

UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM  
INSTITUTO DE CIÊNCIAS EXATAS - ICE  
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

PROBLEMAS DE EQUILÍBRIO E MÉTODOS DE  
LAGRANGEANO AUMENTADO PARA PROBLEMAS DE  
DESIGUALDADES VARIACIONAIS

Jefferson Castro Silva

MANAUS - 2012

UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM  
INSTITUTO DE CIÊNCIAS EXATAS - ICE  
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Jefferson Castro Silva

PROBLEMAS DE EQUILÍBRIO E MÉTODOS DE LAGRANGEANO  
AUMENTADO PARA PROBLEMAS DE DESIGUALDADES  
VARIACIONAIS

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Matemática, na área de concentração em Otimização.

Orientador(a): Profa. Flávia Morgana de Oliveira Jacinto

MANAUS - 2012

Jefferson Castro Silva

PROBLEMAS DE EQUILÍBRIO E MÉTODOS DE LAGRANGEANO  
AUMENTADO PARA PROBLEMAS DE DESIGUALDADES  
VARIACIONAIS

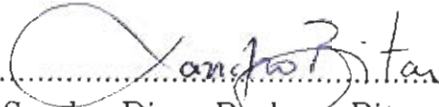
Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Matemática, na área de concentração em Otimização.

Manaus, 21 de Dezembro de 2012.

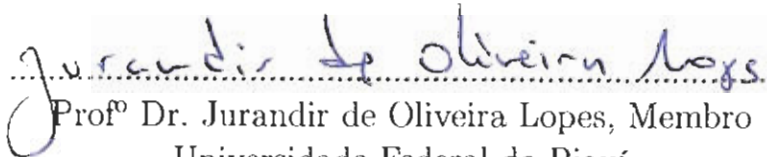
BANCA EXAMINADORA



.....  
Profa Dra. Flávia Morgana de Oliveira Jacinto, Presidente  
Universidade Federal do Amazonas



.....  
Profº Dr. Sandro Dimy Barbosa Bitar, Membro  
Universidade Federal do Amazonas



.....  
Profº Dr. Jurandir de Oliveira Lopes, Membro  
Universidade Federal do Piauí

# AGRADECIMENTOS

- A Deus por todas as graças concebidas em minha vida.
- À minha família pelo apoio e compreensão.
- Aos membros da banca examinadora pelas sugestões dadas à melhoria desta dissertação.
- À professora Flávia pela orientação e pelo estímulo constante.
- Aos professores do Programa de Pós-Graduação em Matemática (PPGM) que contribuíram a minha formação acadêmica.
- Aos colegas do Mestrado pelo apoio constante.
- Ao Fundo de Amparo a Pesquisa na Amazônia- FAPEAM pelo apoio financeiro.

# RESUMO

## PROBLEMAS DE EQUILÍBRIO E MÉTODOS DE LAGRANGEANO AUMENTADO PARA PROBLEMAS DE DESIGUALDADES VARIACIONAIS

Neste trabalho apresentamos métodos de Lagrangeano aumentado que são usados na resolução de Problemas de Desigualdade Variacional (PDV) em subconjuntos do  $\mathbb{R}^n$  definidos por restrições convexas. Analisamos a convergência desses métodos para soluções do problema (PDV), apresentando as condições necessárias a fim de garantir que os algoritmos gerados converjam. Concluimos que os principais algoritmos aqui apresentados, a saber o Método Extragradiante Lagrangeano Aumentado Inexato e o Método Extragradiante Ponto Proximal Inexato, geram a mesma sequência quando aplicados a problemas de desigualdade variacional distintos, porém interligados. Assim, garantida a existência de solução para um dos problemas e a respectiva convergência do algoritmo que lhe é empregado, concluimos que o outro algoritmo também converge fornecendo solução do respectivo problema ao qual é aplicado.

**Palavras-chave:** problema de equilíbrio, desigualdade variacional, métodos de lagrangeano aumentado, lagrangeano aumentado inexato.

# ABSTRACT

## EQUILIBRIUM PROBLEMS AND AUGMENTED LAGRANGIAN METHODS FOR VARIATIONAL INEQUALITIES PROBLEMS

In this paper we present augmented Lagrangian methods that are used in solving Variational Inequalities Problems (VIP) into subsets of  $\mathbb{R}^n$  defined by convex constraints. We analyze the convergence of these methods for (VIP) problem solutions, presenting the conditions necessary to ensure that the generated algorithms converge. We conclude that the main algorithms presented here, namely Inexact Augmented Lagrangian Extragradient Method and Inexact Proximal Point Extragradient Method, generate the same sequence when applied to distinct inequality variational problems but interconnected. Thus assured that solution to one of the problems and its convergence of the algorithm which is used, we conclude that another algorithm also converges providing solution of the respective problem that it is applied.

**Keywords:** equilibrium problem, variational inequality, augmented lagrangian methods, inexact augmented lagrangian.

# Sumário

<b>Introdução</b>	<b>2</b>
<b>1 Preliminares</b>	<b>1</b>
1.1 Elementos de Análise . . . . .	1
1.2 Operadores Monótonos . . . . .	7
1.3 Funções e Distâncias de Bregman . . . . .	8
1.4 O Problema de Desigualdade Variacional (PDV) . . . . .	13
1.5 O Método do Ponto Proximal . . . . .	16
1.6 Métodos de Lagrangeano Aumentado . . . . .	19
1.7 Condições de otimalidade de Karush-Kuhn-Tucker(KKT) para o problema de equilíbrio . . . . .	22
<b>2 Análise de Convergência do Algoritmo IPPEM</b>	<b>27</b>
2.1 Funções Regularizadoras e o Método do Ponto Proximal para $PE(f, \mathbb{K})$ . . . . .	27
2.2 O Método Extragradiante Ponto Proximal Inexato . . . . .	29
2.3 Aplicações dos resultados de convergência de IPPEM . . . . .	36
<b>3 Análise de Convergência do Algoritmo IALEM</b>	<b>39</b>
3.1 Relações entre as soluções dos problemas $PE(f, \mathbb{K})$ e $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . . . . .	39
3.2 O Método Extragradiante Lagrangeano Aumentado Inexato . . . . .	43
3.3 Variantes do algoritmo IALEM . . . . .	53
<b>Considerações Finais</b>	<b>57</b>
<b>Referências Bibliográficas</b>	<b>59</b>

