i(x^*) \leq 0, \quad (1 \leq i \leq m), \quad (3.39)$$

$$\lambda_i^* g_i(x^*) = 0, \quad (1 \leq i \leq m), \quad (3.40)$$

$$\lambda_i^* \geq 0, \quad (1 \leq i \leq m), \quad (3.41)$$

as quais coincidem com as condições exigidas na definição de par ótimo.

Reciprocamente, supondo que (x^*, λ^*) é um par ótimo para $EP(f, C)$, então tomando para cada $i \in \{1, \dots, m\}$, $\eta_i^* = -g_i(x^*)$ obtemos que as equações (3.33)-(3.37) são satisfeitas. Agora, como o problema (3.31)-(3.32) é convexo, então as condições KKT (3.33)-(3.37), são condições suficientes de otimalidade, e assim, (x^*, λ^*) resolve este problema. Em consequência, pela Proposição 2.16 temos que, $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$.

□

Proposição 3.9. *Considere $EP(f, C)$ e suponha que f satisfaz P1-P3 em $\mathbb{R}^n \times \mathbb{R}^n$. Se x^* é um minimizador de F_{x^*} sobre C , com F_{x^*} definida como em (2.31), e a condição de qualificação de Slater dada como na Definição 3.1 é válida para as restrições g_i que definem o conjunto C , então existe $\lambda^* \in \mathbb{R}^m$ tal que (x^*, λ^*) é um par ótimo para $EP(f, C)$. Reciprocamente, se (x^*, λ^*) é um par ótimo para $EP(f, C)$, então x^* é um minimizador de F_{x^*} sobre C , com F_{x^*} definida como em (2.31).*

Demonstração. Suponha que x^* é solução do problema

$$\begin{aligned} \min \quad & F_{x^*}(x) \\ \text{s.a.} \quad & g(x) \leq 0. \end{aligned} \quad (3.42)$$

Como C cumpre a condição de qualificação de Slater, então temos que existem $\lambda^* \in \mathbb{R}_+^m$ tal que as condições (3.27)-(3.30) são satisfeitas, ou seja, (x^*, λ^*) é um par ótimo para $EP(f, C)$. Reciprocamente, se (x^*, λ^*) é um par ótimo para $EP(f, C)$ então, por definição, as condições (3.27)-(3.30) são válidas, que são justamente as condições KKT para o problema (3.42). Como o problema é convexo estas condições são suficientes para garantir que x^* seja solução de (3.42), como queríamos.

□

Corolário 3.10. *Considere $EP(f, C)$, e suponha que f satisfaz P1-P3 em $\mathbb{R}^n \times \mathbb{R}^n$. Se $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ então $x^* \in S(f, C)$. Reciprocamente, se $x^* \in S(f, C)$ e a condição de qualificação de Slater é válida, então existe $\lambda^* \in \mathbb{R}_+^m$ tal que $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$.*

Demonstração. Suponha que $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$, então pela Proposição 3.8 temos que $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}_+^m$ é um par ótimo para $EP(f, C)$ e pela Proposição 3.9 obtemos que x^* é minimizador de F_{x^*} sobre C , que pela Proposição 2.16, equivale a dizer que $x^* \in S(f, C)$. Reciprocamente, se $x^* \in S(f, C)$ então pela Proposição 2.16 temos que $x^* \in \mathbb{R}^n$ é minimizador de F_{x^*} sobre C e como por hipótese a condição de qualificação de Slater vale, então a proposição anterior garante a existência de $\lambda^* \in \mathbb{R}_+^m$ tal que $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}_+^m$ é um par ótimo para o problema $EP(f, C)$, que pela Proposição 3.8, equivale a dizer que $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$. \square

O Corolário 3.10 nos diz que encontrar solução de $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ é suficiente para encontrar solução de $EP(f, C)$. No próximo teorema vamos mostrar que a sequência gerada pelo Algoritmo IPPBPM para resolver $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$, coincide com a sequência gerada por IALPM para $EP(f, C)$.

Teorema 3.11. *Suponha que $EP(f, C)$ satisfaz P1-P4 em $\mathbb{R}^n \times \mathbb{R}^n$. Fixe a sequência $(\gamma_j)_{j \in \mathbb{N}} \subset \mathbb{R}_{++}$ e o erro relativo com tolerância $\sigma \in (0, 1)$. Sejam $((x^j, \lambda^j))_{j \in \mathbb{N}}$ uma sequência gerada pelo algoritmo IALPM (ou IALEM) aplicado a $EP(f, C)$, com erro $e^j \in \mathbb{R}^n$ dado por (3.10) (ou (3.6)) e $(\bar{x}^j, \bar{\lambda}^j)_{j \in \mathbb{N}}$ a sequência gerada pelo algoritmo IPPBPM (ou IPPEM) aplicado ao problema $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ com vetor erro $(e^j, 0) \in \mathbb{R}^n \times \mathbb{R}^m$, e utilizando os mesmos γ_j 's e σ . Se $(x^0, \lambda^0) = (\bar{x}^0, \bar{\lambda}^0)$ então $(x^j, \lambda^j) = (\bar{x}^j, \bar{\lambda}^j)$ para todo $j \in \mathbb{N}$.*

Demonstração. Procederemos por indução sobre j . Para $j = 0$ a igualdade é verdadeira por hipótese. Suponha que o teorema é válido para algum $j \in \mathbb{N}$, isto é, $(x^j, \lambda^j) = (\bar{x}^j, \bar{\lambda}^j)$. Note que pela Proposição 2.7 o subproblema $EP(\widehat{\mathcal{L}}_j^e, \mathbb{R}^n \times \mathbb{R}_+^m)$, com $\widehat{\mathcal{L}}_j^e$ dado como em (3.21), gerado em cada iteração, do Passo 2, do Algoritmo IPPBPM tem solução única, digamos $(\hat{x}^j, \hat{\lambda}^j)$. Agora, pela Proposição 2.16, temos que $(\hat{x}^j, \hat{\lambda}^j)$ resolve o problema convexo definido por

$$\min \widehat{\mathcal{F}}_{(\hat{x}^j, \hat{\lambda}^j)}(x, \lambda) \quad (3.43)$$

$$s.a \ (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_+^m, \quad (3.44)$$

com $\widehat{\mathcal{F}}_{(\hat{x}^j, \hat{\lambda}^j)}(x, \lambda) = \widehat{\mathcal{L}}_j^e((\hat{x}^j, \hat{\lambda}^j), (x, \lambda))$. Como o problema é convexo então existe $\eta^j \in \mathbb{R}^m$, tal que:

$$\gamma_j[\bar{x}^j - \hat{x}^j] + e^j \in \partial F_{\hat{x}^j}(\hat{x}^j) + \sum_{i=1}^m \hat{\lambda}_i^j \partial g_i(\hat{x}^j), \quad (3.45)$$

$$-g_i(\hat{x}^j) + \gamma_j[\log(\hat{\lambda}_i^j) - \log(\bar{\lambda}_i^j)] = \eta_i^j, \quad (1 \leq i \leq m), \quad (3.46)$$

$$\hat{\lambda}^j \geq 0$$

$$\eta^j \geq 0, \quad (3.47)$$

$$\hat{\lambda}_i^j \eta_i^j = 0, \quad (1 \leq i \leq m). \quad (3.48)$$

Logo, combinando a igualdade $\eta_i^j = -g_i(\hat{x}^j) + \gamma_j[\log(\hat{\lambda}_i^j) - \log(\bar{\lambda}_i^j)]$, ($1 \leq i \leq m$), obtida de (3.46), com as equações (3.47) e (3.48) podemos eliminar a variável η^j e assim, o sistema acima é equivalente à

$$\begin{aligned} \gamma_j[\bar{x}^j - \hat{x}^j] + e^j &\in \partial F_{\hat{x}^j}(\hat{x}^j) + \sum_{i=1}^m \hat{\lambda}_i^j \partial g_i(\hat{x}^j), \\ -g_i(\hat{x}^j) + \gamma_j[\log(\hat{\lambda}_i^j) - \log(\bar{\lambda}_i^j)] &\leq 0, \quad (1 \leq i \leq m), \end{aligned} \quad (3.49)$$

$$\hat{\lambda}_i^j \geq 0, \quad (1 \leq i \leq m), \quad (3.50)$$

$$\hat{\lambda}_i^j \left(-g_i(\hat{x}^j) + \gamma_j[\log(\hat{\lambda}_i^j) - \log(\bar{\lambda}_i^j)] \right) = 0, \quad (1 \leq i \leq m). \quad (3.51)$$

Novamente combinando (3.49), (3.50) e (3.51) obtemos

$$\gamma_j[\bar{x}^j - \hat{x}^j] + e^j \in \partial F_{\hat{x}^j}(\hat{x}^j) + \sum_{i=1}^m \hat{\lambda}_i^j \partial g_i(\hat{x}^j), \quad (3.52)$$

$$\hat{\lambda}_i^j = 0, \text{ ou } \hat{\lambda}_i^j = \bar{\lambda}_i^j \exp\left(\frac{g_i(\hat{x}^j)}{\gamma_j}\right), \quad (1 \leq i \leq m).$$

Como pela Proposição 2.17, os pontos gerados por esta sequência sempre estão no interior de $\mathbb{R}^n \times \mathbb{R}_+^m$, então

$$\hat{\lambda}_i^j = \bar{\lambda}_i^j \exp\left(\frac{g_i(\hat{x}^j)}{\gamma_j}\right), \quad (1 \leq i \leq m). \quad (3.53)$$

Assim, substituindo (3.53) em (3.52) concluímos que

$$\gamma_j[\bar{x}^j - \hat{x}^j] + e^j \in \partial F_{\hat{x}^j}(\hat{x}^j) + \sum_{i=1}^m \bar{\lambda}_i^j \exp\left(\frac{g_i(\hat{x}^j)}{\gamma_j}\right) \partial g_i(\hat{x}^j). \quad (3.54)$$

Por outro lado, observe que o Passo 2 do Algoritmo IALPM exige uma solução \tilde{x}^j do subproblema $\text{EP}(\tilde{\mathcal{L}}_j^e, \mathbb{R}^n)$, que pela Proposição 2.16 sabemos que equivale a que o sistema não linear e não diferenciável

$$\gamma_j[\bar{x}^j - \tilde{x}^j] + e^j \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right) \partial g_i(\tilde{x}^j), \quad (3.55)$$

seja satisfeito. Como pela hipótese indutiva $x^j = \bar{x}^j$ e $\lambda^j = \bar{\lambda}^j$, então as equações (3.55) são equivalentes à

$$\gamma_j[\bar{x}^j - \tilde{x}^j] + e^j \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^m \bar{\lambda}_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right) \partial g_i(\tilde{x}^j). \quad (3.56)$$

Agora, note que se substituirmos \tilde{x}^j por \hat{x}^j em (3.56) obtemos exatamente o sistema (3.54), e portanto, $\hat{x}^j \in S(\tilde{\mathcal{L}}_j^e, \mathbb{R}^n)$. Como pela Proposição 2.7 as soluções do sistema (3.54) são

determinadas de forma única então concluímos que

$$\hat{x}^j = \tilde{x}^j. \quad (3.57)$$

Por outro lado, se combinamos a equação (3.12) no Passo 4 de IALPM (ou (3.8) no algoritmo IALEM) para resolver $EP(f, C)$, a equação (3.23) no Passo 3 do algoritmo IPPBPM (ou (3.25) em IPPEM) para $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$, a hipótese indutiva e a igualdade (3.57) concluímos que $x^{j+1} = \bar{x}^{j+1}$.

Para a atualização das variáveis duais, basta olhar para as equações (3.26) e (3.53), e assim obter que:

$$\bar{\lambda}_i^{j+1} = \hat{\lambda}_i^j = \bar{\lambda}_i^j \exp\left(\frac{g_i(\hat{x}^j)}{\gamma_j}\right), \quad (1 \leq i \leq m).$$

Logo, comparando a última igualdade com (3.11), levando em conta (3.57) e $\bar{\lambda}^j = \lambda^j$ concluímos que $\bar{\lambda}^{j+1} = \lambda^{j+1}$. □

Vejamos agora que o critério de parada estabelecido nos algoritmos IALEM e IALPM está bem definido.

Proposição 3.12. *Suponha que o algoritmo IALPM (ou IALEM) para em uma iteração j . Então o vetor \tilde{x}^j gerado por este algoritmo é uma solução para $EP(f, C)$.*

Demonstração. Se o algoritmo IALPM (ou IALEM) para em uma iteração j , então pelo Passo 4 temos que $(\tilde{x}^j, \lambda^{j+1}) = (x^j, \lambda^j)$, em que \tilde{x}^j é solução do subproblema j , gerado no Passo 2 do algoritmo IALPM (ou IALEM). Sabemos do Teorema 3.11, que $\tilde{x}^j \in S(\mathcal{L}_j^e, \mathbb{R}^n)$ se e somente se

$$\gamma_j[x^j - \tilde{x}^j] + e^j \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right) \partial g_i(\tilde{x}^j).$$

Note que se $x^j = \tilde{x}^j$ então pelo Lema 3.2 vale que $\gamma_j[x^j - \tilde{x}^j] + e^j = v^j = 0$ e assim, obtemos que

$$0 \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right) \partial g_i(\tilde{x}^j).$$

Por outro lado, usando que $\lambda^j = \lambda^{j+1}$ em (3.11), segue que

$$\lambda_i^j = \lambda_i^{j+1} = \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right) \quad (1 \leq i \leq m),$$

ou seja,

$$\lambda_i^j \left[1 - \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right)\right] = 0 \quad (1 \leq i \leq m).$$

Portanto, obtemos que

$$0 \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^m \lambda_i^j \partial g_i(\tilde{x}^j), \quad (3.58)$$

$$\lambda_i^j \geq 0, \quad \lambda_i^j g_i(\tilde{x}^j) = 0, \quad g_i(\tilde{x}^j) \leq 0, \quad (1 \leq i \leq m). \quad (3.59)$$

Deste modo, se combinamos (3.58) e (3.59) podemos garantir que o par (\tilde{x}^j, λ^j) cumpre as condições de KKT para o problema convexo de minimizar a função $F_{\tilde{x}^j}$ no conjunto C e portanto, da Proposição 2.16, concluímos que $\tilde{x}^j \in S(f, C)$.

□

Para finalizar este capítulo, vamos estudar a convergência de IALEM e IALPM. A ideia é utilizar o seguinte corolário, o qual é uma adaptação do Teorema (2.15) para $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$.

Corolário 3.13. *Sejam $EP(f, C)$ com C dado como em (3.1) e f satisfazendo P1-P4, P4' em $\mathbb{R}^n \times \mathbb{R}_+^m$. Tome $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma}]$ para algum $\bar{\gamma} > \theta$, com θ a constante dada na hipótese P4. Seja $(x^j, \lambda^j)_{j \in \mathbb{N}}$ a sequência gerada pelo algoritmo IPPEM (ou IPPBPM) aplicado a $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$. Se o problema $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ tem solução, então a sequência $(x^j, \lambda^j)_{j \in \mathbb{N}}$ resolve assintoticamente $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ e todo ponto de acumulação $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$.*

Demonstração. Como por hipótese f satisfaz P1-P4 e P4' então, pela Proposição 3.5, \mathcal{L} também cumpre estas propriedades em $(\mathbb{R}^n \times \mathbb{R}_+^m) \times (\mathbb{R}^n \times \mathbb{R}_+^m)$. Com isto, as hipóteses do Teorema 2.15 se cumprem para $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$, e desta forma obtemos o resultado desejado.