# Notações

$\mathbb{R}_+$	Conjunto dos números reais não-negativos.
$\langle \cdot, \cdot \rangle$	Produto interno usual de $\mathbb{R}^n$ .
$\  \cdot \ $	Norma euclidiana de $\mathbb{R}^n$ .
$\partial f(x)$	Subdiferencial da função $f$ no ponto $x$ .
$\mathcal{N}_D(x)$	Cone das direções normais no ponto $x$ .
$\mathcal{P}(\mathbb{R}^n)$	Conjunto das partes de $\mathbb{R}^n$ .
$D_g$	Distância de Bregman com respeito a função $g$ .
$\nu_g(x, t)$	Módulo de total convexidade da função $g$ .
$\text{PDV}(F, \mathbb{K})$	Problema de Desigualdade Variacional definido por um operador $F$ sobre um subconjunto $\mathbb{K}$ de $\mathbb{R}^n$ .
$S(F, \mathbb{K})$	Conjunto de soluções do problema $\text{PDV}(F, \mathbb{K})$ .
$\text{PE}(f, \mathbb{K})$	Problema de Equilíbrio definido por $f$ em um conjunto $\mathbb{K}$ .
$S(f, \mathbb{K})$	Conjunto de soluções do problema $\text{PE}(f, \mathbb{K})$ .
(MPP)	Método do Ponto Proximal.
(MLA)	Método do Lagrangeano Aumentado.
$\bar{L}_\alpha$	Função Lagrangeana Aumentada associada ao problema de minimização.
$\mathcal{L}$	Lagrangeano Exato ou Lagrangeano associado ao problema $\text{PE}(f, \mathbb{K})$ .
$I(x)$	Conjunto de índices das restrições ativas no ponto $x$ .
(CS)	Condição de Slater para as restrições do problema de minimização.
$\partial_{x_i} f(x_1, \dots, x_i, \dots, x_n)$	Subdiferencial da função $f$ na variável $x_i$ no ponto $(x_1, \dots, x_i, \dots, x_n)$ .
IPPEM	Método Extragradiante Ponto Proximal Inexato.
IALEM	Método Extragradiante Lagrangeano Aumentado Proximal.
$\tilde{f}_k^e$	Forma Inexata da função regularizadora de $f$ para o problema $\text{PE}(f, \mathbb{K})$ .
$\hat{\mathcal{L}}_k^e$	Forma Inexata da função regularizadora de $\mathcal{L}$ para o problema $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ .
$\text{PE}(\hat{\mathcal{L}}_k^e, \mathbb{R}^n \times \mathbb{R}_+^m)$	Forma regularizada inexata do problema $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ na iteração $k$ .
$\tilde{\mathcal{L}}_k^e$	Função Lagrangeana Aumentada Proximal Inexata.
$\text{int } \mathbb{K}$	Interior relativo do conjunto $\mathbb{K}$ .
EALM	Método do Lagrangeano Aumentando Exato.
IALEM	Método Extragradiante Lagrangeano Aumentado Linearizado Inexato.
$\bar{\mathcal{L}}_k^e$	Função Lagrangeana Linearizada associada ao problema $\text{PE}(f, \mathbb{K})$ .



# Introdução

Um dos problemas de interesse em Otimização é a resolução de problemas de equilíbrio em espaços de dimensão finita. Esta classe de problemas é a forma generalizada do clássico problema de minimização. O interesse deste trabalho é resolver um certo tipo de problema de equilíbrio, o qual é chamado Problema de Desigualdade Variacional. Estes problemas são, em geral, determinados por um operador contínuo  $F$  e suas soluções (quando existem) atendem a uma condição de viabilidade determinada por um subconjunto  $\mathbb{K}$  de  $\mathbb{R}^n$ .

Neste trabalho pedimos que o operador  $F$  seja monótono e que o conjunto viável  $\mathbb{K}$  (que define o problema) seja não-vazio, fechado e convexo. Apresentamos métodos que buscam soluções ao problema de desigualdade variacional, através da resolução de subproblemas, os quais também são problemas de desigualdade variacional, que envolvem funções Lagrangeanas apropriadas obtidas diretamente do problema principal. Estes métodos são conhecidos na literatura como métodos de Lagrangeano aumentado, em particular, tais métodos combinam idéias que advêm de métodos clássicos como os de penalização e de dualidade.

O objetivo deste trabalho é estudar a convergência de um método de Lagrangeano aumentado, apresentado por Iusem em [5], aplicado a um problema de desigualdade variacional restrito que é nosso problema principal. Este algoritmo é chamado Método Lagrangeano Extragradiante Aumentado Inexato (IALEM), o qual tem por objetivo resolver o problema de desigualdade variacional restrito através de uma sequência de subproblemas de desigualdade variacional irrestritos, onde cada subproblema é definido por uma função Lagrangeana, obtida através da forma regularizada inexata da Lagrangeana aumentada do problema principal.

A análise de convergência do algoritmo IALEM é feita baseada na análise de convergência do Método Extragradiante Ponto Proximal Inexato (IPPEM) (ver [5]). Observamos que a monotonicidade do operador  $F$  é essencial na análise de convergência do algoritmo IPPEM. De fato, a monotonicidade estabelece uma ligação entre os algoritmos IALEM e IPPEM, quando mostramos que, aplicando IPPEM a um problema de desigualdade variacional definido pela função Lagrangeana, a sequência gerada coincide com a sequência gerada por IALEM aplicada ao problema principal. A partir da monotonicidade de  $F$ , garantimos a convergência de IPPEM a uma solução do problema e, como as sequências de ambos coincidem, IALEM também converge a esta solução.

A fim de obter a conclusão apresentada acima, este trabalho está organizado como segue. No capítulo 1 começamos com a apresentação de elementos necessários a análise de conver-

gência de IPPEM, e conseqüentemente de IALEM. Apresentamos os problemas de equilíbrio, de desigualdade variacional e os elementos analíticos clássicos presentes nestes problemas, além das suas respectivas condições de otimalidade de Karush-Kuhn-Tucker.