□

Note que no caso em que a função usada nos algoritmos de ponto proximal, para as variáveis primal e dual, seja a função quadrática, basta usar o Corolário 3.13 para obter convergência nos algoritmos IALEM e IALPM através do teorema de equivalência como mostrado em [16]. Por outro lado, o caso em que usamos $\psi(x, \lambda) = \frac{1}{2} \|x\|^2 + \sum_{i=1}^m \lambda_i \log(\lambda_i) - \lambda_i + 1$, sabemos que $(\lambda^j)_{j \in \mathbb{N}}$ é sempre positiva, pois $\lambda^{j+1} = \lambda^j \exp(\frac{g(\tilde{x}^j)}{\gamma})$ e $\lambda^0 > 0$. Analisamos o que acontece quando $\lambda_i^j \rightarrow 0$, para alguma componente i . Para este i , suponha, sem perda de generalidade que $\lambda_i^{j+1} < \lambda_i^j$ para todo $j \in \mathbb{N}$. Como $\lambda_i^{j+1} = \lambda_i^j \exp(\frac{g_i(\tilde{x}^j)}{\gamma})$ segue que $\exp(\frac{g_i(\tilde{x}^j)}{\gamma}) = \frac{\lambda_i^{j+1}}{\lambda_i^j} < 1$ o qual implica que $g(\tilde{x}^j) < 0$, ou seja, a sequência $(\tilde{x}^j)_{j \in \mathbb{N}}$ se torna viável quando j cresce. Em consequência qualquer ponto de acumulação da sequência gerada pelo algoritmo IALEM ou IALPM cumpre as condições (3.59).

Do ponto de vista computacional, ou seja quando os algoritmos IALEM e IALPM são implementados, usamos como critério de parada $\|x^{j+1} - x^j\| < \varepsilon$ e $\|\lambda^{j+1} - \lambda^j\| < \varepsilon$, para algum $\varepsilon > 0$. Isto significa que mesmo aquelas componentes do multiplicador de Lagrange que serão zero na solução estarão sendo satisfeitas para uma precisão ε dada. Em outras palavras, nunca encontramos a solução exata e sim aproximada. Logo, o problema

$EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ para este caso será resolvido de forma aproximada pelo $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_\varepsilon^m)$ para ε arbitrariamente pequeno. Assim, note que, do Corolário 3.13 e do Teorema 3.11, obtemos convergência dos algoritmos IALPM e IALEM.

A convergência do Algoritmo IALPM é estabelecida no seguinte teorema.

Teorema 3.14. *Considere $EP(f, C)$ e suponha que valem*

- i) f satisfaz $P1$ - $P4$, $P4'$ em $\mathbb{R}^n \times \mathbb{R}^n$*
- ii) C é dado como em (3.1)*
- iii) A condição de qualificação de Slater, dada como na Definição 3.1, é válida para o conjunto C*
- iv) $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma}]$ para algum $\bar{\gamma} > \theta$, com θ a constante de submonotonicidade definida em $P4$.*

Seja $((x^j, \lambda^j))_{j \in \mathbb{N}}$ a sequência gerada pelo algoritmo IALPM para resolver $EP(f, C)$. Se $EP(f, C)$ tem solução, então a sequência $((x^j, \lambda^j))_{j \in \mathbb{N}}$ converge para algum par ótimo (x^*, λ^*) do $EP(f, C)$ e conseqüentemente $x^* \in S(f, C)$.

Demonstração. Pelo Teorema 3.11 a sequência $((x^j, \lambda^j))_{j \in \mathbb{N}}$ coincide com a sequência gerada pelo algoritmo IPPBPM aplicado ao problema $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$. Por hipótese $EP(f, C)$ tem solução e a condição de qualificação de Slater vale, então do Corolário 3.10 obtemos que $EP(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ também tem solução. No Corolário 3.13 mostramos que a sequência $((x^j, \lambda^j))_{j \in \mathbb{N}}$ é convergente, digamos para $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ que pela Proposição 3.8 é um par ótimo para $EP(f, C)$. Assim, se utilizamos novamente o Corolário 3.10 concluímos que $x^* \in S(f, C)$. □

O Teorema 3.14 também pode ser reescrito para o algoritmo IALEM, a demonstração é essencialmente a mesma e por isto a omitimos. Além disso, se considerarmos $e^j = 0$ para todo $j \in \mathbb{N}$ no teorema anterior, então este algoritmo pode ser usado para concluir que a sequência gerada pelo algoritmo exato EALM também converge para uma solução do problema de equilíbrio $EP(f, C)$.

O corolário seguinte fornece uma garantia de que sempre é possível resolver o Passo 2 dos algoritmos que estamos considerando neste trabalho.

Corolário 3.15. *Considere o problema $EP(f, C)$, com f continuamente diferenciável satisfazendo $P1$ - $P4$, $P4'$ em $\mathbb{R}^n \times \mathbb{R}^n$ e as funções g_i , $1 \leq i \leq m$ que definem o conjunto convexo C diferenciáveis. Seja $((x^j, \lambda_j))_{j \in \mathbb{N}}$ a sequência gerada pelo algoritmo IALPM (ou IALEM). Suponha que x^j não é solução de $EP(f, C)$ e que \tilde{x}^j é solução única do problema $EP(\tilde{\mathcal{L}}_j^e, \mathbb{R}^n)$, definido como em (3.9) (ou (3.5)), com $e^j = 0$. Então existe $\delta_j > 0$ tal que qualquer $x \in B(\tilde{x}^j, \delta_j)$ resolve as equações (3.9) e (3.10) (ou (3.5) e (3.6)).*

Demonstração. Ver Corolário 3.13 de [16].

□

Este corolário estabelece que sob hipótese de diferenciabilidade sempre é possível resolver o Passo 2 do algoritmo IALPM (ou IALEM). Usaremos este fato no próximo capítulo para definir algoritmos que resolvam o Passo 2 dos Algoritmos de Lagrangeano aumentado.

Capítulo 4

Implementação dos algoritmos propostos

Como mencionamos no início do capítulo anterior, o objetivo principal deste capítulo é desenvolver algoritmos que permitem resolver o Passo 2 do Algoritmo IALPM e assim estabelecer sua implementação. Em [25], os autores apresentam dois algoritmos para resolver o Passo 2 do Algoritmo IALEM. Neste capítulo, vamos propor três esquemas iterativos para resolver os subproblemas de equilíbrio irrestritos gerados no Passo 2 do Algoritmo IALPM, utilizando as penalidades quadrática e exponencial. No primeiro algoritmo adaptamos as mesmas ideias de [25], isto é, aplicamos o método de Newton para resolver um sistema não linear e para os dois últimos consideramos uma adaptação do método de projeção de subgradiente proposto em [18] para resolver os subproblemas de forma exata, isto é, considerando $e^j = 0$, para todo $j \in \mathbb{N}$.

4.1 Método de Newton para resolver o Passo 2 de IALPM

Vimos na demonstração do Teorema 3.11, ou na equação (3.55) do capítulo anterior, que $\tilde{x}^j \in S(\tilde{\mathcal{L}}_j^e, \mathbb{R}^n)$ se e somente se

$$e^j \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right) \partial g_i(\tilde{x}^j) + \gamma_j[\tilde{x}^j - x^j], \quad (4.1)$$

para o caso em que a função de penalidade seja a função exponencial. Analogamente, para o caso da penalidade quadrática, obtemos

$$e^j \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^m \max\left\{0, \lambda_i^j + \frac{g_i(\tilde{x}^j)}{\gamma_j}\right\} \partial g_i(\tilde{x}^j) + \gamma_j[\tilde{x}^j - x^j], \quad (4.2)$$

em que assumimos as notações, $\partial F_{\tilde{x}^j}(\tilde{x}^j)$ para nos referir ao subdiferencial de $F_{\tilde{x}^j}$ avaliado no ponto \tilde{x}^j e $F_{\tilde{x}^j}$ definida como em (2.31), $\partial g_i(\tilde{x}^j)$, para $i = 1, 2, \dots, m$, denota o subdiferencial de g_i avaliado no ponto \tilde{x}^j e $S(\tilde{\mathcal{L}}_j^e, \mathbb{R}^n)$ é o conjunto solução de EP($\tilde{\mathcal{L}}_j^e, \mathbb{R}^n$). O Corolário 3.15 estabelece que, sob certas condições de diferenciabilidade, podemos resolver o Passo 2 do algoritmo IALPM em um número finito de passos, isto é, podemos garantir a existência do par (\tilde{x}^j, e^j) tal que \tilde{x}^j é solução de EP($\tilde{\mathcal{L}}_j^e, \mathbb{R}^n$) e e^j satisfaz (3.10). Tendo em vista este resultado, suponha que existe uma sub-rotina que resolva o j -ésimo subproblema (3.9) gerado no Passo 2 do algoritmo IALPM com $e^j = 0$. Através desta configuração, são geradas soluções aproximadas $\tilde{x}^{j,\ell}$ cujos vetores erro são $e^{j,\ell}$ para $\ell = 1, \dots$. Uma vez que $\tilde{x}^{j,\ell}$ é calculado na iteração ℓ da sub-rotina, checamos se o critério dado em (3.9) é satisfeito, para assim obter o par $(\tilde{x}^{j,\ell}, e^{j,\ell})$ como solução do Passo 2 do algoritmo IALPM. Caso contrário, se o vetor erro $e^{j,\ell}$ não satisfaz (3.9) e (3.10) então chamamos a sub-rotina para gerar mais uma iteração e fazemos $\ell = \ell + 1$. Em particular, se consideramos o caso em que f e g_i são diferenciáveis, então temos que $\tilde{x}^{j,\ell}$ satisfaz

$$e^j = \gamma_j(\tilde{x}^{j,\ell} - x^j) + \nabla F_{\tilde{x}^{j,\ell}}(\tilde{x}^{j,\ell}) + \sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^{j,\ell})}{\gamma_j}\right) \nabla g_i(\tilde{x}^{j,\ell}), \quad (4.3)$$

e no caso em que usamos a penalidade quadrática:

$$e^j = \gamma_j(\tilde{x}^{j,\ell} - x^j) + \nabla F_{\tilde{x}^{j,\ell}}(\tilde{x}^{j,\ell}) + \sum_{i=1}^m \max\left\{0, \lambda_i^j + \frac{g_i(\tilde{x}^{j,\ell})}{\gamma_j}\right\} \nabla g_i(\tilde{x}^{j,\ell}). \quad (4.4)$$

Por outro lado, se consideramos que $\tilde{x}^{j,\ell}$ é uma solução exata para o j -ésimo subproblema, então o lado direito de (4.3) (ou (4.4)) será igual a zero. Caso contrário $\tilde{x}^{j,\ell}$ estará próximo da solução exata e assim o lado direito de (4.3) (ou (4.4)) é não nulo e o chamaremos de $e^{j,\ell}$. Portanto, se $e^{j,\ell}$ satisfaz (3.10) com $x^j = \tilde{x}^j$, então tomamos $(\tilde{x}^{j,\ell}, e^{j,\ell}) = (x^j, e^j)$ e o algoritmo continua para o Passo 3. Caso a solução $\tilde{x}^{j,\ell}$ não seja suficientemente boa, ou seja, não satisfaça (3.9) e (3.10), então faremos um passo adicional na sub-rotina e o teste será repetido para os valores $\tilde{x}^{j,\ell+1}$ e $e^{j,\ell+1}$ obtidos da sub-rotina. Assim, vamos trabalhar o caso em f e g_i sejam funções suaves. Neste caso, resolver o Passo 2 em cada iteração de IALPM consiste em resolver uma equação não linear $G^j(x) = 0$, com $G^j : \mathbb{R}^n \rightarrow \mathbb{R}^n$ dada por

$$G^j(x) = \gamma_j(x - x^j) + \nabla F_x(x) + \sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(x)}{\gamma_j}\right) \nabla g_i(x), \quad (4.5)$$

e no caso da penalidade quadrática

$$G^j(x) = \gamma_j(x - x^j) + \nabla F_x(x) + \sum_{i=1}^m \max\left\{0, \lambda_i^j + \frac{g_i(x)}{\gamma_j}\right\} \nabla g_i(x). \quad (4.6)$$

Observe que a diferença entre as equações (4.5) e (4.6) é que a penalidade G^j é uma função contínua mas dependendo da penalidade escolhida pode não ser diferenciável, pois o termo $\max \left\{ 0, \lambda_i^j + \frac{g_i(x)}{\gamma_j} \right\}$ da penalidade quadrática não é diferenciável na origem, o que impede o uso de métodos envolvendo derivadas como no caso do método de Newton. Assim, para implementar o algoritmo IALPM utilizando a penalidade quadrática, vamos considerar a mesma técnica apresentada em [25] para contornar a não diferenciabilidade presente na definição de G^j em (4.6).

Para cada índice $i = 1, \dots, m$ consideremos a função

$$p_i^j(x) = \max \left\{ 0, \lambda_i^j + \frac{g_i(x)}{\gamma_j} \right\}. \quad (4.7)$$

Como esse é o fator que aparece na definição da função $G^j(x)$ em (4.6), que o torna não diferenciável, a ideia é substituir (4.7) por

$$\bar{p}_i^j(x) = \begin{cases} 0, & \text{se } \lambda_i^j + \frac{g_i(x)}{\gamma_j} < -t, \\ \frac{1}{4t} \left(\lambda_i^j + \frac{g_i(x)}{\gamma_j} \right)^2 + \frac{1}{2} \left(\lambda_i^j + \frac{g_i(x)}{\gamma_j} \right) + \frac{t}{4}, & \text{se } -t \leq \lambda_i^j + \frac{g_i(x)}{\gamma_j} \leq t, \\ \lambda_i^j + \frac{g_i(x)}{\gamma_j}, & \text{se } \lambda_i^j + \frac{g_i(x)}{\gamma_j} > t, \end{cases} \quad (4.8)$$

em que t é um parâmetro real positivo. Note que a função \bar{p}_i^j é uma função diferenciável e a sua derivada, denotada por $(\bar{p}_i^j)'$, é dada por

$$\begin{aligned} (\bar{p}_i^j)'(x) &= \begin{cases} 0, & \text{se } \lambda_i^j + \frac{g_i(x)}{\gamma_j} < -t, \\ \frac{1}{2t} \left(\lambda_i^j + \frac{g_i(x)}{\gamma_j} \right) \frac{\nabla g_i(x)}{\gamma_j} + \frac{\nabla g_i(x)}{2\gamma_j}, & \text{se } -t \leq \lambda_i^j + \frac{g_i(x)}{\gamma_j} \leq t, \\ \frac{\nabla g_i(x)}{\gamma_j}, & \text{se } \lambda_i^j + \frac{g_i(x)}{\gamma_j} > t, \end{cases} \\ &= \frac{\nabla g_i(x)}{2\gamma_j} \begin{cases} 0, & \text{se } \lambda_i^j + \frac{g_i(x)}{\gamma_j} < -t, \\ \frac{1}{t} \left(\lambda_i^j + \frac{g_i(x)}{\gamma_j} \right) + 1, & \text{se } -t \leq \lambda_i^j + \frac{g_i(x)}{\gamma_j} \leq t, \\ 2, & \text{se } \lambda_i^j + \frac{g_i(x)}{\gamma_j} > t. \end{cases} \end{aligned} \quad (4.9)$$

Assim, definimos a função diferenciável \bar{G}^j dada por

$$\bar{G}^j(x) = \gamma_j(x - x^j) + \nabla F_x(x) + \sum_{i=1}^m \bar{p}_i^j(x) \nabla g_i(x). \quad (4.10)$$

Com esta definição, vamos ver que quando $t \rightarrow 0$ então temos que \bar{p}_i^j se aproxima de p_i^j .