No capítulo 2 fazemos uma análise de convergência de IPPEM num contexto mais geral (como em [6]), através da distância de Bregman definida em  $\mathbb{R}^n$ .

No capítulo 3 é feita a análise de convergência propriamente dita de IALEM, onde são apresentados os principais resultados deste trabalho, a saber estes garantem que a sequência gerada por IALEM aplicado ao problema  $PE(f, \mathbb{K})$ , e a sequência gerada por IPPEM aplicado ao problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , são as mesmas. Além disso, a convergência a uma solução fornecida por um destes algoritmos garante a convergência do outro a esta solução.

Ressaltamos que faremos apenas a análise de convergência destes algoritmos, questões como a ordem de convergência e implementação não serão abordadas neste trabalho.

# Capítulo 1

## Preliminares

Neste capítulo apresentamos conceitos imprescindíveis aos algoritmos estudados no decorrer deste trabalho. Começamos com elementos clássicos da Análise Convexa e de Análise no espaço  $\mathbb{R}^n$ . Em seguida, descrevemos os conceitos de operadores e funções de Bregman em  $\mathbb{R}^n$  e explicitamos a formulação geral dos problemas de desigualdade variacional e equilíbrio neste mesmo espaço. A fim de estudarmos os métodos para problemas de desigualdade variacional no escopo de equilíbrio, mostramos os métodos de ponto proximal e Lagrangeano aumentado usados para resolver problemas de minimização convexa em subconjuntos convexos de  $\mathbb{R}^n$ . Finalizamos, este capítulo, estudando as condições de Karush-Kuhn-Tucker para o problema de minimização e apresentamos o formato geral destas condições para o problema de equilíbrio.

### 1.1 Elementos de Análise

Nesta seção apresentamos os elementos necessários a compreensão dos conceitos a serem abordados nas seções seguintes. A primeira definição a ser apresentada é uma extensão ao conceito de continuidade de uma função.

**Definição 1.1.** ([8], p.13) *Uma função  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  é dita semicontínua inferiormente no ponto  $\bar{x} \in D$ , quando para qualquer sequência  $\{x^k\} \subset D$  tal que  $\lim_{k \rightarrow \infty} x^k = \bar{x}$ , tem-se*

$$\liminf_{k \rightarrow \infty} f(x^k) \geq f(\bar{x}).$$

*Nas mesmas condições,  $f$  é dita semicontínua superiormente em  $\bar{x} \in D$  quando*

$$\limsup_{k \rightarrow \infty} f(x^k) \leq f(\bar{x}).$$

*A função  $f$  é semicontínua inferiormente(superiormente) no conjunto  $D$ , quando ela é semicontínua inferiormente(superiormente) em todos os pontos de  $D$ .*

Convém observar que se  $f$  é uma função contínua em  $D \subset \mathbb{R}^n$ , então  $f$  é semicontínua

superiormente e inferiormente em  $D$ . Apenas a condição de ser semicontínua inferior ou superior não garante que  $f$  seja contínua, de fato, isto ocorre no seguinte exemplo.

**Exemplo 1.1.** Considere a função  $f: \mathbb{R} \rightarrow \mathbb{R}$  definida por:

$$f(x) = \begin{cases} 1 & \text{se } x \geq 0; \\ -1 & \text{se } x < 0 \end{cases}$$

Vê-se que  $f$  é semicontínua superiormente em  $x = 0$ , porém  $f$  não é semicontínua inferiormente neste ponto. Assim,  $f$  não é contínua em  $x = 0$ .

Assim como o conceito de continuidade pode ser estendido, também podemos estender a noção de derivada de uma função. A esta extensão de derivada denominamos de subdiferencial. Aqui, é interessante identificar o subdiferencial de uma classe específica de função, a saber, das funções convexas. Antes de apresentarmos tal noção, precisamos das definições de conjunto convexo em  $\mathbb{R}^n$  e de função convexa. Precisamente, temos:

**Definição 1.2.** ([8], p.67) Um conjunto  $D \subset \mathbb{R}^n$  é chamado conjunto convexo se para quaisquer  $(x, y) \in D \times D$  e  $\alpha \in [0, 1]$ , tem-se que  $\alpha x + (1 - \alpha)y \in D$ .

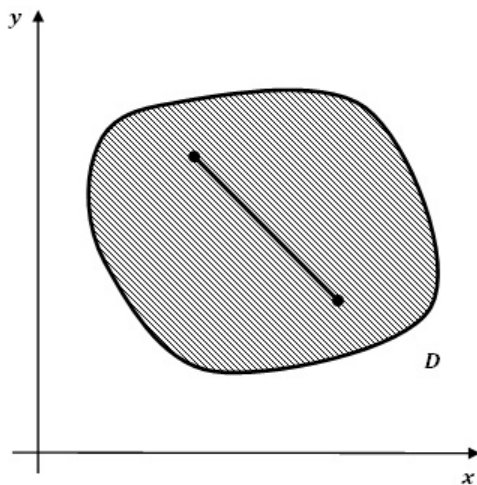


Figura 1.1:  $D$  é um conjunto convexo

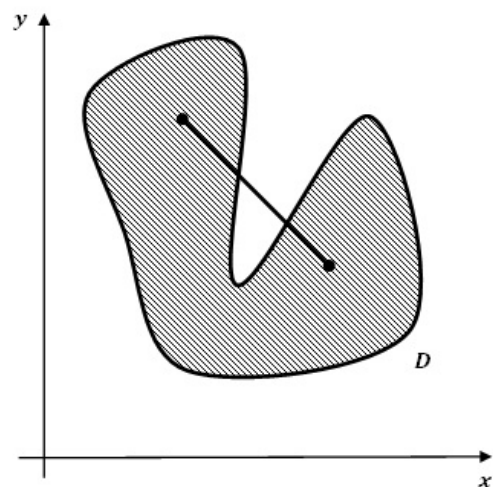


Figura 1.2:  $D$  não é um conjunto convexo

**Definição 1.3.** ([8], p.74) Seja  $D \subset \mathbb{R}^n$  um conjunto convexo. Dizemos que uma função  $f: D \rightarrow \mathbb{R}$  é convexa em  $D$  quando

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

para todo  $(x, y) \in D \times D$  e  $\alpha \in [0, 1]$ . A função  $f$  é dita estritamente convexa quando a desigualdade acima é estrita para todo  $x \neq y$  e  $\alpha \in (0, 1)$ .

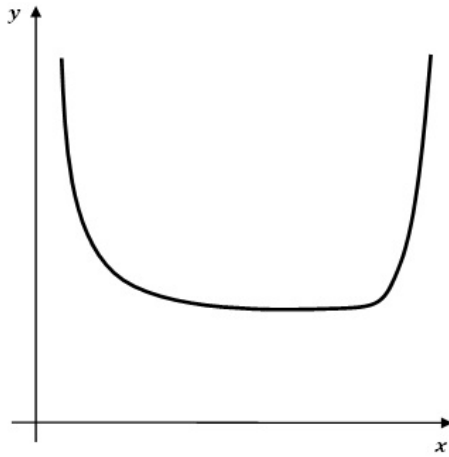


Figura 1.3: Função convexa

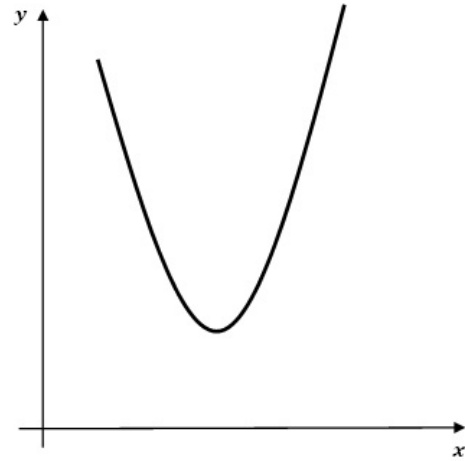


Figura 1.4: Função estritamente convexa

**Exemplo 1.2.** Todo intervalo  $I$  da reta real  $\mathbb{R}$  é um conjunto convexo. O produto cartesiano de dois intervalos  $I$  e  $J$  da reta é um conjunto convexo. A bola  $B = \{x \in \mathbb{R}^n; \|x\| \leq 1\}$  também é um conjunto convexo.

**Exemplo 1.3.** A função  $f : [-1, 1] \rightarrow \mathbb{R}$  dada por  $f(x) = |x|$  é convexa. Com efeito, para todo  $(x, y) \in [-1, 1] \times [-1, 1]$  e  $\alpha \in [0, 1]$ , tem-se

$$f(\alpha x + (1 - \alpha)y) = |\alpha x + (1 - \alpha)y| \leq |\alpha||x| + |1 - \alpha||y| = \alpha f(x) + (1 - \alpha)f(y).$$

Agora podemos definir o subdiferencial de uma função convexa como segue:

**Definição 1.4.** ([8], p.171) Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa. Dizemos que  $y \in \mathbb{R}^n$  é um subgradiente de  $f$  no ponto  $x \in \mathbb{R}^n$  se

$$f(z) \geq f(x) + \langle y, z - x \rangle, \quad \forall z \in \mathbb{R}^n.$$

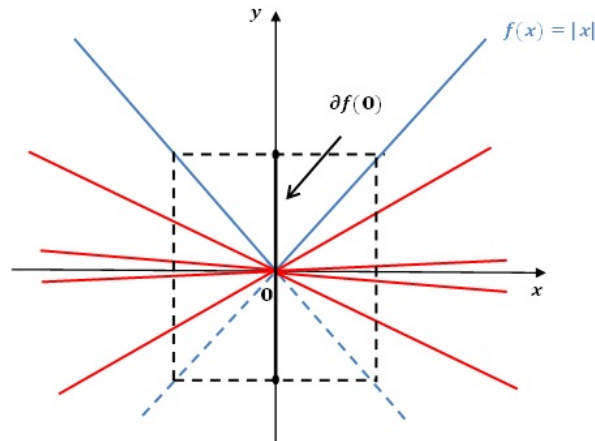
O conjunto de todos os subgradientes de  $f$  em  $x$  se chama o subdiferencial de  $f$  em  $x$ , e o denotamos por  $\partial f(x)$ .

**Exemplo 1.4.** Seja a função  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $f(x) = |x|$ . O subdiferencial de  $f$  no ponto  $x = 0$  é o conjunto de todos os valores reais  $y$  tais que  $f(z) \geq f(0) + y(z - 0)$ , isto é,  $|z| \geq yz$ . Daí, temos três casos a avaliar:

- (i) se  $z = 0$ , então a desigualdade é satisfeita para qualquer valor  $y \in \mathbb{R}$ ;
- (ii) se  $z > 0$ , então  $1 \geq y$ ;
- (iii) se  $z < 0$ , então  $-1 \leq y$ .

Assim, concluímos que  $\partial f(0) = [-1, 1]$ . Em geral, para  $a \neq 0$ , temos  $\partial f(a) = \begin{cases} \{1\}, & a > 0; \\ \{-1\}, & a < 0. \end{cases}$

**Observação 1.1.** Observe que o conjunto  $\partial f(0)$ , dado no Exemplo 1.4, nos fornece um conjunto de valores reais. Estes são os coeficientes angulares de retas em  $\mathbb{R}^2$ , que passam pela origem, que substituem aquilo que seria uma reta tangente para a função  $f(x) = |x|$  no ponto  $x = 0$ . Em geral, o subdiferencial de uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , no ponto  $x_0 \in \mathbb{R}^n$ , nos fornece vetores  $y \in \mathbb{R}^n$  que determinam hiperplanos  $H \subset \mathbb{R}^n$  que substituem a noção de espaço tangente ao gráfico de  $f$  no ponto  $(x_0, f(x_0))$ .



**Figura 1.5:** Retas determinadas pelo subdiferencial de  $f$  em  $x = 0$

Note que o fato de uma função convexa ser contínua no ponto  $x_0$  não é necessário para que ela possua um subdiferencial em  $x_0$ . Também poderíamos excluir a convexidade de  $f$ , e ainda seria possível determinar o seu subdiferencial. A importância da convexidade é o teor do teorema a seguir, onde são apresentadas características muito especiais do subdiferencial para este caso.