Proposição 4.1. *Para cada $i = 1, \dots, m$, sejam \bar{p}_i^j e p_i^j as funções definidas em (4.7) e (4.8), respectivamente. Então*

$$\lim_{t \rightarrow 0} \bar{p}_i^j(x) = p_i^j(x), \forall i = 1, \dots, m \text{ e } \forall x \in \mathbb{R}^n. \quad (4.11)$$

Demonstração. Sejam $i \in \{1, \dots, m\}$ e $a_i^j = \lambda_i^j + \frac{g_i(x)}{\gamma_j}$ fixados. Observe que se $-t \leq a_i^j \leq t$, então temos que

$$\left| \frac{1}{4t} a_i^j \right| \leq \frac{1}{4},$$

e além disso,

$$\lim_{t \rightarrow 0} a_i^j = 0.$$

Em consequência,

$$\lim_{t \rightarrow 0} \frac{1}{4t} (a_i^j)^2 + \frac{1}{2} (a_i^j) + \frac{t}{4} = \lim_{t \rightarrow 0} \left(\frac{1}{4t} a_i^j + \frac{1}{2} \right) (a_i^j) + \frac{t}{4} = 0.$$

Logo, da definição de \bar{p}_i^j podemos concluir que

$$\begin{aligned} \lim_{t \rightarrow 0} \bar{p}_i^j &= \begin{cases} 0, & \text{se } a_i^j < 0, \\ 0, & \text{se } a_i^j = 0, \\ a_i^j, & \text{se } a_i^j > 0, \end{cases} \\ &= \begin{cases} 0, & \text{se } a_i^j \leq 0, \\ a_i^j, & \text{se } a_i^j > 0, \end{cases} \\ &= \max \{0, a_i^j\} \\ &= \max \left\{ 0, \lambda_i^j + \frac{g_i(x)}{\gamma_j} \right\} \\ &= p_i^j. \end{aligned}$$

□

Agora, estamos em condições de apresentar o método de Newton para resolver o problema $\bar{G}^j(x) = 0$, considerando \bar{G}^j como em (4.10) ou $\bar{G}^j(x) = G^j(x)$ com $G^j(x)$ dada em (4.5).

Consideremos a seguinte hipótese, introduzida com o objetivo de garantir unicidade da solução no Passo 2 do algoritmo IALPM.

P7: Suponha que $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, para $i = 1, \dots, m$, definida como em (3.1) e $U : \mathbb{R}^n \rightarrow \mathbb{R}^n$ dada por

$$U(x) = \nabla F_x(x), \quad \forall x \in \mathbb{R}^n,$$

são funções duas vezes continuamente diferenciáveis, em que $F_x : \mathbb{R}^n \rightarrow \mathbb{R}$ é dada como em (2.31).

Lema 4.2. *Suponha que a hipótese P7 é válida. Então a matriz Jacobiana de $\bar{G}^j(x)$, com*

\bar{G}^j definida como em (4.10) ou (4.5), são dadas respectivamente, por

$$J\bar{G}^j(x) = \gamma_j I_{n \times n} + JU(x) + \sum_{i=1}^m (\bar{p}^j)'_i(x) [\nabla g_i(x)]^T + \sum_{i=1}^m \bar{p}_i^j(x) \nabla^2 g_i(x) \quad (4.12)$$

$$J\bar{G}^j(x) = \gamma_j I_{n \times n} + JU(x) + \sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(x)}{\gamma_j}\right) \nabla g_i(x) [\nabla g_i(x)]^T + \sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(x)}{\gamma_j}\right) \nabla^2 g_i(x) \quad (4.13)$$

em que $I_{n \times n}$ é a matriz identidade, $JU(x)$ é a jacobiana de U no ponto x , e (\bar{p}_i^j) e $(\bar{p}_i^j)'$ são dadas como em (4.8) e (4.9) respectivamente.

Demonstração. Como g_i e U são diferenciáveis, basta aplicar as propriedades de diferenciação nas equações (4.10) e (4.5) para garantir o resultado desejado. \square

Na sequência reescrevemos o algoritmo IALPM utilizando o método de Newton para resolver o Passo 2.

Algoritmo 4.3 (IALPMN). *IALPM + Método de Newton*

Dados: $x^0 \in \mathbb{R}^n$, $\lambda^0 \in \mathbb{R}_+^m$, $t > 0$, $\sigma \in (0, 1)$, $\epsilon_1, \epsilon_2 > 0$, e uma sequência limitada $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma}]$ para algum $\bar{\gamma} > \theta$, com θ a constante de submonotonicidade. $j = 0$ e $aux = 1$.

ENQUANTO $aux = 1$

Faça $\ell = 0$, $z^\ell = x^j$, $aux1 = 1$ e calcule $\bar{G}^j(z^\ell)$ utilizando (4.10) (ou (4.5))

ENQUANTO $aux1 = 1$

Calcule $J\bar{G}^j(z^\ell)$ utilizando (4.12) (ou (4.13)) e determine

a direção de Newton $d^\ell = - (J\bar{G}^j(z^\ell))^{-1} \bar{G}^j(z^\ell)$

Faça $z^{\ell+1} = z^\ell + d^\ell$

Calcule $\bar{G}^j(z^{\ell+1})$ como em (4.10) (ou (4.5)) e defina $e^\ell = \bar{G}^j(z^{\ell+1})$.

SE $\|e^\ell\| \leq \min\{\epsilon_2, \frac{1}{2}\sigma\gamma_j(\min\{\|x^j - z^{\ell+1}\|, 1\})^2\}$ ou $\ell > 1000$

Faça $aux1 = 0$.

SENÃO

Faça $\ell = \ell + 1$.

FIM

FIM

ATUALIZE para cada $1 \leq i \leq m$,

$$\lambda_i^{j+1} = \max\left\{0, \lambda_i^j + \frac{g_i(z^{\ell+1})}{\gamma_j}\right\}, \quad \text{ou} \quad \left(\lambda_i^{j+1} = \lambda_i^j \exp\left(\frac{g_i(z^{\ell+1})}{\gamma_j}\right)\right)$$

SE $\|(x^j, \lambda^j) - (z^{\ell+1}, \lambda^{j+1})\| \leq \epsilon_1$

$aux = 0$

SENÃO Tome $v^j = \gamma_j(x^j - z^{\ell+1}) + e^\ell$

$$\text{Atualize } x^{j+1} = x^j - \frac{1}{\gamma_j} \left(1 - \frac{1}{\|v^j\|^2} \langle e^\ell, v^j \rangle\right) v^j \quad \text{e} \quad j = j + 1.$$

FIM

FIM

Note que o algoritmo descrito acima possui dois laços; um interno (quando $aux1 = 1$) e outro externo (quando $aux = 1$). O laço interno utiliza o método de Newton para resolver o sistema não linear $\bar{G}^j = 0$ dado como em (4.10) (ou $G^j = 0$ em (4.5)), que como sabemos é equivalente a resolver o Passo 2 de IALPM. Assim, uma vez resolvido o sistema, vamos ao laço externo para atualizar os multiplicadores e verificar se o critério de parada é satisfeito. Caso não se cumpra que $\|(x^j, \lambda^j) - (z^{\ell+1}, \lambda^{j+1})\| \leq \epsilon_1$ então utilizamos a Observação 3.3 para atualizar o novo ponto x^{j+1} .

Em cada passo do laço interno é necessário calcular a direção de Newton $d^\ell = -(J\bar{G}^j(z^\ell))^{-1} \bar{G}^j(z^\ell)$, a qual depende da função de penalidade considerada. Assim, é preciso estabelecer condições para que a inversa da matriz $J\bar{G}^j$ exista (ou, que a solução do sistema $(J\bar{G}^j(z^\ell)) d^\ell = -\bar{G}^j(z^\ell)$ seja única) e assim garantir que o Algoritmo IALPMN está bem definido.

Teorema 4.4. *Suponha que f satisfaz P1-P4, P4' e P7, com constante de submonotonicidade $\theta = 0$ (isto é, f é monótona). Então o Algoritmo 4.3 está bem definido.*

Demonstração. Vamos verificar que a matriz $J\bar{G}^j$ definida em (4.12) é definida positiva para cada $x \in \mathbb{R}^n$. Para isto analisaremos cada termo no somatório (4.12). Sejam $t > 0$, $x \in \mathbb{R}^n$ e $\gamma_j > 0$ para cada j . Observe inicialmente que o primeiro termo em (4.12) $\gamma_j I_{n \times n}$ é uma matriz definida positiva, pois cada termo da sequência γ_j é positivo. Para os termos restantes na soma basta mostrar que são semidefinidos positivos. Na Proposição 6.3 de [15], foi mostrado que se f é monótona então o operador U , dado na hipótese P7, é monótono e da Proposição 4.1.6 de [13] concluímos que a matriz jacobiana de U é semidefinida positiva. Note que se $-t \leq \lambda^j + \frac{g_i(x)}{\gamma_j} \leq t$, então temos que

$$\left| \frac{1}{2t} \left(\lambda^j + \frac{g_i(x)}{\gamma_j} \right) \right| \leq \frac{1}{2}.$$

Somando $\frac{1}{2}$ a ambos lados da desigualdade acima temos que

$$0 \leq \frac{1}{2t} \left(\lambda^j + \frac{g_i(x)}{\gamma_j} \right) + \frac{1}{2} \leq 1.$$

Assim, combinando esta última desigualdade com (4.9), obtemos que $\sum_{i=1}^m (\bar{p}^j)'_i(x) [\nabla g_i(x)]^T$ é uma soma de matrizes semidefinidas positivas. Finalmente, como para $1 \leq i \leq m$, as funções g_i são convexas e \bar{p}_i são não negativas, então temos que $\sum_{i=1}^m \bar{p}_i^j(x) \nabla^2 g_i(x)$ também é uma soma de matrizes semidefinidas positivas, tendo assim que a matriz $J\bar{G}^j$ em (4.12) é definida positiva e consequentemente o Algoritmo 4.3 está bem definido. \square

Observação 4.5. Note que o teorema anterior também é válido se consideramos a função exponencial, pois por convexidade o termo $\sum_{i=1}^m \lambda_i^j \exp\left(\frac{g_i(x)}{\gamma_j}\right) \nabla g_i(x) [\nabla g_i(x)]^T$ é soma de matrizes semidefinidas positivas. A convergência destes Algoritmos está garantida pois, como vimos na Proposição 4.1, se $t \rightarrow 0$ então $\bar{p}_i^j(x) \rightarrow p_i^j(x)$ e conseqüentemente $x^j(t) \rightarrow x^j$, com $x^j(t)$ sendo raiz de \bar{G}^j dado como em (4.10) e x^j sendo raiz de G^j dado em (4.6).

4.2 Método de Projeção do Subgradiente para resolver o Passo 2 de IALPM

Nesta seção, apresentamos uma nova metodologia para resolver os subproblemas gerados no Passo 2, do Algoritmo IALPM. O método que vamos utilizar é baseado em [18] e consiste basicamente de dois passos; uma maximização inexata de uma função contínua sobre um conjunto compacto e uma projeção aproximada sobre um hiperplano que separa o ponto calculado na maximização do conjunto solução do problema. Para o desenvolvimento e convergência deste algoritmo precisamos modificar um pouco as hipóteses iniciais da função f . Em princípio vamos exigir que a função f seja contínua em ambas variáveis, e como um segundo caso, consideramos posteriormente que f seja diferenciável.

4.2.1 Método de Projeção do Subgradiente

Para este método vamos considerar que f seja uma função contínua em ambas variáveis e que satisfaça as propriedades P1-P3. Em sua forma geral este método considera um conjunto $\Gamma \subset \mathbb{R}^n$ não vazio, convexo e fechado (não necessariamente da mesma forma que C). Como em [18] consideremos o conjunto $L_f(y) = \{x \in \Gamma : f(y, x) \leq 0\}$ que gera um problema de viabilidade convexa, denotado por (CFP), e definido como:

$$\text{Encontrar } \bar{x} \in \Gamma \text{ tal que } \bar{x} \in \bigcap_{y \in \Gamma} L_f(y), \quad (4.14)$$

o qual equivale a

$$\text{Encontrar } \bar{x} \in \Gamma \text{ tal que } f(y, \bar{x}) \leq 0, \text{ para todo } y \in \Gamma.$$

Este problema foi introduzido em [19] como um problema auxiliar para resolver o problema de equilíbrio. A relação que evidencia este fato é dada no seguinte teorema mostrado em [18], que devido à sua importância, para o nosso trabalho, reescrevemos aqui.

Teorema 4.6. *Seja $\Gamma \subset \mathbb{R}^n$ um conjunto não vazio, convexo e fechado. Considere o problema de equilíbrio $EP(f, \Gamma)$ dado como em (1.1), com f contínua satisfazendo P1-P3. Então o conjunto solução do problema CFP dado em (4.14) é um subconjunto do conjunto solução do problema $EP(f, \Gamma)$.*

Demonstração. Seja \bar{x} solução de CFP e considere $y \in \Gamma$. Defina $z_t = ty + (1 - t)\bar{x}$, para cada $t \in (0, 1)$. Note que $z_t \in \Gamma$, pois Γ é convexo. Agora, combinando P1 e a convexidade de f , na segunda variável, obtemos que

$$0 = f(z_t, z_t) \leq tf(z_t, y) + (1 - t)f(z_t, \bar{x}), \quad (4.15)$$

para cada $t \in (0, 1)$. Assim, tomando $\bar{x} \in \bigcap_{y \in \Gamma} L_f(y)$ e $z_t \in \Gamma$, então por definição temos que $f(z_t, \bar{x}) \leq 0$ e por (4.15) $0 \leq f(z_t, y)$, para todo $t \in (0, 1)$. Logo, como f é contínua na primeira variável concluímos que

$$0 \leq \lim_{t \rightarrow 0} f(z_t, y) = f(\bar{x}, y),$$

para todo $y \in \Gamma$, isto é, $\bar{x} \in S(f, \Gamma)$.

□

A recíproca deste teorema não é necessariamente verdadeira. Para ver isto, consideremos os seguintes exemplos, apresentados em [18].

1. Seja $f(x, y) = (x - y)^2$ e $\Gamma = [0, 1]$. O conjunto solução do problema (4.14) é vazio, enquanto que $S(f, \Gamma) = \Gamma$.
2. Considere $f(x, y) = \max\{0, |x - y| - 1\}$ e $\Gamma = [0, 2]$. Como f é não negativa em Γ então $S(f, \Gamma) = \Gamma$. Por outro lado, observe que para todo $x \in \Gamma - \{1\}$ existe $y \in \Gamma$ tal que $f(x, y) > 0$, assim, a solução do problema (4.14) é $x = 1$.

A seguinte proposição demonstrada em [14], estabelece condições suficientes para que a recíproca do Teorema 4.6 seja verdadeira.

Proposição 4.7. *Considere f satisfazendo P1-P3 e uma dentre as hipóteses $P4^*$ e $P4'$, então o conjunto solução de $EP(f, \Gamma)$ e CFP coincidem.*

Demonstração. Basta mostrar que $S(f, \Gamma)$ está contido no conjunto solução do problema CFP. Suponha inicialmente que f cumpre $P4'$. Tome $\bar{x} \in S(f, \Gamma)$, $y \in \Gamma$ e para cada $t \in (0, 1)$, defina $z_t = ty + (1 - t)\bar{x}$. Por $P4'$,

$$tf(y, z_t) + (1 - t)f(\bar{x}, z_t) \leq 0. \quad (4.16)$$

Agora, como $\bar{x} \in S(f, \Gamma)$ e $z_t \in \Gamma$ então $f(\bar{x}, z_t) \geq 0$ e portanto, segue de (4.16) que $f(y, z_t) \leq 0$. Logo, tomando limite e usando a continuidade de f obtemos que

$$\lim_{t \rightarrow 0} f(y, z_t) = f(y, \bar{x}) \leq 0,$$

para todo $y \in \Gamma$, ou seja, \bar{x} é solução de CFP. Agora, suponha que f cumpre P4*. Tome $\bar{x} \in S(f, \Gamma)$, assim, $f(\bar{x}, y) \geq 0$ para todo $y \in \Gamma$ e por P4* concluimos que $f(y, \bar{x}) \leq 0$ para todo $y \in \Gamma$, ou seja, \bar{x} resolve CFP. \square

A seguir, consideramos uma propriedade importante, dada em [18], para o problema de equilíbrio. Nesta propriedade é dado o conceito de solução local, e a partir disto, é mostrado que soluções locais do $EP(f, \Gamma)$ são soluções globais.