**Teorema 1.1.** ([8], p.173) *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa. Então, para todo  $x \in \mathbb{R}^n$ , o conjunto  $\partial f(x)$  é convexo, compacto e não-vazio. Além disso, para todo  $d \in \mathbb{R}^n$ , tem-se*

$$\frac{\partial f}{\partial d}(x) = \max_{y \in \partial f(x)} \langle y, d \rangle.$$

*Demonstração.* Ver Teorema 3.4.12 em [8]. ■

No Exemplo 1.4, vemos que  $f$  é convexa e que seu subdiferencial em  $x = 0$  é um conjunto não-vazio, compacto e convexo. Porém, isto não ocorre no seguinte exemplo.

**Exemplo 1.5.** *Seja a função  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $f(x) = x^3$ . O subdiferencial de  $f$  no ponto  $x = 0$  é um conjunto vazio. Com efeito, temos que  $\partial f(0) = \{y \in \mathbb{R}; z^3 \geq yz, \forall z \in \mathbb{R}\}$  e daí*

(i) se  $z = 0$ , então a desigualdade é satisfeita para todo  $y \in \mathbb{R}$ ;

(ii) se  $z > 0$ , então  $y \leq z^2$ ;

(iii) se  $z < 0$ , então  $yz \leq z^3 \leq 0$ .

Observe que (ii) é verdadeira para todo  $y \leq 0$ , porém (iii) será falsa. De fato, em (iii), se  $y = 0$  então  $0 \leq z^3 \leq 0$ , isto é,  $z = 0$ , o que é um absurdo, pois  $z < 0$ . Portanto,  $\partial f(0) = \emptyset$ .

Podemos caracterizar a diferenciabilidade de uma função convexa a partir do seu subdiferencial. De fato, temos:

**Proposição 1.1.** ([8], p.176) *Uma função convexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é diferenciável no ponto  $x \in \mathbb{R}^n$  se, e somente se, o conjunto  $\partial f(x)$  contém um único elemento. Neste caso,*

$$\partial f(x) = \{\nabla f(x)\}.$$

*Demonstração.* Ver Proposição 3.4.5 em [8]. ■

Veja que no Exemplo 1.5 mesmo sendo a função dada diferenciável, tem-se  $\partial f(0) \neq \{f'(0)\}$ . Isto se deve ao fato de  $f(x) = x^3$  não ser uma função convexa em todo intervalo da forma  $[-t, t]$ , para  $t \geq 0$ .

Antes de prosseguirmos, vamos apresentar as definições de *cone* e *cone normal* em  $\mathbb{R}^n$ .

**Definição 1.5.** *Um cone  $C \subset \mathbb{R}^n$  é o conjunto de todos os vetores  $d \in C$  tais que  $\alpha d \in C$  para todo  $\alpha > 0$ .*

**Definição 1.6.** ([8], p.73) *Sejam  $D \subset \mathbb{R}^n$  um conjunto convexo e  $\bar{x} \in D$ . O cone normal (cone de direções normais) no ponto  $\bar{x}$  em relação ao conjunto  $D$  é dado por*

$$\mathcal{N}_D(\bar{x}) = \{d \in \mathbb{R}^n; \langle d, x - \bar{x} \rangle \leq 0, \forall x \in D\}.$$

**Exemplo 1.6.** *No espaço  $\mathbb{R}^n$ , qualquer subespaço de  $\mathbb{R}^n$  e o ortante não-negativo*

$$\mathbb{R}_+^n = \{x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n; x_i \geq 0, i = 1, \dots, n\}$$

*são exemplos de cone.*

**Exemplo 1.7.** *Sejam  $D_1 = \{(x, y) \in \mathbb{R}^2; y = -x\}$  e  $D_2 = \{(x, y) \in \mathbb{R}^2; x \geq 0, y \geq 0\}$ . O cone normal no ponto  $(0, 0)$  em relação a  $D_1$  é o conjunto*

$$\mathcal{N}_{D_1}(0, 0) = \{(x, y) \in \mathbb{R}^2; y = x\}.$$

*Em relação ao conjunto  $D_2$ , o cone normal em  $(0, 0)$  é*

$$\mathcal{N}_{D_2}(0, 0) = \{(x, y) \in \mathbb{R}^2; x \leq 0, y \leq 0\}.$$

*Os conjuntos  $D_1$  e  $D_2$  e os seus respectivos cones normais, em  $(0, 0)$ , são ilustrados nas Figura 1.6 e Figura 1.7.*

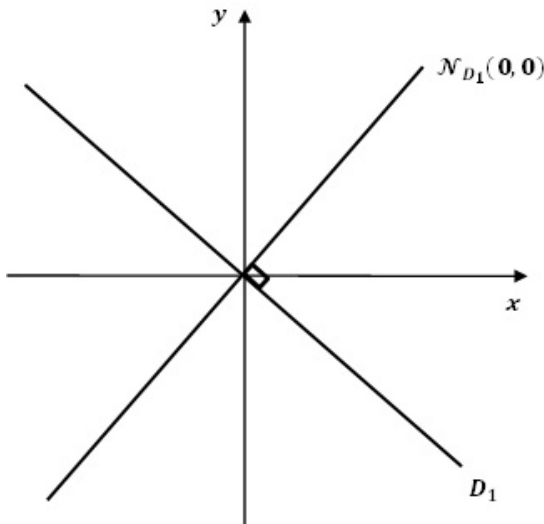


Figura 1.6: Cone Normal ao conjunto  $D_1$

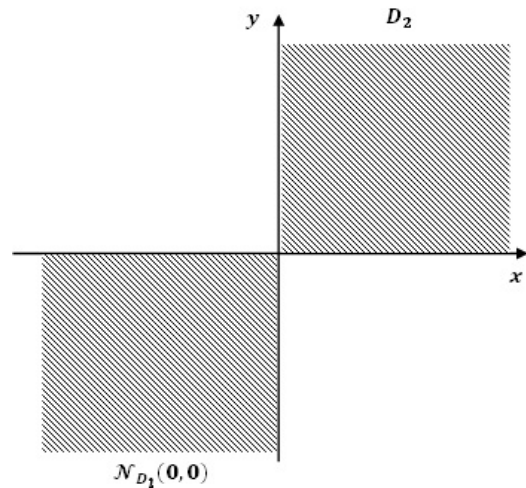


Figura 1.7: Cone Normal ao conjunto  $D_2$

Com a definição de cone normal, podemos caracterizar as condições de otimalidade para o caso em que se deseja minimizar uma função convexa, não necessariamente diferenciável, em um subconjunto de  $\mathbb{R}^n$  convexo. Essa caracterização é apresentada no teorema a seguir.

**Teorema 1.2.** ([8], p.177) *Sejam  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa e  $D \subset \mathbb{R}^n$  um conjunto convexo. Então,  $\bar{x} \in \mathbb{R}^n$  é um minimizador de  $f$  em  $D$  se, e somente se,*

$$\exists y \in \partial f(\bar{x}) \text{ tal que } \langle y, x - \bar{x} \rangle \geq 0 \quad \forall x \in D,$$

ou, equivalentemente,

$$0 \in \partial f(\bar{x}) + \mathcal{N}_D(\bar{x}).$$

Em particular,  $\bar{x}$  é um minimizador de  $f$  no  $\mathbb{R}^n$  se, e somente se,

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

*Demonstração.* Ver Teorema 3.4.13 em [8]. ■

Esta caracterização será usada com bastante frequência ao longo deste trabalho, assim como a propriedade operatória do subdiferencial de funções convexas dada a seguir.

**Proposição 1.2.** ([8], p.180) *Sejam  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$ , funções convexas. Então*

$$\partial \left( \sum_{i=1}^p f_i(x) \right) = \sum_{i=1}^p \partial f_i(x).$$

*Demonstração.* Ver Proposição 3.4.8 em [8]. ■



## 1.2 Operadores Monótonos

Neste trabalho necessitamos conhecer operadores com características bem específicas, uma destas é a monotonicidade.

**Definição 1.7.** ([2], p.6) *Um operador  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  é monótono se*

$$\langle F(x) - F(y), x - y \rangle \geq 0,$$

*para todo  $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ . Caso a desigualdade acima seja estrita para todo  $x \neq y$ ,  $F$  é dito ser um operador estritamente monótono.*

**Exemplo 1.8.** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa diferenciável. O operador*

$$F = \nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

*é monótono (ver Teorema 3.4.7 em [8]).*

Um fato interessante é que para uma transformação linear  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  semidefinida positiva, temos

$$\langle F(v), v \rangle \geq 0, \quad \forall v \in \mathbb{R}^n,$$

se, e somente se,

$$\langle F(x) - F(y), x - y, \rangle = \langle F(x - y), x - y \rangle \geq 0, \forall x, y \in \mathbb{R}^n.$$

Assim, vemos que a matriz associada a todo operador monótono é semidefinida positiva.

Prosseguindo, vejamos uma classe de aplicações que também é importante neste trabalho, a saber, as aplicações ponto-conjunto.

**Definição 1.8.** *Sejam  $\Omega_1$  e  $\Omega_2$  conjuntos quaisquer não-vazios. Uma aplicação  $F : \Omega_1 \rightarrow \Omega_2$  é uma aplicação ponto-conjunto quando todo elemento  $a \in \Omega_1$  está em correspondência com um subconjunto  $A \subset \Omega_2$  através da aplicação  $F$ , isto é, existe  $A \subset \Omega_2$  tal que  $F(a) = A$  qualquer que seja  $a \in \Omega_1$ .*

O subdiferencial dado pela Definição 1.4 é um exemplo de aplicação ponto conjunto. No Exemplo 1.4, o subdiferencial de  $f$  num ponto  $a \in \mathbb{R}$  é a aplicação ponto-conjunto  $\partial f(a) : \mathbb{R} \rightarrow [-1, 1]$  definida pela relação

$$\partial f(a) = \begin{cases} \{1\}, & a > 0; \\ [-1, 1], & a = 0; \\ \{-1\}, & a < 0. \end{cases}$$

Assim como foi definida a monotonicidade de operadores em  $\mathbb{R}^n$ , também podemos definir tal propriedade para aplicações ponto-conjunto. Seja  $F$  uma aplicação que associa a cada

elemento de  $\mathbb{R}^n$  um elemento no seu conjunto de partes  $\mathcal{P}(\mathbb{R}^n)$ , isto é,  $F : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$ . A monotonicidade de  $F$  é, nesse caso, caracterizada da seguinte forma:

**Definição 1.9.** ([2], p.6) *Uma aplicação  $F : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$  é monótona se*

$$\langle u - v, x - y \rangle \geq 0,$$

para todo  $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ , e todo  $(u, v) \in F(x) \times F(y)$ . Se a desigualdade acima é estrita para todo  $x \neq y$ , dizemos que  $F$  é uma aplicação estritamente monótona.

Outra característica importante para aplicações ponto-conjunto é a chamada monotonicidade maximal. Esta garante que o gráfico de um operador monótono não é "coberto" pelo gráfico de outro operador monótono que contenha a sua imagem.

**Definição 1.10.** ([2], p.7) *Uma aplicação  $F : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$  é monótona maximal se, e somente se*

(i)  $F$  é monótona;

(ii) Para toda  $F' : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$  monótona tal que  $F(x) \subset F'(x)$  para todo  $x \in \mathbb{R}^n$ , vale que  $F = F'$ .

**Exemplo 1.9.** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa diferenciável. A aplicação*

$$F = \partial f(x) : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$$

é monótona maximal para todo  $x \in \mathbb{R}^n$  (ver a equação (1.7) e o Teorema 1.1 em [12]).

Finalizamos esta seção, definindo, respectivamente, os zeros de um operador e de uma aplicação ponto-conjunto.

**Definição 1.11.** *Seja  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  um operador. Dizemos que  $z_0 \in \mathbb{R}^n$  é um zero de  $F$  se  $F(z_0) = 0$ .*

**Definição 1.12.** *Seja  $F : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$  uma aplicação ponto-conjunto. Dizemos que  $z_0 \in \mathbb{R}^n$  é um zero de  $F$  se  $0 \in F(z_0)$ .*

## 1.3 Funções e Distâncias de Bregman

Nesta parte apresentamos alguns conceitos que são necessários a análise de convergência do algoritmo do Método Extragradiante Ponto Proximal Inexato (IPPEM) apresentado no capítulo 2. Estes conceitos são *função* e *distância de Bregman*.