Teorema 4.8. *Seja f contínua satisfazendo P1-P2. Se existe um aberto $U \subset \mathbb{R}^n$ e $x^* \in \Gamma \cap U$ tal que $f(x^*, y) \geq 0$ para todo $y \in \Gamma \cap U$ então x^* resolve $EP(f, \Gamma)$.*

Demonstração. Seja $I_\Gamma : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ a função indicadora de Γ , isto é,

$$I_\Gamma(x) = \begin{cases} 0 & \text{se } x \in \Gamma \\ +\infty & \text{se } x \notin \Gamma. \end{cases}$$

Defina $h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ como $h(y) = f(x^*, y) + I_\Gamma(y)$. De P1, $f(x^*, x^*) = 0 \leq f(x^*, y)$, para todo $y \in \Gamma \cap U$ e assim, $h(x^*) \leq h(y)$ para $y \in U$, ou seja, x^* é minimizador local de h em U . Além disso, como h é convexa, pois f cumpre P2 e I_Γ é convexa então x^* é um minimizador global de h e em particular, $0 = f(x^*, x^*) = h(x^*) \leq h(y) = f(x^*, y)$ para todo $y \in \Gamma$. Portanto, $x^* \in S(f, \Gamma)$. \square

Tendo em vista o problema (4.14) e sua relação com $EP(f, \Gamma)$, consideramos a seguir uma adaptação do Algoritmo 2, apresentado em [18], para resolver problemas de viabilidade convexa (4.14) (e conseqüentemente o problema de equilíbrio) para uma função contínua f que satisfaça as hipóteses P1-P3. Como $\Gamma \subset \mathbb{R}^n$ é um conjunto não vazio, convexo e fechado, podemos definir a projeção ortogonal de um ponto sobre o conjunto Γ . Denotaremos esta função por $P_\Gamma : \mathbb{R}^n \rightarrow \Gamma$.

Algoritmo 4.9 (SPMEP). *Método de Subgradiente projetado para $EP(f, \Gamma)$.*

Dados $\alpha \in (0, 1)$, $x^0 \in \Gamma$, $\rho_0 = \|x^0\|$, $\varepsilon_0 \geq 0$ e $r > 0$. Considere duas sequências, uma de parâmetros de relaxação $(\beta_\ell)_{\ell \in \mathbb{N}} \subset [\alpha, 1]$ e outra de precisão de parâmetros $(\varepsilon_\ell)_{\ell \in \mathbb{N}} \subset \mathbb{R}_+$ tal que $\varepsilon_\ell \rightarrow 0$ quando $\ell \rightarrow \infty$.

$\ell = 0$, $aux = 1$.

ENQUANTO $aux = 1$

$$\text{DEFINA } \Gamma_\ell = \{x \in \Gamma : \|x\| \leq \rho_\ell + r\} = \overline{B(0, \rho_\ell + r)} \cap \Gamma. \quad (\text{a})$$

$$\text{ENCONTRE } y^\ell \in \Gamma_\ell \text{ tal que } f(y^\ell, x^\ell) \geq 0 \text{ e} \quad (\text{b})$$

$$\max_{y \in \Gamma_\ell} f(y, x^\ell) \leq f(y^\ell, x^\ell) + \varepsilon_\ell$$

$$\text{SE } 0 \in \partial F_{y^\ell}(x^\ell) \quad (\text{c})$$

$$aux = 0, x^{\ell+1} = y^\ell.$$

$$\text{SENÃO tome } \xi^\ell \in \partial F_{y^\ell}(x^\ell), \text{ defina} \quad (\text{d})$$

$$z^\ell = x^\ell - \beta_\ell \left(\frac{f(y^\ell, x^\ell)}{\|\xi^\ell\|^2} \right) \xi^\ell \text{ e atualize,}$$

$$\begin{aligned}x^{\ell+1} &= z^\ell + \beta_\ell(P_\Gamma(z^\ell) - z^\ell), \\ \rho_{\ell+1} &= \max \{ \rho_\ell, \|x^{\ell+1}\| \},\end{aligned}$$

FIM

$$\ell = \ell + 1$$

FIM

Observe que, para este algoritmo, o conjunto $\Gamma_\ell = \{x \in \Gamma : \|x\| \leq \rho_\ell + r\} = \overline{B(0, \rho_\ell + r)} \cap \Gamma$ em que $\rho_\ell = \|x^\ell\|$ e $r > 0$, apresenta uma pequena diferença com respeito ao algoritmo definido em [18], pois em [18] os autores consideram o conjunto $\Gamma_\ell = \{x \in \Gamma : \|x\| \leq \rho_\ell + 1\} = \overline{B(0, \rho_\ell + 1)} \cap \Gamma$. Afirmamos que, embora tenha sido feita uma adaptação no raio da bola sobre o conjunto que queremos maximizar, os resultados obtidos em [18] continuam sendo válidos. Para esclarecer isto, vamos reescrever a análise de convergência do algoritmo SPMEP com a adaptação feita no conjunto Γ_ℓ . Começamos esta tarefa mostrando que o algoritmo SPMEP está bem definido.

Proposição 4.10. *O algoritmo SPMEP está bem definido.*

Demonstração. Sabemos que por definição a sequência $(\rho_\ell)_{\ell \in \mathbb{N}}$ é não decrescente, então por definição $\Gamma_\ell \subset \Gamma_{\ell+1}$ para todo $\ell \in \mathbb{N}$. Como $x^0 \in \Gamma$, então temos que $x^0 \in \Gamma_0$ e portanto $x^0 \in \Gamma_\ell$ para todo $\ell \in \mathbb{N}$, o que mostra que cada conjunto Γ_ℓ é não vazio e compacto, pois Γ é fechado. Logo, pela continuidade de f obtemos que a função $f(\cdot, x^\ell)$ atinge seu máximo em Γ_ℓ o que garante a existência de y^ℓ como solução do subproblema aproximado

$$\max_{y \in \Gamma_\ell} f(y, x^\ell) \leq f(y^\ell, x^\ell) + \varepsilon_\ell.$$

Finalmente, utilizando a convexidade de $f(y^\ell, \cdot)$ e o fato de que Γ é convexo e fechado obtemos a unicidade na projeção exigida na atualização de $x^{\ell+1}$.

□

O seguinte lema será utilizado na demonstração do teorema de convergência do algoritmo SPMEP.

Lema 4.11. *Considere f contínua, satisfazendo P1-P3 e $0 \neq \xi \in \partial F_{\bar{y}}(\bar{x})$, com $\bar{x}, \bar{y} \in \Gamma$. Se $P_S : \mathbb{R}^n \rightarrow S$ é a projeção ortogonal sobre $S = \{v \in \mathbb{R}^n \mid \langle \xi, v - \bar{x} \rangle + f(\bar{y}, \bar{x}) \leq 0\}$ então*

$$P_S(\bar{x}) = \bar{x} - \left[\frac{f(\bar{y}, \bar{x})}{\|\xi\|^2} \right] \xi. \quad (4.17)$$

Demonstração. Basta considerar o problema

$$\min_{z \in S} \frac{1}{2} \|z - \bar{x}\|^2$$

e ver que a solução é dada por $z = \bar{x} - \left[\frac{f(\bar{y}, \bar{x})}{\|\xi\|^2} \right] \xi$, isto é, $P_S(\bar{x}) = \bar{x} - \left[\frac{f(\bar{y}, \bar{x})}{\|\xi\|^2} \right] \xi$.

□

Além do Lema 4.11, para o teorema de convergência do Algoritmo SPMEP, vamos precisar da seguinte hipótese auxiliar considerada em [18].

A : Se A, B são dois subconjuntos compactos de Γ então $\bigcup_{(x,y) \in A \times B} \partial F_x(y)$ é limitado.

Como mencionamos anteriormente o teorema a seguir é dado essencialmente em [18]. Rescreveremos a prova, para evidenciar que embora modificamos o conjunto Γ_ℓ no algoritmo SPMEP, os resultados obtidos pelos autores em [18] continuam sendo válidos.

Teorema 4.12. *Sejam $(x^\ell)_{\ell \in \mathbb{N}}, (y^\ell)_{\ell \in \mathbb{N}}$ sequências geradas pelo Algoritmo SPMEP. Suponha que f satisfaz P1-P3 e a hipótese **A**. Então ou o Algoritmo para em um passo ℓ e em tal caso y^ℓ é solução de $EP(f, \Gamma)$, ou gera uma sequência infinita $(x^\ell)_{\ell \in \mathbb{N}}$, e em tal caso:*

- i) *Se o conjunto solução de CFP é não vazio então $(x^\ell)_{\ell \in \mathbb{N}}$ converge para uma solução de $EP(f, \Gamma)$.*
- ii) *Se $EP(f, \Gamma)$ não admite solução então $(x^\ell)_{\ell \in \mathbb{N}}$ não é convergente.*

Demonstração. Vamos discutir inicialmente o caso de terminação finita que acontece quando $0 \in \partial F_{y^\ell}(x^\ell)$. Neste caso, pela hipótese P2, x^ℓ é um minimizador da função $F_{y^\ell}(\cdot)$ em \mathbb{R}^n e em particular em Γ . Portanto, usando a relação (b) do algoritmo SPMEP concluímos que

$$0 \leq f(y^\ell, x^\ell) = F_{y^\ell}(x^\ell) \leq F_{y^\ell}(w) = f(y^\ell, w),$$

para todo $w \in \Gamma$. Assim, por definição significa que $y^\ell \in S(f, \Gamma)$.

(i) Suponhamos que x^* é solução do problema CFP e sejam $S_\ell = \{v \in \mathbb{R}^n \mid \langle \xi^\ell, v - x^\ell \rangle + f(y^\ell, x^\ell) \leq 0\}$ e $P_{S_\ell} : \mathbb{R}^n \rightarrow S_\ell$ a projeção ortogonal sobre S_ℓ . Note que $x^* \in S_\ell$ pois, em particular, $x^* \in \bigcap_{\ell=1}^{\infty} L_f(y^\ell)$ e como $\xi^\ell \in \partial F_{y^\ell}(x^\ell)$, então pela definição do subgradiente para uma função convexa obtemos $0 \geq f(y^\ell, x^*) \geq f(y^\ell, x^\ell) + \langle \xi^\ell, x^* - x^\ell \rangle$. Do Lema 4.11 segue que a sequência $(z_\ell)_{\ell \in \mathbb{N}}$, definida no algoritmo SPMEP, pode ser reescrita como $z^\ell = x^\ell + \beta_\ell [P_{S_\ell}(x^\ell) - x^\ell]$, e com isto, obtemos que

$$\begin{aligned} \|z^\ell - x^*\|^2 &= \|x^\ell - x^* + \beta_\ell [P_{S_\ell}(x^\ell) - x^\ell]\|^2 \\ &= \|x^\ell - x^*\|^2 + 2\beta_\ell \langle x^\ell - x^*, P_{S_\ell}(x^\ell) - x^\ell \rangle + \beta_\ell^2 \|P_{S_\ell}(x^\ell) - x^\ell\|^2 \\ &= \|x^\ell - x^*\|^2 + \beta_\ell^2 \|P_{S_\ell}(x^\ell) - x^\ell\|^2 + 2\beta_\ell \langle x^\ell - x^* + P_{S_\ell}(x^\ell) - P_{S_\ell}(x^\ell), P_{S_\ell}(x^\ell) - x^\ell \rangle \\ &= \|x^\ell - x^*\|^2 + \beta_\ell^2 \|P_{S_\ell}(x^\ell) - x^\ell\|^2 - 2\beta_\ell \|x^\ell - P_{S_\ell}(x^\ell)\|^2 + 2\beta_\ell \langle P_{S_\ell}(x^\ell) - x^*, P_{S_\ell}(x^\ell) - x^\ell \rangle \\ &= \|x^\ell - x^*\|^2 + \beta_\ell(\beta_\ell - 2) \|x^\ell - P_{S_\ell}(x^\ell)\|^2 + 2\beta_\ell \langle P_{S_\ell}(x^\ell) - x^*, P_{S_\ell}(x^\ell) - x^\ell \rangle \\ &\leq \|x^\ell - x^*\|^2 - \beta_\ell(2 - \beta_\ell) \|x^\ell - P_{S_\ell}(x^\ell)\|^2, \end{aligned}$$

em que a última desigualdade vale pois $x^* \in S_\ell$, e com isto, $\langle P_{S_\ell}(x^\ell) - x^*, P_{S_\ell}(x^\ell) - x^\ell \rangle \leq 0$. Agora, usando a igualdade (4.17), na desigualdade acima concluímos que

$$\|z^\ell - x^*\|^2 \leq \|x^\ell - x^*\|^2 - \beta_\ell(2 - \beta_\ell) \left(\frac{f(y^\ell, x^\ell)}{\|\xi^\ell\|} \right)^2. \quad (4.18)$$

Um procedimento similar ao realizado acima para concluir a equação (4.18) pode ser realizado para a atualização $x^{\ell+1}$ do Algoritmo SPMEP e assim obter que

$$\|x^{\ell+1} - x^*\|^2 \leq \|z^\ell - x^*\|^2 - \beta_\ell(2 - \beta_\ell) \|z^\ell - P_\Gamma(z^\ell)\|^2. \quad (4.19)$$

Logo, utilizando a desigualdade (4.18) em (4.19) obtemos

$$\begin{aligned} \|x^{\ell+1} - x^*\|^2 &\leq \|x^\ell - x^*\|^2 - \beta_\ell(2 - \beta_\ell) \left[\left(\frac{f(y^\ell, x^\ell)}{\|\xi^\ell\|} \right)^2 + \|z^\ell - P_\Gamma(z^\ell)\|^2 \right] \\ &\leq \|x^\ell - x^*\|^2, \end{aligned} \quad (4.20)$$

em que a segunda desigualdade vale porque $\beta_\ell(2 - \beta_\ell) \left[\left(\frac{f(y^\ell, x^\ell)}{\|\xi^\ell\|} \right)^2 + \|z^\ell - P_\Gamma(z^\ell)\|^2 \right] \geq 0$. De (4.20), obtemos que a sequência $(\|x^\ell - x^*\|)_{\ell \in \mathbb{N}}$ é monótona, limitada e portanto convergente. Além disso, a sequência $(x^\ell)_{\ell \in \mathbb{N}}$ é limitada e conseqüentemente $(\rho_\ell)_{\ell \in \mathbb{N}}$ também. Assim, seja $\rho^* = \sup_{\ell \in \mathbb{N}} \rho_\ell$ o limitante superior da sequência $(\rho_\ell)_{\ell \in \mathbb{N}}$. Então como $y^\ell \in \Gamma_\ell$ obtemos que $\|y^\ell\| \leq \rho_\ell + r \leq \rho^* + r$ para todo $\ell \in \mathbb{N}$, e portanto $(y^\ell)_{\ell \in \mathbb{N}}$ também é uma sequência limitada. Logo, pelo algoritmo sabemos que $(\beta_\ell)_{\ell \in \mathbb{N}} \subset [\alpha, 1]$, com $\alpha \in (0, 1)$ assim,

$$-\beta_\ell(2 - \beta_\ell) = -2\beta_\ell + (\beta_\ell)^2 \leq -2\beta_\ell + \beta_\ell = -\beta_\ell \leq -\alpha.$$

Agora, se combinamos esta desigualdade com (4.20) obtemos que

$$0 \leq \alpha \left[\left(\frac{f(y^\ell, x^\ell)}{\|\xi^\ell\|} \right)^2 + \|z^\ell - P_\Gamma(z^\ell)\|^2 \right] \leq \|x^\ell - x^*\|^2 - \|x^{\ell+1} - x^*\|^2,$$

e como $\alpha > 0$ temos que

$$0 \leq \left(\frac{f(y^\ell, x^\ell)}{\|\xi^\ell\|} \right)^2 + \|z^\ell - P_\Gamma(z^\ell)\|^2 \leq \frac{1}{\alpha} \left[\|x^\ell - x^*\|^2 - \|x^{\ell+1} - x^*\|^2 \right]. \quad (4.21)$$

Assim, tomando limite nesta última desigualdade e usando o fato de que $(\|x^\ell - x^*\|)_{\ell \in \mathbb{N}}$ converge, concluímos que

$$\lim_{\ell \rightarrow \infty} \|z^\ell - P_\Gamma(z^\ell)\| = \lim_{\ell \rightarrow \infty} \frac{f(y^\ell, x^\ell)}{\|\xi^\ell\|} = 0. \quad (4.22)$$

Por outro lado, como as sequências $(x^\ell)_{\ell \in \mathbb{N}} \subset \Gamma$ e $(y^\ell)_{\ell \in \mathbb{N}} \subset \Gamma$ são limitadas e $\xi^\ell \in \partial F_{y^\ell}(x^\ell)$ então pela hipótese **A**, $(\xi^\ell)_{\ell \in \mathbb{N}}$ é limitada. Logo, de (4.22) obtemos que $\lim_{\ell \rightarrow \infty} f(y^\ell, x^\ell) = 0$. Agora, se $(x^{\ell_k})_{k \in \mathbb{N}}$ e $(y^{\ell_k})_{k \in \mathbb{N}}$ são subsequências de $(x^\ell)_{\ell \in \mathbb{N}}$ e $(y^\ell)_{\ell \in \mathbb{N}}$, respectivamente, tais que $x^{\ell_k} \rightarrow \bar{x}$ e $y^{\ell_k} \rightarrow \bar{y}$ então pela continuidade de f , concluímos que

$$f(\bar{x}, \bar{y}) = 0. \quad (4.23)$$

Vejamos que \bar{x} é solução de $EP(f, \Gamma)$. Com efeito, dado $\delta \in (0, r)$, defina $C(\delta) = B(0, \rho^* + r - \delta)$, com $\rho^* = \sup_{\ell \in \mathbb{N}} \rho_\ell$. Como $\|x^\ell\| \leq \rho_\ell \leq \rho^*$, para todo $\ell \in \mathbb{N}$, então tomando limite na subsequência obtemos que $\|\bar{x}\| \leq \rho^* < \rho^* + r - \delta$, isto é, $\bar{x} \in C(\delta) \cap \Gamma$. Logo, pelo Teorema 4.8 basta mostrar que $f(\bar{x}, y) \geq 0$, para todo $y \in \Gamma \cap C(\delta)$. Para isto, tome $L \in \mathbb{N}$ tal que $\rho_L \geq \rho^* - \delta$ e assim, $\rho^* + r - \delta \leq \rho_\ell + r$, para todo $\ell \geq L$ pois $(\rho_\ell)_{\ell \in \mathbb{N}}$ é não decrescente e portanto, $\Gamma \cap C(\delta) \subset \Gamma_\ell$, para todo $\ell \geq L$. Tome $z \in \Gamma \cap C(\delta)$ então $z \in \Gamma_\ell$, para todo $\ell \geq L$ e portanto, segue que

$$f(z, x^\ell) \leq \max_{y \in \Gamma_\ell} f(y, x^\ell) \leq f(y^\ell, x^\ell) + \varepsilon_\ell, \quad \forall \ell \geq L.$$

Logo, tomando limite na desigualdade acima sobre as subsequências $(x^{\ell_k})_{k \in \mathbb{N}}$, $(y^{\ell_k})_{k \in \mathbb{N}}$ e usando o fato de que f é contínua e $\varepsilon_\ell \rightarrow 0$, obtemos que

$$f(z, \bar{x}) = \lim_{k \rightarrow \infty} f(z, x^{\ell_k}) \leq \lim_{k \rightarrow \infty} [f(y^{\ell_k}, x^{\ell_k}) + \varepsilon_{\ell_k}] = f(\bar{x}, \bar{y}) = 0, \quad (4.24)$$

por (4.23). Como z é arbitrário em $\Gamma \cap C(\delta)$ então a desigualdade (4.24) estabelece que \bar{x} é solução do problema (4.14) no conjunto $\Gamma \cap C(\delta)$ e conseqüentemente, pelo Teorema 4.6, \bar{x} é solução de $EP(f, \Gamma \cap C(\delta))$. Finalmente, como $C(\delta)$ é aberto e usando o Teorema 4.8, obtemos que \bar{x} é solução de $EP(f, \Gamma)$.

Para finalizar este item, vejamos a unicidade dos pontos de acumulação da sequência $(x^\ell)_{\ell \in \mathbb{N}}$. Provamos que se \bar{x} é um ponto de acumulação da sequência $(x^\ell)_{\ell \in \mathbb{N}}$ então $\bar{x} \in \bigcap_{z \in \Gamma \cap C(\delta)} L_f(z)$, ou seja, $f(z, \bar{x}) \leq 0$, para todo $z \in \Gamma$ tal que $\|z\| \leq \rho^* + r - \delta$. Assim, como $\delta \in (0, r)$ foi escolhido de forma arbitrária, então podemos concluir que $f(z, \bar{x}) \leq 0$ para todo $z \in \Gamma$ e $\|z\| < \rho^* + r$ e, pela continuidade de f , $f(z, \bar{x}) \leq 0$ para todo $z \in \Gamma$ e $\|z\| \leq \rho^* + r$. Agora, como $y^\ell \in \Gamma_\ell$ vale que $\|y^\ell\| \leq \rho_\ell + r \leq \rho^* + r$ e portanto $f(y^\ell, \bar{x}) \leq 0$ para todo $\ell \in \mathbb{N}$, ou seja, $\bar{x} \in \bigcap_{\ell=1}^{\infty} L_f(y^\ell)$, e assim, podemos obter a desigualdade (4.20) com $x^* = \bar{x}$. Portanto, $(\|x^\ell - \bar{x}\|)_{\ell \in \mathbb{N}}$ é não crescente, não negativa e possui uma subsequência que converge para zero. Logo, $x^\ell \rightarrow \bar{x}$, ou seja, \bar{x} é o único ponto de acumulação de $(x^\ell)_{\ell \in \mathbb{N}}$.

(ii) Suponha por absurdo que $(x^\ell)_{\ell \in \mathbb{N}}$ converge, digamos para $x^* \in \mathbb{R}^n$. Logo, obtemos que $(x^\ell)_{\ell \in \mathbb{N}}$ é limitada e conseqüentemente como fizemos no item (i) podemos ver que $(y^\ell)_{\ell \in \mathbb{N}}$ também é limitada. Agora, note que, por P1, vale que $0 = f(y^\ell, y^\ell) \geq f(y^\ell, x^\ell) + \langle \xi^\ell, y^\ell - x^\ell \rangle$, com $\xi^\ell \in \partial F_{y^\ell}(x^\ell)$, ou seja, $y^\ell \in S_\ell$ para todo $\ell \in \mathbb{N}$. Assim, se

utilizarmos novamente o mesmo raciocínio que em (i), obtemos que a desigualdade (4.21) vale, tomando $x^* = y^\ell$. Logo, se (x^*, y^*) é um ponto de acumulação de $(x^\ell, y^\ell)_{\ell \in \mathbb{N}}$ então obtemos que $f(x^*, y^*) = 0$. A partir desta igualdade, o resto da prova do item anterior pode ser feita para concluir que x^* é solução do $\text{EP}(f, \Gamma)$ o que é um absurdo. \square

O Algoritmo SPMEP gera duas sequências $(x^\ell)_{\ell \in \mathbb{N}}$ e $(y^\ell)_{\ell \in \mathbb{N}}$ e o teorema anterior garante que alguma destas convergirá para uma solução do problema de equilíbrio, desde que as hipóteses sejam satisfeitas. Uma destas é que a hipótese **A** seja satisfeita. A seguinte proposição é uma caracterização dessa hipótese.

Proposição 4.13. *Se $f : \Gamma \times \Gamma \rightarrow \mathbb{R}$ é a restrição a $\Gamma \times \Gamma$ de uma função contínua $f : U \times U \rightarrow \mathbb{R}$ que satisfaz P1 e P2 em $U \times U$ em que U é um conjunto aberto que contém o conjunto Γ , então f satisfaz a hipótese **A**.*

Demonstração. Ver Proposição 4.3 de [18]. \square

Como o objetivo desta seção é utilizar o Algoritmo SPMEP para resolver o Passo 2 do algoritmo IALPM e lembrando que este passo consiste em resolver um problema irrestrito $\text{EP}(\tilde{\mathcal{L}}_j^e, \mathbb{R}^n)$, com $\tilde{\mathcal{L}}_j^e$ definida como em (3.9), então vejamos que as hipóteses do Teorema 4.12 são satisfeitas para este problema. De fato, sabemos que se f satisfaz P1-P3, então $\tilde{\mathcal{L}}_j^e$ também. Por outro lado, como os problemas gerados no Passo 2 são irrestritos então a Proposição 4.13 garante que $\tilde{\mathcal{L}}_j^e$ satisfaz a hipótese **A**, e como consequência o Teorema 4.12 afirma que o Algoritmo SPMEP converge para uma solução do problema $\text{EP}(\tilde{\mathcal{L}}_j^e, \mathbb{R}^n)$ sempre que o conjunto solução do problema CFP seja não vazio.

Note que para resolver o problema (3.9) é preciso fornecer o vetor e^j que posteriormente será usado para verificar se a condição (3.10) é satisfeita. Assim, isto nos leva a propor dois algoritmos para resolver o Passo 2 de IALPM. O primeiro destes algoritmos é definido para funções contínuas não necessariamente diferenciáveis e na sequência definiremos outro algoritmo para trabalhar o caso diferenciável. Para ambos algoritmos, vamos exigir que a sequência $(e^j)_{j \in \mathbb{N}}$ seja a sequência nula, pois vimos que o Algoritmo SPMEP tem a capacidade de resolver o Passo 2 do Algoritmo IALPM independentemente do valor de e^j e com esta escolha a condição (3.10) é satisfeita trivialmente. Por simplicidade vamos denotar como $\tilde{\mathcal{L}}_j$ para nos referir à função $\tilde{\mathcal{L}}_j^e$ com $e^j = 0$, além disso, como em (2.31), em alguns casos usaremos a notação $(\tilde{\mathcal{L}}_j)_y(x)$ em vez de $\tilde{\mathcal{L}}_j(y, x)$.

Algoritmo 4.14 (IALPMP). *IALPM + Método Do Subdiferencial Projetado.*

Dados: $x^0 \in \mathbb{R}^n$, $\lambda^0 > 0$, $\epsilon_1 > 0$, $\epsilon_2 > 0$, $\alpha \in (0, 1)$ e uma sequência limitada $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma}]$ para algum $\bar{\gamma} > \theta$, com θ a constante de submonotonicidade. $j = 0$ e $aux = 1$.

ENQUANTO $aux = 1$

Sejam $\ell = 0$, $r > 0$, $aux = 1$, $z^\ell = x^j$, $\rho_\ell = \|z^\ell\|$, $(\epsilon_\ell)_{\ell \in \mathbb{N}} \subset (0, +\infty)$ e $\beta_\ell \subset [\alpha, 1]$.

ENQUANTO $aux1 = 1$

DEFINA $\Gamma_\ell = \{x \in \mathbb{R}^n : \|x\| \leq \rho_\ell + r\} = \overline{B(0, \rho_\ell + r)}$.

ENCONTRE $y^\ell \in \Gamma_\ell$ tal que $\tilde{\mathcal{L}}_j(y^\ell, z^\ell) \geq 0$ e

$$\max_{y \in \Gamma_\ell} \tilde{\mathcal{L}}_j(y, z^\ell) \leq \tilde{\mathcal{L}}_j(y^\ell, z^\ell) + \varepsilon_\ell$$

SE $0 \in \partial(\tilde{\mathcal{L}}_j)_{y^\ell}(z^\ell) = \partial\tilde{\mathcal{L}}_j(y^\ell, z^\ell)$

$$aux = 0, z^{\ell+1} = y^\ell.$$

SENÃO tome $\xi^\ell \in \partial(\tilde{\mathcal{L}}_j)_{y^\ell}(z^\ell)$, e atualize

$$z^{\ell+1} = z^\ell - \beta_\ell \left(\frac{\tilde{\mathcal{L}}_j(y^\ell, z^\ell)}{\|\xi^\ell\|^2} \right) \xi^\ell$$

SE $\beta_\ell \frac{\tilde{\mathcal{L}}_j(y^\ell, z^\ell)}{\|\xi^\ell\|} < \epsilon_2$ ou $\ell > 10000$

$$aux1 = 0.$$

FIM

$$\rho_{\ell+1} = \max \{ \rho_\ell, \|z^{\ell+1}\| \},$$

FIM

Tome $\ell = \ell + 1$.

FIM

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

$$\lambda_i^{j+1} = \max \left\{ 0, \lambda_i^j + \frac{g_i(z^{\ell+1})}{\gamma_j} \right\}, \quad \text{ou} \quad \left(\lambda_i^{j+1} = \lambda_i^j \exp \left(\frac{g_i(z^{\ell+1})}{\gamma_j} \right) \right)$$

SE $\|(x^j, \lambda^j) - (z^{\ell+1}, \lambda^{j+1})\| \leq \epsilon_1$,

$$aux = 0$$

SENÃO atualize

$$x^{j+1} = z^{\ell+1}.$$

FIM

$j=j+1$

FIM

Note que neste algoritmo colocamos de forma implícita, na função $\tilde{\mathcal{L}}$, o tipo de penalidade que define o método, isto é, escolhida a penalidade, quadrática ou exponencial definimos a função $\tilde{\mathcal{L}}$. Além disso, a atualização das variáveis duais λ^{j+1} é determinada de acordo com a penalidade escolhida. No Algoritmo 4.14 foram feitas algumas simplificações tanto no Algoritmo interno SPMEP como no externo IALPM, além de definir os critérios de parada no algoritmo interno. As simplificações foram feitas na definição do conjunto Γ_ℓ e na atualização da sequência $(x^\ell)_{\ell \in \mathbb{N}}$; a primeira, deriva-se do fato de que os subproblemas gerados (3.9) são irrestritos e a segunda, combina o fato de que $e^j = 0$ para todo j com a observação 3.3, pois neste caso

$$x^{j+1} = x^j - \frac{1}{\gamma_j} \left(1 - \frac{1}{\|x^j\|^2} \langle e^j, v^j \rangle \right) v^j = z^{\ell+1},$$

em que $v^j = \gamma_j(x^j - z^{\ell+1})$.

Os critérios de parada do algoritmo interno foram estabelecidos levando em conta o Teorema 4.12 e a igualdade $\|z^{\ell+1} - z^\ell\| = \beta_\ell \frac{\tilde{\mathcal{L}}_j(y^\ell, z^\ell)}{\|\xi^\ell\|}$, além da já mencionada condição inicial $e^j = 0$, para todo $j \in \mathbb{N}$, pois neste caso, ambas equações (3.9) e (3.10) se reduzem a encontrar uma solução do EP($\tilde{\mathcal{L}}_j, \mathbb{R}^n$) e qualquer solução encontrada para uma tolerância ϵ_2 suficientemente pequeno.