**Definição 1.13.** ([6], p.4) *Seja  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa e diferenciável. A distância de Bregman com respeito a  $g$  é a função  $D_g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  definida por*

$$D_g(x, y) = g(x) - g(y) - \langle \nabla g(y), x - y \rangle.$$

A função  $g$ , nesse caso, é dita uma função de Bregman.

Neste trabalho, sempre que dissermos que  $g$  é uma função de Bregman, estamos abreviadamente falando que  $g$  é uma função convexa e diferenciável associada a uma função de distância dada pela Definição 1.13.

**Exemplo 1.10.** *Seja  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  definida por  $g(x) = \frac{1}{2}\|x\|^2$ . É fácil ver que  $g$  é uma função de Bregman associada a distância  $D_g$  definida por*

$$D_g(x, y) = \frac{1}{2}\|x\|^2 - \frac{1}{2}\|y\|^2 - \langle y, x - y \rangle = \frac{1}{2}\|x - y\|^2.$$

A respeito da função  $D_g$  é importante mencionar o seguinte resultado.

**Teorema 1.3.** *Sejam  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função de Bregman e  $D_g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  uma distância de Bregman. Então, as seguintes afirmações são equivalentes:*

- (i) a função  $g$  é convexa e diferenciável;
- (ii)  $D_g$  é uma aplicação não negativa;
- (iii)  $\nabla g : \mathbb{R}^n \rightarrow \mathbb{R}^n$  é um operador monótono.

*Em particular,  $g$  é estritamente convexa e diferenciável  $\Leftrightarrow D_g$  é uma aplicação positiva  $\Leftrightarrow \nabla g$  é um operador estritamente monótono.*

*Demonstração.* Ver Teorema 3.4.7 em [8]. ■

Um conceito importante a respeito da convexidade de  $g$  é o *módulo de total convexidade* de uma função convexa. Tal conceito é descrito como segue:

**Definição 1.14.** ([6], p.4) *A função  $\nu_g : \mathbb{R}^n \times [0, +\infty) \rightarrow [0, +\infty]$  definida por*

$$\nu_g(x, t) = \inf_{y \in \mathbb{R}^n} \{D_g(y, x); \|y - x\| = t\}$$

*é denominada o módulo de total convexidade da função  $g$ .*

Para interpretarmos geometricamente o significado das Definições 1.13 e 1.14, necessitamos da seguinte definição.

**Definição 1.15.** ([8], p.26) *Dizemos que  $v \in \mathbb{R}^n$  é uma direção tangente em relação ao conjunto  $D$  no ponto  $\bar{x} \in D$  quando*

$$d(\bar{x} + tv, D) = o(t),$$

*onde  $\lim_{t \rightarrow 0^+} \frac{o(t)}{t} = 0, t \in \mathbb{R}_+$ . O conjunto de todas as direções tangentes ao conjunto  $D$  no ponto  $\bar{x}$  é denominado cone tangente.*

O módulo de total convexidade mede o desvio mínimo do gráfico da função  $g$ , no ponto  $(x, g(x))$ , em relação ao cone tangente neste ponto.

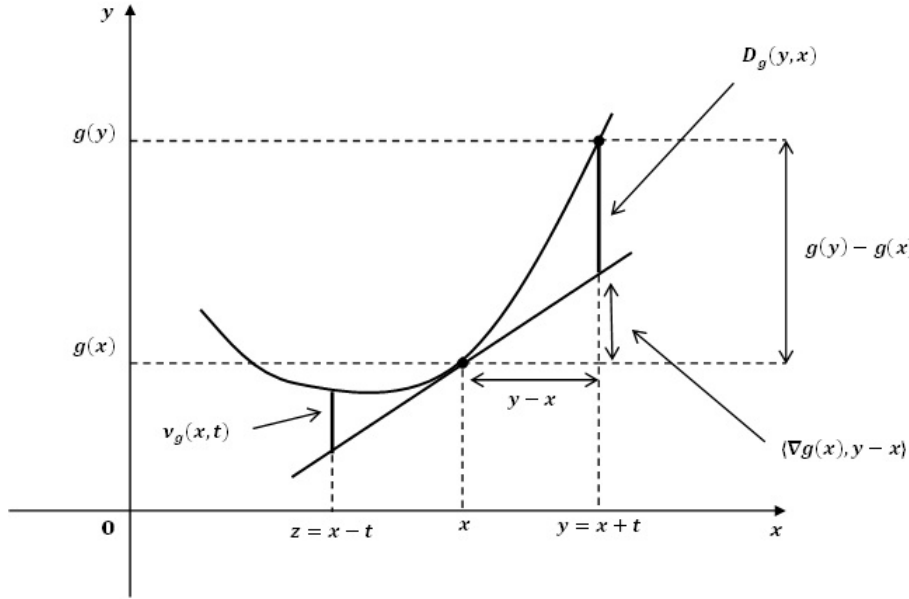


Figura 1.8: Módulo de Total Convexidade e Distância de Bregman

**Exemplo 1.11.** Seja  $g(x) = \frac{1}{2}\|x\|^2$ . O módulo de total convexidade de  $g$  é, neste caso,

$$\nu_g(x, t) = \begin{cases} 0, & \text{se } x = y \\ \frac{t^2}{2}, & \text{se } x \neq y. \end{cases}$$

Vejam agora uma lista de propriedades, que são de grande importância a análise de convergência a ser feita na Seção 2.2, que devem ser satisfeitas pelas funções  $g$ ,  $D_g$  e  $\nu_g$  dadas na Definição 1.13 e Definição 1.14.