Teorema 4.15. *Considere EP(f, C) e suponha que f é contínua e satisfaz P1-P5, com $\theta = 0$, isto é, f é monótona. Então a sequência $(x^j)_{j \in \mathbb{N}}$ gerada pelo algoritmo IALPMP converge para uma solução de EP(f, C) se este tem solução.*

Demonstração. A prova segue do fato de que $e^j = 0$ para todo j , e dos Teoremas 3.14 e 4.12. □

O algoritmo que acabamos de descrever tem a vantagem de resolver problemas de equilíbrio não diferenciáveis. No entanto, no restante desta seção realizaremos um estudo para o caso em que as funções f e g envolvidas em EP(f, C) são continuamente diferenciáveis. Sabemos que dado $\rho_\ell = \|z^\ell\|$ com $z^\ell \in \mathbb{R}^n$, o problema

$$\text{Encontre } y^\ell \in \Gamma_\ell \text{ tal que } \max_{y \in \Gamma_\ell} \tilde{\mathcal{L}}_j(y, z^\ell) \leq \tilde{\mathcal{L}}_j(y^\ell, z^\ell) + \varepsilon_\ell, \quad (4.25)$$

com $\Gamma_\ell = \overline{B(0, \rho_\ell + r)}$ e $r > 0$, definidos no algoritmo interno IALPMP, tem que ser resolvido a cada passo juntamente com a condição adicional $\tilde{\mathcal{L}}_j(y^\ell, z^\ell) \geq 0$. Esta última condição é satisfeita quando obtemos uma solução exata do problema de maximização (4.25), pois, se y^ℓ é solução deste problema então $\tilde{\mathcal{L}}_j(y^\ell, z^\ell) \geq \tilde{\mathcal{L}}_j(y, z^\ell)$ para $y \in \overline{B(0, \rho_\ell + r)}$ e como por definição $z^\ell \in \overline{B(0, \rho_\ell + r)}$ então de P1 concluímos que $\tilde{\mathcal{L}}_j(y^\ell, z^\ell) \geq \tilde{\mathcal{L}}_j(z^\ell, z^\ell) = 0$. Devido a que (4.25) é um problema de otimização, então temos que definir uma ferramenta que permita resolvê-lo, considerando as hipóteses P1-P3, para a função $\tilde{\mathcal{L}}_j$ que queremos maximizar. Embora existam Softwares como *Matlab* que possui ferramentas como FMINCON para resolver problemas de otimização com restrições de maneira eficiente, vamos explorar um pouco mais as hipóteses P1 - P3 no caso diferenciável. Como em (4.25) é exigida uma solução aproximada, então podemos considerar a aproximação linear da função objetivo ao redor do ponto z^ℓ , pois este ponto está fixado no problema. Usando P1 temos que $\tilde{\mathcal{L}}_j(z^\ell, z^\ell) = 0$ então pela diferenciabilidade da função $\tilde{\mathcal{L}}_j(\cdot, z^\ell)$ podemos esperar que para pontos próximos de z^ℓ a função objetivo cresça e assim obter um candidato que resolva (4.25) e cumpra $\tilde{\mathcal{L}}_j(y^\ell, z^\ell) \geq 0$. Em consequência, obteríamos uma reformulação bem mais simples do problema (4.25) dada por

$$\max_{y \in \tilde{C}_\ell} (\nabla_1 \tilde{\mathcal{L}}_j(z^\ell, z^\ell))^T (y - z^\ell), \quad (4.26)$$

em que $z^\ell \in \mathbb{R}^n$, $\nabla_1 \tilde{\mathcal{L}}_j(z^\ell, z^\ell)$ denota o gradiente da função $\tilde{\mathcal{L}}_j$ com respeito à primeira

variável avaliado no ponto z^ℓ e $\tilde{C}_\ell \subset \Gamma_\ell$ é um conjunto tal que z^ℓ está em seu interior. Por exemplo, se $r > 0$ é escolhido tal que a cada iteração $r \geq \|z^\ell\| + 1 = \rho_\ell + 1$, então $\overline{B(0, 2\rho_\ell + 1)} \subset \Gamma_\ell = \overline{B(0, \rho_\ell + r)}$ e portanto basta considerar $\tilde{C}_\ell = \overline{B(z^\ell, \omega_\ell)}$, em que $0 < \omega_\ell \leq \rho_\ell + 1 \leq r$. Com esta escolha, obtemos que o problema (4.26) tem uma solução analítica dada por

$$y^\ell = z^\ell + \frac{\omega_\ell}{\|\nabla_1 \tilde{\mathcal{L}}_j(z^\ell, z^\ell)\|} \nabla_1 \tilde{\mathcal{L}}_j(z^\ell, z^\ell). \quad (4.27)$$

Esta solução, depende da escolha do valor ω_ℓ em cada passo, isto é, ω_ℓ será escolhido tal que (4.25) e a condição $\tilde{\mathcal{L}}_j(y^\ell, z^\ell) \geq 0$ sejam satisfeitas. Assim, vamos definir uma versão do algoritmo IALPMP para o caso diferenciável.

Algoritmo 4.16 (IALPMG). *IALPM + Método Do Gradiente.*

Dados: $x^0 \in \mathbb{R}^n$, $\lambda^0 > 0$, $\epsilon_1 > 0$, $\epsilon_2 > 0$, $\alpha \in (0, 1)$ e uma sequência limitada $(\gamma_j)_{j \in \mathbb{N}} \subset (\theta, \bar{\gamma}]$ para algum $\bar{\gamma} > \theta$, com θ a constante de submonotonicidade. $j = 0$ e $aux = 1$.

ENQUANTO $aux = 1$

Sejam $\ell = 0$, $aux1 = 1$, $z^\ell = x^j$, $\rho_\ell = \|z^\ell\|$, $\omega_\ell \in (0, \rho_\ell + 1]$, e $\beta_\ell \in [\alpha, 1]$.

ENQUANTO $aux1 = 1$

CALCULE $y^\ell = z^\ell + \frac{\omega_\ell}{\|\nabla_1 \tilde{\mathcal{L}}_j(z^\ell, z^\ell)\|} \nabla_1 \tilde{\mathcal{L}}_j(z^\ell, z^\ell)$

SE $\tilde{\mathcal{L}}_j(y^\ell, z^\ell) \geq 0$

SE $\nabla(\tilde{\mathcal{L}}_j)_{y^\ell}(z^\ell) = \nabla \tilde{\mathcal{L}}_j(y^\ell, z^\ell) = 0$

$aux1 = 0$, $z^{\ell+1} = y^\ell$

SENÃO tome $\xi^\ell = \nabla(\tilde{\mathcal{L}}_j)_{y^\ell}(z^\ell)$, e atualize

$$z^{\ell+1} = z^\ell - \beta_\ell \left(\frac{\tilde{\mathcal{L}}_j(y^\ell, z^\ell)}{\|\xi^\ell\|^2} \right) \xi^\ell$$

SE $\beta_\ell \frac{\tilde{\mathcal{L}}_j(y^\ell, z^\ell)}{\|\xi^\ell\|} < \epsilon_2$ ou $\ell > 10000$

$aux1 = 0$

FIM

Atualize $\rho_{\ell+1} = \max\{\rho_\ell, \|z^{\ell+1}\|\}$, e

Escolha $\omega_{\ell+1} \in (0, \rho_{\ell+1} + 1]$.

Faça $\ell = \ell + 1$

FIM

SENÃO

$\omega_\ell = \omega_\ell/2$ (poderia ser $\omega_\ell = (\eta * \omega_\ell)$ com $\eta \in (0, 1)$)

FIM

FIM

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

$$\lambda_i^{j+1} = \max \left\{ 0, \lambda_i^j + \frac{g_i(z^{\ell+1})}{\gamma_j} \right\}, \quad \text{ou} \quad \left(\lambda_i^{j+1} = \lambda_i^j \exp \left(\frac{g_i(z^{\ell+1})}{\gamma_j} \right) \right)$$

SE $\|(x^j, \lambda^j) - (z^{\ell+1}, \lambda^{j+1})\| \leq \epsilon_1,$
 $aux = 0$
 SENÃO
 $x^{j+1} = z^{\ell+1}$
 FIM
 $j = j + 1.$
 FIM

Note que este algoritmo é um caso particular diferenciável do IALPMP, então sua convergência é garantida pelo Teorema 4.15. Além disso, observe que foi estabelecido de forma implícita que r é suficientemente grande e tal que $r \geq \rho_\ell + 1$ para cada $\ell \in \mathbb{N}$, deixando assim, como única dificuldade a escolha de ω_ℓ que usamos em cada iteração, pois dependendo da escolha deste valor podemos obter ou não uma solução aproximada do sub-problema (4.26) gerado no laço interno. Também salientamos que a principal vantagem do algoritmo IALPMG é que não vamos precisar de nenhuma ferramenta adicional, ou seja, uma vez definidas as sequências e parâmetros iniciais o algoritmo só precisa realizar operações básicas, isto é, adição, multiplicação e divisão, que diferentemente do algoritmo em sua forma pura, IALPMP, não precisaremos escolher outro algoritmo externo para resolver o problema de maximização aproximada, nem resolver um sistema não linear e com cálculos da Jacobiana como é o caso de IALPMN. Sendo assim, obtemos teoricamente uma vantagem sobre os dois primeiros algoritmos, especialmente para problemas de grande porte. Desta forma, esperamos que para uma boa escolha de ω_ℓ , em cada passo, possamos obter um algoritmo de Lagrangeano aumentado competitivo do ponto de vista computacional.

No próximo capítulo realizaremos testes numéricos para os Algoritmos IALPMN, IALPMG e IALPMP, utilizando a penalidade quadrática e exponencial. Os dois primeiros serão usados para resolver problemas de equilíbrio com funções diferenciáveis e o último para o caso não diferenciável. Os problemas considerados em nossos testes podem ser encontrados em [8, 24, 25, 26, 29].

Capítulo 5

Testes Numéricos

Nesta seção apresentaremos o desempenho numérico dos Algoritmos de Lagrangeano aumentado para resolver os problemas de equilíbrio, descritos no Capítulo 4. Para obter os resultados numéricos, utilizamos o software *Matlab* como linguagem de programação. Os resultados obtidos dos Algoritmos IALPMN, IALPMP e IALPMG foram disponibilizados em tabelas, seguidas de alguns comentários. Para diferenciar os algoritmos com penalidade exponencial dos algoritmos com penalidade quadrática vamos considerar o subíndice e nestas tabelas.

Em todos os algoritmos consideramos os valores $\epsilon_1 = 10^{-4}$ e $\epsilon_2 = 10^{-8}$ como tolerância mínima para aceitar uma solução, ou seja, no caso IALPMN, uma solução do Passo 2 (no laço interno) é aceita quando a desigualdade

$$\|\bar{G}^j(z^\ell)\| \leq \min \left\{ \epsilon_2, \frac{1}{2} \sigma \gamma_j (\min\{\|x^j - z^{\ell+1}\|, 1\})^2 \right\},$$

é satisfeita. Ao passo que nos algoritmos IALPMP e IALPMG uma solução é aceita quando a desigualdade

$$\|z^{\ell+1} - z^\ell\| = \beta_\ell \frac{\tilde{\mathcal{L}}_j(y^\ell, z^\ell)}{\|\xi^\ell\|} < \epsilon_2$$

se cumpre. Além disso, interrompemos o algoritmo se a solução não foi encontrada na precisão dada, após 1000 iterações para IALPMN e 10000 para IALPMG e IALPMP. Finalmente, dizemos que o algoritmo encontra uma solução para $EP(f, C)$ quando

$$\|(z^{\ell+1}, \lambda^{j+1}) - (x^j, \lambda^j)\| \leq \epsilon_1,$$

com $z^{\ell+1}$ obtido como solução no algoritmo interno.

A seguir vamos considerar seis problemas para testar e comparar o desempenho entre os algoritmos IALPMN, IALPMP e IALPMG. Os primeiros três problemas se referem a problemas de equilíbrio de Nash generalizado, em que na verdade, os dois primeiros são problemas de poluição de bacias hidrográficas [8, 24, 26, 29] e o terceiro é uma va-

riante do modelo de comunicação de internet (the internet switching model) introduzido por Kesselman em [21]. Todos estes exemplos podem ser reformulados como problemas de equilíbrio. Os problemas quatro e cinco são quadráticos tratados em [25, 29] e finalizamos com um problema não diferenciável analisado em [29].

O problema de equilíbrio de Nash generalizado é um jogo não cooperativo com N jogadores em que cada jogador i controla a sua variável de decisão dada por $x^i \in \mathbb{R}^{n_i}$. As variáveis de decisão para os demais jogadores são dadas por

$$x^{-i} = (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^N).$$

Logo,

$$x = (x^i, x^{-i}) = (x^1, \dots, x^{i-1}, x^i, x^{i+1}, \dots, x^N) \subset X,$$

denota as variáveis de todos os jogadores, sendo X um subconjunto convexo e fechado de \mathbb{R}^n , em que $n = \sum_{i=1}^N n_i$. Assim, o problema de equilíbrio de Nash generalizado consiste em resolver

$$\begin{aligned} \min_{x^i} \quad & \theta_i(x^i, x^{-i}) \\ \text{s.a.} \quad & x^i \in X_i(x^{-i}), \end{aligned} \tag{5.1}$$

em que $X_i(x^{-i}) \subset \mathbb{R}^{n_i}$ é o conjunto estratégia do jogador i e θ_i é a função perda do jogador i , para cada $i = 1 \dots, N$. Nos exemplos seguintes, consideramos o conjunto X como

$$X = \{x \in \mathbb{R}^n : g_k(x) \leq 0, \quad \forall k = 1, \dots, m\}, \tag{5.2}$$

com $g_k : \mathbb{R}^n \rightarrow \mathbb{R}$ funções convexas para cada $k = 1, \dots, m$. Sabemos de [8, 26] que o problema de equilíbrio de Nash generalizado pode ser escrito como um problema de desigualdade variacional (consequentemente como um problema de equilíbrio) que tem a função $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, dada por

$$f(x, y) = \sum_{i=1}^N \langle \nabla_{x^i} \theta_i(x), y^i - x^i \rangle \quad \text{e} \quad C = X, \tag{5.3}$$

com X dado como em (5.2).

Desta forma, consideremos o problema de poluição de bacias hidrográficas escrito como um problema de equilíbrio Nash generalizado, o qual pela análise feita acima pode ser reescrito como um problema de equilíbrio. Este problema é muito importante e tem sido muito estudado na literatura. Ver, por exemplo, [24, 25, 29].

Exemplo 5.1. Consideremos o problema de equilíbrio de Nash, para 3 jogadores $i =$

1, 2, 3. O jogo consiste em minimizar as funções objetivos dadas por

$$\theta_i(x) = c_i (x^i)^2 + d_1 x^i (x^1 + x^2 + x^3) - b_i x^i \quad i = 1, 2, 3, \quad (5.4)$$

em que $c_1 = c_3 = 0.01$, $c_2 = 0.05$, $d_1 = 0.01$, $b_1 = 2.90$, $b_2 = 2.88$ e $b_3 = 2.85$. As restrições g_i definidas em (5.2) são dadas por $g_1(x) = 3.25x^1 + 1.25x^2 + 4.125x^3 - 100$ e $g_2(x) = 2.291x^1 + 1.5625x^2 + 2.8125x^3 - 100$. Com estes dados obtemos que a bifunção de equilíbrio pode ser escrita na forma (5.3) como

$$f(x, y) = \langle Ax - b, y - x \rangle, \quad (5.5)$$

em que

$$A = \begin{pmatrix} 2(c_1 + d_1) & d_1 & d_1 \\ d_1 & 2(c_2 + d_1) & d_1 \\ d_1 & d_1 & 2(c_3 + d_1) \end{pmatrix} = \begin{pmatrix} 0.04 & 0.01 & 0.01 \\ 0.01 & 0.12 & 0.01 \\ 0.01 & 0.01 & 0.04 \end{pmatrix}$$

e

$$b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 2.90 \\ 2.88 \\ 2.85 \end{pmatrix}.$$

Neste exemplo, consideramos como pontos iniciais $x^0 = [5, 9, 3]^T$ e $\lambda^0 = [1, 1.5]^T$. O resto dos parâmetros foram escolhidos como, $t = 0.1$, $\sigma = 0.5$, $\beta_\ell = 1$, para todo ℓ , e a sequência $(\gamma_j)_{j \in \mathbb{N}}$ foi tomada como uma sequência constante em todos os Algoritmos, isto é, $\gamma_j = 1/250$ para IALPMN e $\gamma_j = 1/2$ para IALPMG e IALPMP. Salientamos que os algoritmos também resolvem o problema da poluição para outros valores de γ , mas aqui simplesmente selecionamos estes valores pois observamos que com eles, os algoritmos têm uma melhor performance. O valor de r foi escolhido como sendo $r = 1$ para IALPMP (note que com esta escolha o algoritmo é o definido em [18]) e $\omega_\ell = \|z^\ell\| + 1$ para IALPMG.

A Tabela 5.1 a seguir mostra os resultados obtidos para este problema. Na primeira coluna representamos o nome do método utilizado para resolver o problema, mais precisamente, usamos IALPMN, IALPMG e IALPMP para nos referir aos métodos de Lagrangeano aumentado com penalidade quadrática e IALPMNe, IALPMGe e IALPMPe para os métodos de Lagrangeano aumentado com penalidade exponencial. A segunda coluna mostra a solução obtida por cada algoritmo. As colunas três e quatro são, respectivamente, o número de iterações externas e internas empregado pelos algoritmos para encontrar a solução. A última coluna mostra o tempo, em segundos, empregado para resolver o problema considerado. Lembrando que uma das diferenças entre os Algoritmos IALPMG e IALPMP é dada pela forma em que o problema de maximização aproximado (4.25) é resolvido, ou seja, no Algoritmo IALPMG empregamos a aproximação linear descrita no capítulo anterior e para o Algoritmo IALPMP utilizamos a ferramenta FMINCON

disponível em *Matlab* para resolver os problemas de otimização gerados como subproblemas de IALPMP.

Algoritmo	Solução obtida pelo Algoritmo	Iterações Externas	Iterações Internas	Tempo
IALPMN	$[21.143, 16.028, 2.727]^T$	4	7	0.04
IALPMNe	$[21.144, 16.028, 2.727]^T$	153	216	0.04
IALPMG	$[21.103, 16.029, 2.758]^T$	380	3648	0.32
IALPMGe	$[21.143, 16.027, 2.727]^T$	160	31386	0.98
IALPMP	$[21.016, 16.030, 2.827]^T$	252	4585	41.81
IALPMPe	$[21.137, 16.028, 2.731]^T$	267	9090	61.38

Tabela 5.1: Resultados do Exemplo 5.1

Nos resultados obtidos na Tabela 5.1 podemos observar que os algoritmos IALPMN e IALPMNe foram os algoritmos com melhor desempenho. Apesar de IALPMG e IALPMGe ter realizado muitas iterações internas e externas para encontrar a solução, devemos salientar que este algoritmo ainda pode ser considerado uma boa alternativa para problemas diferenciáveis, pois, este algoritmo emprega apenas operações básicas (soma, multiplicação e divisão) enquanto nos outros dois algoritmos precisamos ou resolver um sistema linear que envolve a cálculo de Hessiana (no caso de IALPMN e IALPMNe) ou resolver um problema de maximização (caso de IALPMP e IALPMPe). Evidenciamos esta afirmação no seguinte exemplo.

Exemplo 5.2. Como em [25] vamos considerar uma generalização do Exemplo 5.1. Aumentamos o número de jogadores para N e o número de restrições para m . A função a ser minimizada é definida por

$$\theta_k(x) = u_k(x^k)^2 + \beta x^k \left(\sum_{i=1}^N x^i \right) - v_k x^k \quad \forall k = 1, \dots, N,$$

com, $\beta = 0.01$ e $u_k, v_k \in \mathbb{R}$ para cada $k = 1, \dots, N$. Consideramos m restrições, que representam as constantes de emissão de poluição, definidas pelo sistema $Ax \leq b$, onde $A \in \mathbb{R}^{m \times N}$ e $b \in \mathbb{R}^m$. O multiplicador de Lagrange inicial λ^0 , a matriz A , o vetor b e os escalares u_k e v_k foram gerados de maneira aleatória. Tomamos como ponto inicial $x^0 = [1, \dots, 1]^T \in \mathbb{R}^N$. O resto dos parâmetros foram escolhidos como, $t = 0.1$, $\sigma = 0.5$, $\beta_\ell = 0.75$ para todo ℓ e a sequência $(\gamma_j)_{j \in \mathbb{N}}$ foi tomada como uma sequência constante $\gamma_j = 1/250$ para os algoritmos IALPMN e IALPMG e para IALPMNe e IALPMGe usamos $\gamma_0 = 10000$ e $\gamma_k = 1/250$. Excluímos deste teste o Algoritmo IALPMP pelo fato de o mesmo não ter obtido um bom desempenho para 3 jogadores, e assim, para N jogadores acreditamos que seu desempenho seria ainda pior.

Apresentamos na Tabela 5.2 os resultados obtidos com os algoritmos IALPMN, IALPMNe, IALPMG e IALPMGe. Como no exemplo anterior para 3 jogadores, a primeira

coluna indica o algoritmo implementado. As colunas dois e três, representam respectivamente, o número de jogadores e restrições de cada problema. As colunas quatro e cinco fornecem informações sobre o número de iterações internas e o número de iterações externas, respectivamente. A última coluna indica o tempo em segundos empregado por cada algoritmo para resolver o problema.

Algoritmo	N	m	Iterações externas	Iterações Internas	Tempo
IALPMN	10	5	2	3	0.03
IALPMNe			3	5	0.01
IALPMG			2	126	0.05
IALPMGe			3	389	0.04
IALPMN	30	20	3	4	0.03
IALPMNe			4	6	0.02
IALPMG			3	548	0.11
IALPMGe			4	510	0.05
IALPMN	50	10	4	5	0.03
IALPMNe			6	8	0.03
IALPMG			4	1367	0.22
IALPMGe			4	441	0.09
IALPMN	100	10	5	6	0.06
IALPMNe			7	9	0.09
IALPMG			4	2190	0.34
IALPMGe			4	472	0.11
IALPMN	200	20	3	4	0.12
IALPMNe			6	8	0.11
IALPMG			3	821	0.35
IALPMGe			3	110	0.09
IALPMN	1000	20	4	6	21.2
IALPMNe			NR	NR	NR
IALPMG			6	1839	0.94
IALPMGe			NR	NR	NR
IALPMN	1000	100	6	24	85.3
IALPMNe			NR	NR	NR
IALPMG			16	161	0.18
IALPMGe			NR	NR	NR
IALPMN	2000	100	6	14	384.9
IALPMNe			NR	NR	NR
IALPMG			11	13543	14.1
IALPMGe			NR	NR	NR

Tabela 5.2: Resultados do Exemplo 5.2

Na Tabela 5.2 podemos observar que para os problemas em que o número de jogadores e restrições, é pequeno, o Algoritmo IALPMN tem desempenho melhor do que o algoritmo IALPMG. Além disso, para problemas com maior dimensão podemos

observar que ocorre o contrário, isto é, o desempenho do Algoritmo IALPMG é superior ao do IALPMN. Por outro lado, nos algoritmos IALPMGe e IALPMNe note que nos problemas de dimensão maior usamos as letras NR para indicar que problema não foi resolvido pelo método. Uma justificativa do porque isto acontece é que no caso em que usamos a penalidade exponencial notamos que os pontos encontrados nas iterações não podem ser muito inviáveis nem os valores da sequência $(\gamma_j)_{j \in \mathbb{N}}$ ser muito pequenos. Mais precisamente, é difícil controlar numericamente este parâmetro neste algoritmo, quando os valores dados são aleatórios. Um exemplo simples em que podemos notar a diferença entre essas penalidades é que supondo que em uma determinada iteração j , $\lambda_i^j = 1$ e x^j é um ponto inviável tal que o valor de uma restrição associado ao índice i é $g_i(x^j) = 5$. Se consideramos o valor $\gamma_j = 1/250$, note que no caso em que usamos a penalidade quadrática o valor $g(x^j)/\gamma_j = 1250$ e assim, o valor do termo $\max\{0, \lambda^j + g(x^j)/\gamma_j\} = 1251$. Porém, quando usamos a penalidade exponencial notamos que ao calcularmos em Matlab o termo $\lambda^j \exp(g(x^j)/\gamma_j) = +\infty$. Isto sugere que os algoritmos com penalidade quadrática têm mais estabilidade numérica do que os que usam penalidade exponencial. Uma saída para este problema numérico é utilizar um procedimento semelhante ao de otimização ordinária, ou seja, fazer um ajuste quadrático à penalidade exponencial. Fazemos mais comentários sobre isso na seção de trabalhos futuros.

Exemplo 5.3. O seguinte problema é apresentado em [9] como um problema de equilíbrio de Nash para N jogadores. A função objetivo para este problema é dada por:

$$\theta^i(x) = \frac{-x^i}{x^1 + \dots + x^N} \left(1 - \frac{x^1 + \dots + x^N}{B} \right), \quad (5.6)$$

para todo $i = 1, \dots, N$ e B constante. Para este problema, e por questões de comparação, consideramos os mesmos valores de [9], ou seja $B = 1$ e $N = 10$. Por (5.3) obtemos que a função que define o problema de equilíbrio é dado por:

$$f(x, y) = \sum_{i=1}^{10} \langle \nabla_{x^i} \theta_i(x), y^i - x^i \rangle,$$

em que $\nabla_{x^i} \theta_i(x) = 1 - \frac{1}{x^1 + \dots, x^{10}} + \frac{x^i}{(x^1 + \dots + x^{10})^2}$. As restrições para o primeiro jogador são dadas por $0.3 \leq x^1 \leq 0.5$ e para o restante dos jogadores consideramos que $x^1 + \dots + x^{10} \leq 1$ e $x^i \geq 0.01$ $i = 2, \dots, 10$. A sequência de parâmetros de penalidade foi escolhida de acordo com o método usado na implementação, isto é, para os algoritmos IALPMN e IALPMG, com penalidades quadráticas, tomamos $\gamma_0 = 1$ e $\gamma_j = 1/250$ para todo $j \in \mathbb{N}$ e para os métodos IALPMNe e IALPMGe, os quais usam a penalidade exponencial, tomamos $\gamma_0 = 1$ e $\gamma_j = 1/2$ para todo $j \in \mathbb{N}$. Além disso, neste exemplo, consideramos no algoritmo IALPMG, para o caso da penalidade quadrática, $\beta_\ell = 0.1$, para todo $\ell \in \mathbb{N}$ e para o caso da penalidade exponencial $\beta_\ell = 0.75$, para todo $\ell \in \mathbb{N}$.

Este problema foi testado em cada algoritmo de acordo com o estabelecido em [9], isto é, consideramos 3 pontos iniciais dados na Tabela 5.3 e em cada método o multiplicador inicial escolhido foi $\lambda^0 = [1, \dots, 1]^T \in \mathbb{R}^{12}$. A solução determinada por todos os métodos testados e para todos os pontos iniciais foi

$$x^* = [0.3000, 0.0694, 0.0694, 0.0694, 0.0694, 0.0694, 0.0694, 0.0694, 0.0694, 0.0694, 0.0694, 0.0694]^T$$

os demais resultados referentes ao número de iterações e tempo estão disponíveis da terceira à quinta colunas da Tabela 5.3.

Algoritmo	Ponto inicial	Iterações externas	Iterações internas	Tempo
IALPMN	$0.01 * [1, \dots, 1]^T$	3	12	0.05
IALPMNe	$0.01 * [1, \dots, 1]^T$	19	40	0.05
IALPMN	$0.1 * [1, \dots, 1]^T$	3	9	0.04
IALPMNe	$0.1 * [1, \dots, 1]^T$	19	37	0.05
IALPMN	$[1, \dots, 1]^T$	3	12	0.05
IALPMNe	$[1, \dots, 1]^T$	19	47	0.06
IALPMG	$0.01 * [1, \dots, 1]^T$	4	260	0.08
IALPMGe	$0.01 * [1, \dots, 1]^T$	31	940	0.27
IALPMG	$0.1 * [1, \dots, 1]^T$	4	304	0.09
IALPMGe	$0.1 * [1, \dots, 1]^T$	35	827	0.23
IALPMG	$[1, \dots, 1]^T$	5	373	0.09
IALPMGe	$[1, \dots, 1]^T$	36	1384	0.39

Tabela 5.3: Resultados do Exemplo 5.3.

Note que, de acordo com a escolha do ponto inicial, o número de iterações aumenta na medida que nos aproximamos do valor $[1, \dots, 1]^T$. Além disso, obtemos que o melhor desempenho, como no exemplo anterior, foi obtido no algoritmo IALPMN, independente da penalidade escolhida.

Exemplo 5.4. Neste exemplo consideramos dois problemas de equilíbrio quadráticos que denotaremos por $EP(f_i, C)$, $i=1,2$, os quais também foram descritos e implementados em [25, 29, 30]. As bifunções de equilíbrio que definem os problemas são dadas por $f_i(x, y) = \langle P_i x + Qy + q, y - x \rangle$, em que

$$P_1 = \begin{bmatrix} 3.1 & 2 & 0 & 0 & 0 \\ 2 & 3.6 & 0 & 0 & 0 \\ 0 & 0 & 3.5 & 2 & 0 \\ 0 & 0 & 2 & 3.3 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 3.1 & 2 & 0 & 0 & 0 \\ 2 & 3.6 & 0 & 0 & 0 \\ 0 & 0 & 3.5 & 2 & 0 \\ 0 & 0 & 2 & 3.3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$Q = \begin{bmatrix} 1.6 & 1 & 0 & 0 & 0 \\ 1 & 1.6 & 0 & 0 & 0 \\ 0 & 0 & 1.5 & 1 & 0 \\ 0 & 0 & 1 & 1.5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \text{ e } q = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \\ -1 \end{bmatrix}.$$

O conjunto de restrições para este problema é dado por

$$C = \{x \in \mathbb{R}^5 : \sum_{k=1}^5 x^k \geq -1, \quad -5 \leq x^k \leq 5, \quad k = 1, \dots, 5\}. \quad (5.7)$$

Os valores iniciais considerados para este exemplo foram $x^0 = [1, 3, 1, 1, 2]^T$ e $\lambda^0 = [1, \dots, 1]^T \in \mathbb{R}^{11}$. Note que cada restrição de caixa em (5.7) é transformada em duas restrições de desigualdade. Os parâmetros escolhidos para estes problema foram, $t = 0.1$, $\sigma = 0.5$, $\beta_\ell = 0.75$ para todo ℓ e, como no exemplo anterior, a sequência $(\gamma_j)_{j \in \mathbb{N}}$ foi tomada como a sequência $\gamma_0 = 1$ e $\gamma_j = 1/250$, para todo $j \geq 1$, para os três algoritmos testados. As Tabelas 5.4 e 5.5 exibem os resultados numéricos para cada problema, sendo que o significado das colunas são os mesmos das tabelas anteriores.

Algoritmo	Solução obtida pelo Algoritmo	Iterações externas	Iterações internas	Tempo
IALPMN	$[-0.725, 0.803, 0.720, -0.867, 0.250]^T$	2	6	0.03
IALPMNe	$[-0.725, 0.803, 0.720, -0.867, 0.250]^T$	3	6	0.01
IALPMG	$[-0.725, 0.803, 0.720, -0.867, 0.250]^T$	3	59	0.04
IALPMGe	$[-0.725, 0.803, 0.720, -0.867, 0.250]^T$	3	178	0.03
IALPMP	$[-0.725, 0.803, 0.720, -0.867, 0.250]^T$	3	47	1.84
IALPMPe	$[-0.725, 0.803, 0.720, -0.867, 0.250]^T$	3	63	1.95

Tabela 5.4: Resultados do Exemplo 5.4 associado com $EP(f_1; C)$

Algoritmo	Solução obtida pelo Algoritmo	Iterações externas	Iterações internas	Tempo
IALPMN	$[-0.725, 0.803, 0.720, -0.867, 0.200]^T$	2	6	0.03
IALPMNe	$[-0.725, 0.803, 0.720, -0.867, 0.200]^T$	3	6	0.01
IALPMG	$[-0.725, 0.803, 0.720, -0.867, 0.200]^T$	3	59	0.04
IALPMGe	$[-0.725, 0.803, 0.720, -0.867, 0.200]^T$	3	178	0.05
IALPMP	$[-0.725, 0.803, 0.720, -0.867, 0.200]^T$	3	58	1.87
IALPMPe	$[-0.725, 0.803, 0.720, -0.867, 0.200]^T$	3	63	1.99

Tabela 5.5: Resultados do Exemplo 5.4 associado com $EP(f_2; C)$

Como podemos observar nas colunas três, quatro e cinco das Tabelas 5.4 e 5.5, os

algoritmos IALPMN e IALPMG têm tempos similares, embora o número de iterações empregado por cada algoritmo sejam diferentes. Por outro lado, diferentemente do exemplo anterior, o Algoritmo IALPMP obteve um desempenho razoável. Porém, quando comparado com os tempos de IALPMN e IALPMG observamos que o mesmo continua sendo o mais lento.

Exemplo 5.5. Neste exemplo, consideramos uma generalização do Exemplo 5.4. As matrizes P_i , Q e o vetor q foram geradas de maneira aleatória, mantendo um formato similar de esparsidade. Os valores iniciais para cada problema são dados por $x^0 = [1, \dots, 1]^T \in \mathbb{R}^N$, $\lambda = [1, \dots, 1]^T \in \mathbb{R}^{2N+1}$, a sequência foi escolhida como $\gamma_0 = 1$ e para $j \geq 1$, $\gamma_j = 1/250$ para IALPMN, IALPMG e $\gamma_j = 1/2$ para IALPMNe e IALPMGe. Note que, o número de restrições, seguindo a definição do C em (5.7), é $m = 2N + 1$. Os resultados numéricos são apresentados na Tabela 5.6 a seguir e, assim como no Exemplo 5.2, excluímos dos testes o Algoritmo IALPMP.

A Tabela 5.6 mostra que o número de iterações executadas pelos algoritmos IALPMG e IALPMGe são maiores do que o número de iterações executadas pelos algoritmos IALPMN e IALPMNe, respectivamente. Porém, como no exemplo 2, vemos que em relação ao tempo de CPU não são muito diferentes.

Exemplo 5.6. Para finalizar, consideremos o seguinte exemplo não diferenciável, apresentado em [29]. Este problema é definido pela bifunção $f(x, y) = |y_1| - |x_1| + y_2^2 - x_2^2$ e o conjunto de restrições $C = \{x \in \mathbb{R}^2 : x_1 + x_2 \leq 1, -x_1 - x_2 \leq -1, x_1 \geq 0, x_2 \geq 0\}$. A solução deste problema é $x^* = (\frac{1}{2}, \frac{1}{2})$ e para verificar isto, aplicamos os Algoritmos IALPMP e IALPMPe. Neste caso, utilizamos a ferramenta FMINCON para resolver o subproblema de maximização aproximada. Note que o subdiferencial da bifunção f com respeito à segunda variável é dado por

$$\partial F_x(y) = \begin{cases} (1, 2y_2) & \text{se } y_1 > 0, \\ ([-1, 1], 2y_2) & \text{se } y_1 = 0, \\ (-1, 2y_2) & \text{se } y_1 < 0. \end{cases}$$

Lembrando que $F_x(y) = f(x, y)$. Para este exemplo consideramos $\lambda^0 = [1, 1, 1, 1]^T$ como valor do multiplicador inicial, $\gamma_j = 0.25$ para todo j , $\alpha = \beta = 0.75$ e $r = 1$. Na Tabela 5.7 exibimos os resultados para diferentes valores de x^0 .

Note que uma das vantagens do método de Lagrangeano aumentado com respeito ao algoritmo apresentado em [29], para resolver este problema, é que aqui não precisamos considerar que o valor inicial x^0 esteja no conjunto viável C o que é exemplificado pelos quatro valores iniciais diferentes. Os dois primeiros valores foram tomados de [29] e portanto observe que eles estão no conjunto viável. Os dois últimos foram escolhidos com a intenção de mostrar que o Algoritmo IALPMP obtém a solução desejada, mesmo que o valor inicial não esteja dentro do conjunto viável C .

Algoritmo	N	m	Iterações Externas	Iterações Internas	Tempo
IALPMN	10	21	2	6	0.04
IALPMNe			3	4	0.04
IALPMG			2	33	0.08
IALPMGe			2	58	0.16
IALPMN	20	41	2	6	0.04
IALPMNe			7	10	0.05
IALPMG			2	38	0.08
IALPMGe			4	91	0.29
IALPMN	50	101	2	6	0.05
IALPMNe			8	14	0.08
IALPMG			2	41	0.16
IALPMGe			6	74	0.42
IALPMN	100	201	2	6	0.07
IALPMNe			7	8	0.09
IALPMG			2	24	0.16
IALPMGe			2	69	0.82
IALPMN	150	301	2	6	0.09
IALPMNe			4	5	0.10
IALPMG			2	36	0.24
IALPMGe			2	58	0.54
IALPMN	300	601	3	12	0.44
IALPMNe			13	37	1.18
IALPMG			3	64	0.63
IALPMGe			2	114	1.22
IALPMN	500	1001	2	6	1.03
IALPMNe			7	8	1.94
IALPMG			2	80	2.04
IALPMGe			2	84	1.87
IALPMN	1000	2001	2	6	7.44
IALPMNe			3	5	10.2
IALPMG			9	42	1.09
IALPMGe			2	109	6.14
IALPMN	2000	4001	2	6	57.7
IALPMNe			5	6	114.9
IALPMG			9	150	39.6
IALPMGe			2	129	21.74
IALPMN	3000	6001	2	6	198.1
IALPMNe			4	6	327.1
IALPMG			6	591	387.1
IALPMGe			2	158	70.11
IALPMN	4000	8001	2	6	468.3
IALPMNe			4	6	779.8
IALPMG			6	553	602.6
IALPMGe			2	166	128.1

Tabela 5.6: Resultados do Exemplo 5.5

Algoritmo	Ponto inicial	Solução obtida pelo Algoritmo	Iterações externas	Iterações internas	Tempo
IALPMP	$[0, 1]^T$	$[0.4999, 0.5001]^T$	6	158	2.75
IALPMPe	$[0, 1]^T$	$[0.4999, 0.5001]^T$	7	152	2.59
IALPMP	$[1/3, 2/3]^T$	$[0.4999, 0.5001]^T$	6	149	2.62
IALPMPe	$[1/3, 2/3]^T$	$[0.4999, 0.5001]^T$	6	151	2.59
IALPMP	$[-1, -5]^T$	$[0.5001, 0.4999]^T$	7	191	3.8
IALPMPe	$[-1, -5]^T$	$[0.4999, 0.5001]^T$	8	193	3.35
IALPMP	$[10, 10]^T$	$[0.4999, 0.5001]^T$	7	203	3.85
IALPMPe	$[10, 10]^T$	$[0.4999, 0.5001]^T$	7	254	5.25

Tabela 5.7: Resultados do Exemplo 5.6.

Conclusão

Neste trabalho, pesquisamos métodos de Lagrangeano aumentado para problemas de equilíbrio com funções diferenciáveis e não diferenciáveis, definidas em um conjunto convexo, fechado e não vazio, formado por restrições de desigualdade. Os métodos apresentados se diferenciam, entre si, pela função de penalidade adotada, a qual é uma característica dos métodos baseados em penalização. Sob hipóteses razoáveis no contexto de problemas de equilíbrio, incluindo uma hipótese adicional introduzida em [4], mostramos a convergência de métodos exatos e inexatos, através da utilização de uma forte relação existente entre os métodos de ponto proximal e Lagrangeano aumentado.

Desta forma, as principais contribuições da tese foram:

(a) O desenvolvimento de três algoritmos implementáveis para resolver os subproblemas gerados pelos métodos de Lagrangeano aumentado, aplicados ao problema de equilíbrio e desenvolvidos em [16]. O primeiro destes é baseado no método de Newton para resolução de um sistema não linear e não diferenciável. Os outros dois são baseados em projeção do subgradiente os quais são oriundos do conceito de viabilidade convexa e foram desenvolvidos, na versão aqui apresentada, em [18]. Sendo que primeiro não exige hipóteses de diferenciabilidade para as funções envolvidas o que nos permite resolver problemas de equilíbrio não diferenciáveis. O segundo exige diferenciabilidade mas com uma vantagem de só executar operações simples, basicamente adições e multiplicações, devido a isso, se comportou melhor para problemas de dimensões mais altas.

(b) Estendemos os métodos de Lagrangeano aumentado apresentados em [16], no sentido de usar penalidades mais gerais do que a quadrática. É bem conhecido na literatura que a penalidade quadrática proposta por Hestenes, Powell e Rockafellar [12, 27, 28], para o problema de otimização não linear com restrições de desigualdade, e em [16] para o problema de equilíbrio, só possuem derivadas de primeira ordem. Aqui, nesta tese, desenvolvemos as penalidades de maneira que fossem independentes das funções que as definem. Desta forma, quaisquer funções que satisfaçam as hipóteses consideradas podem ser utilizadas com essa nova metodologia que apresentamos. Para efeito de desenvolvimento da metodologia utilizamos a penalidade exponencial, a qual tem a vantagem em relação à quadrática, por possuir derivadas de todas as ordens. Com poucas adaptações é possível fazer o desenvolvimento para o caso da penalidade barreira logarítima.

(c) Implementação, em Matlab, de todos os algoritmos desenvolvidos. Analisamos numericamente o comportamento destes aplicando-os a problemas conhecidos na

literatura, como é o caso do problema de poluição da bacia de um rio, que é clássico e apresentado em vários trabalhos da literatura, inclusive podendo ser tratado como um problema de equilíbrio de Nash generalizado. Além deste, aplicamos a outros problemas de equilíbrio diferenciáveis e um caso não diferenciável, para exemplificar o caso do algoritmo que não exigia diferenciabilidade das funções envolvidas. Observamos que para problemas menores o algoritmo baseado no método de Newton foi melhor. Já para problemas maiores o método baseado em viabilidade convexa teve um melhor desempenho. O que ficou claro, e, também, não deveria ser uma novidade, pois o mesmo ocorre com o caso de otimização ordinária, a penalidade exponencial sofre do problema de mal condicionamento, ou seja, cresce muito rápido para pontos inviáveis. Para remediar isso, imitando o caso de otimização ordinária, é necessário um ajuste de uma função quadrática. Esse assunto ficou como proposta de desenvolvimentos futuros.

Por fim, destacamos que um artigo está sendo submetido, em parceria com o pesquisador Mostafa Nasri do Canadá, referente aos itens (a) e (c) acima. Outro artigo está em fase de redação e é referente ao item (b) acima.

Trabalhos futuros

Algumas possibilidades de continuação de pesquisa desta tese são:

- Um ajuste quadrático na penalidade exponencial, semelhante ao que é realizado em otimização ordinária em vários trabalhos da literatura, como por exemplo em [22, 23]. É comum quando se usa penalidade exponencial ou logarítmica fazer este ajuste. Isto evita o crescimento muito rápido da função exponencial quando os pontos gerados pelo algoritmo estão inviáveis. Para pontos viáveis estas penalidades funcionam muito bem pois o efeito da penalização, neste caso, é pequeno.
- Desenvolver outras penalidades diferentes da exponencial, como por exemplo a penalidade barreira logarítmica, que é utilizada em vários trabalhos da literatura para o caso de otimização ordinária. Desta forma, haveria uma gama de possibilidades combinando estas, tais como: exponencial-quadrática; logarítmica-quadrática; exponencial-logarítmica, dentre outras.
- Um melhor estudo sobre o parâmetro de penalidade γ , tanto do ponto de vista teórico como numérico. Os métodos se mostraram muito sensíveis a pequenas variações neste parâmetro. A convergência dos algoritmos analisados depende que este parâmetro seja limitado. Portanto, uma boa escolha poderá deixar os métodos mais rápidos.
- Por fim, aprofundar nos experimentos numéricos. São poucos os exemplos encontrados na literatura sobre problemas de equilíbrio. Muitos exemplos são oriundos de desigualdade variacional ou teoria de jogos. Métodos de Lagrangeano aumentado para otimização ordinária têm se mostrado muito eficientes para problemas de grande porte. O mesmo deve ocorrer no caso de problemas de equilíbrio.

Referências Bibliográficas

- [1] G. Bigi, M. Castellani, M. Pappalardo, and M. Passacantando. Existence and solution methods for equilibria. *European Journal of Operational Research*, 227(1):1 – 11, 2013.
- [2] E. G. Birgin and J. M. Martínez. *Practical Augmented Lagrangian Methods for Constrained Optimization*. SIAM, 2014.
- [3] E. Blum and W. Oettli. From optimization and variational inequalities to equilibrium problems. *The Mathematics Student*, 63:123–145, 1994.
- [4] R. Burachik and G. Kassay. On a generalized proximal point method for solving equilibrium problems in Banach spaces. *Nonlinear Analysis: Theory, Methods & Applications*, 75(18):6456 – 6464, 2012.
- [5] D. Butnariu and A. N. Iusem. *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization*, volume 40. 2008.
- [6] A. R. De Pierro and A. N. Iusem. A relaxed version of bregman’s method for convex programming. *Journal of Optimization Theory and Applications*, 51(3):421–440, 1986.
- [7] J. Eckstein. Nonlinear proximal point algorithms using bregman functions, with applications to convex programming. *Mathematics of Operations Research*, 18(1):202–226, 1993.
- [8] F. Facchinei and C. Kanzow. Generalized Nash equilibrium problems. *4OR*, 5(3):173–210, 2007.
- [9] F. Facchinei and C. Kanzow. Penalty methods for the solution of generalized Nash equilibrium problems. *SIAM Journal on Optimization*, 20(5):2228–2253, 2010.
- [10] K. Fan. A generalization of Tychonoff’s fixed point theorem. *Mathematische Annalen*, 142(3):305–310, 1961.
- [11] K. Fan. A minimax inequality and applications. In: *Shisha, O. (ed.) Inequality III*, 1972.

- [12] M. R. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and applications*, 4:303–320, 1969.
- [13] A. N. Iusem and R. Burachik. *Set-Valued Mappings and Enlargements of Monotone Operators*, volume 8. 2008.
- [14] A. N. Iusem, G. Kassay, and W. Sosa. On certain conditions for the existence of solutions of equilibrium problems. *Mathematical Programming*, 116(1):259–273, 2009.
- [15] A. N. Iusem and M. Nasri. Inexact proximal point methods for equilibrium problems in Banach spaces. *Numerical Functional Analysis and Optimization*, 28(11-12):1279–1308, 2007.
- [16] A. N. Iusem and M. Nasri. Augmented Lagrangian methods for equilibrium problems. *RAIRO Oper. Res.*, 44(1):5–26, 2010.
- [17] A. N. Iusem and R. G. Otero. Inexact versions of proximal point and augmented Lagrangian algorithms in Banach spaces. *Numerical Functional Analysis and Optimization*, 22(5-6):609–640, 2001.
- [18] A. N. Iusem and W. Sosa. Iterative algorithms for equilibrium problems. *Optimization*, 52(3):301–316, 2003.
- [19] A. N. Iusem and W. Sosa. New existence results for equilibrium problems. *Nonlinear Analysis: Theory, Methods and Applications*, 52(2):621–635, 2003.
- [20] A. N. Iusem and W. Sosa. On the proximal point method for equilibrium problems in Hilbert spaces. *Optimization*, 59(8):1259–1274, 2010.
- [21] A. Kesselman, S. Leonardi, and V. Bonifaci. *Game-Theoretic Analysis of Internet Switching with Selfish Users*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.
- [22] J. M. Martínez. Box-quacan and the implementation of augmented Lagrangian algorithms for minimization with inequality constraints. *Computational and Applied Mathematics*, 19:31–56, 01 2000.
- [23] L. C. Matioli. *Uma nova metodologia para construção de funções de penalização para algoritmos de Lagrangeano aumentado*. PhD thesis, Universidade Federal de Santa Catarina- UFSC, Santa Catarina-Brasil, 2001.
- [24] L. C. Matioli, W. Sosa, and J. Y. Yuan. A numerical algorithm for finding solutions of a generalized Nash equilibrium problem. *Computational Optimization and Applications*, 52(1):281–292, 2012.

- [25] M. Nasri, L. C. Matioli, E. M. F. Ferreira, and A. Silveira. Implementation of augmented Lagrangian methods for equilibrium problems. *Journal of Optimization Theory and Applications*, 168(3):971–991, 2015.
- [26] M. Nasri and W. Sosa. Equilibrium problems and generalized Nash games. *Optimization*, 60(8-9):1161–1170, 2011.
- [27] M. J. D. Powell. A method for nonlinear constraints in minimization problems. In R. Fletcher, editor, *Optimization*, pages 283–298. Academic Press, New York, 1969.
- [28] R. T. Rockafellar. The multiplier method of Hestenes and Powell applied to convex programming. *Journal of Optimization Theory and applications*, 12:555–562, 1973.
- [29] P. Santos and S. Scheimberg. An inexact subgradient algorithm for equilibrium problems. *Computational & Applied Mathematics*, 30(1):91–107, 2011.
- [30] D. Q. Tran, L. D. Muu, and V. H. Nguyen. Extragradient algorithms extended to equilibrium problems. *Optimization*, 57(6):749–776.