**Proposição 1.3.** ([6], p.5) Seja  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  definida por  $g(x) = r\|x\|^s$ . Para todo  $r > 0$  e todo  $s > 1$ , a função  $g$ , assim definida, satisfaz as seguintes propriedades:

- (H1) : Os conjuntos de nível de  $D_g(x, \cdot)$  são limitados para todo  $x \in \mathbb{R}^n$ ;
- (H2) : Para todo conjunto  $\mathbb{L} \subset \mathbb{R}^n$  limitado, tem-se  $\inf_{x \in \mathbb{L}} \nu_g(x, t) > 0, \forall t > 0$ ;
- (H3) :  $\nabla g$  é uniformemente contínua sobre todos os subconjuntos limitados de  $\mathbb{R}^n$ ;
- (H4) :  $\nabla g$  é sobrejetiva, isto é, para todo  $y \in \mathbb{R}^n$ , existe  $x \in \mathbb{R}^n$  tal que  $\nabla g(x) = y$ ;
- (H5) : Seja  $\mathbb{K} \subset \mathbb{R}^n$  um conjunto não-vazio, fechado e convexo. Para todo  $z \in \mathbb{K}$  fixo e todo  $\rho \geq 0$ , tem-se

$$\lim_{\|x\| \rightarrow \infty} (g(x) - \rho\|x - z\|) = +\infty;$$

Além disso, para  $r = \frac{1}{2}$  e  $s = 2$ ,  $g$  também satisfaz

- (H6) : Se  $\{y^k\}$  e  $\{z^k\}$  são seqüências em  $\mathbb{K}$  que convergem para  $y$  e  $z$ , respectivamente, com  $y \neq z$ , então

$$\liminf_{k \rightarrow \infty} |\langle \nabla g(y^k) - \nabla g(z^k), y - z \rangle| > 0.$$

*Demonstração.* Vamos mostrar a proposição para o caso em que  $s = 2$ . Seja  $g(x) = r\|x\|^2$ . Temos:

(H1) : Para todo  $x \in \mathbb{R}^n$  fixado, tem-se  $D_g(x, y) = r\|x - y\|^2$ . Assim, os conjuntos de nível de  $D_g(x, \cdot)$  são da forma

$$\begin{aligned} L_{D_g(x,y)}(c) &= \{y \in \mathbb{R}^n; D_g(x, y) \leq c, \forall x \in \mathbb{R}^n\} \\ &= \{y \in \mathbb{R}^n; r\|x - y\|^2 \leq c, \forall x \in \mathbb{R}^n\} \\ &= B\left[x, \frac{c}{r}\right]. \end{aligned}$$

Sendo  $B[x, \frac{c}{r}]$  uma bola fechada de centro  $x$  e raio  $\frac{c}{r}$ , concluímos que  $L_{D_g(x,y)}(c)$  é limitado, qualquer que seja  $x \in \mathbb{R}^n$ .

(H2) : É fácil verificar que  $\nu_g(x, t) = rt^2$ . Assim, em qualquer conjunto limitado  $\mathbb{L} \subset \mathbb{R}^n$ , tem-se

$$\inf_{x \in \mathbb{L}} \nu_g(x, t) = rt^2 > 0, \quad \forall t > 0.$$

(H3) : Ora, temos  $\nabla g(x) = 2rx$ , isto é,  $\nabla g \equiv 2rI$ . Daí, segue que  $\nabla g$  é uniformemente contínuo em qualquer subconjunto limitado de  $\mathbb{R}^n$ .

(H4) : Como  $\nabla g \equiv 2rI$ , segue que  $\nabla g$  é uma aplicação sobrejetiva.

(H5) : Seja  $\mathbb{K} \subset \mathbb{R}^n$  um conjunto qualquer, não-vazio, fechado e convexo. Para todo  $z \in \mathbb{K}$  fixo e todo  $\rho \geq 0$ , tem-se

$$\begin{aligned} \lim_{\|x\| \rightarrow \infty} (r\|x\|^2 - \rho\|x - z\|) &= \lim_{\|x\| \rightarrow \infty} \|x\|^2 \left( r - \frac{\rho}{\|x\|} \frac{\|x - z\|}{\|x\|} \right) \\ &\geq \lim_{\|x\| \rightarrow \infty} \|x\|^2 \left( r - \frac{\rho}{\|x\|} \frac{\|x\| + \|z\|}{\|x\|} \right) \\ &= \lim_{\|x\| \rightarrow \infty} \|x\|^2 \left( r - \frac{\rho}{\|x\|} \left( 1 + \frac{\|z\|}{\|x\|} \right) \right) \\ &= +\infty. \end{aligned}$$

(H6) : Sejam  $\{y^k\}$  e  $\{z^k\}$  sequências em  $\mathbb{K}$  tais que  $y^k \rightarrow y$  e  $z^k \rightarrow z$ , com  $y \neq z$ . Temos:

$$\liminf_{k \rightarrow \infty} |\langle \nabla g(y^k) - \nabla g(z^k), y - z \rangle| = \liminf_{k \rightarrow \infty} 2r |\langle y^k - z^k, y - z \rangle| = 2r\|y - z\|^2 > 0.$$

■

**Observação 1.2.** *Seja  $\mathbb{L} \subset \mathbb{R}^n$  um conjunto limitado. Veja que se  $\inf_{x \in \mathbb{L}} \nu_g(x, t) > 0$ , para todo  $t > 0$ , então*

$$D_g(y, x) \geq \inf_{y \in \mathbb{R}^n} \{D_g(y, x); \|y - x\| = t\} \geq \inf_{x \in \mathbb{L}} \nu_g(x, t) > 0$$

para todo  $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ . Logo,  $D_g$  é positiva. Portanto, a condição (H2) implica na convexidade estrita de  $g$ .

Finalizamos esta seção apresentando a projeção de Bregman de um ponto sobre um conjunto.

**Proposição 1.4.** ([6], p.6) *Se uma função de Bregman  $g$  satisfaz (H1) – (H2) e  $\mathbb{K} \subset \mathbb{R}^n$  é fechado e convexo, então para todo  $\bar{x} \in \mathbb{R}^n$  existe uma única solução  $\hat{x}$  do problema*

$$\min_{x \in \mathbb{K}} D_g(x, \bar{x})$$

que satisfaz a desigualdade

$$\langle \nabla g(\bar{x}) - \nabla g(\hat{x}), y - \hat{x} \rangle \leq 0, \quad \forall y \in \mathbb{K}.$$

*Demonstração.* Ver Proposição 2.5 em [6]. ■

**Definição 1.16.** *Todo ponto  $\hat{x}$  na Proposição 1.4 é chamado uma projeção de Bregman de  $\bar{x}$  sobre  $\mathbb{K}$ .*

**Exemplo 1.12.** *Seja  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por  $g(x, y) = \frac{1}{2}(x^2 + y^2)$ . A projeção de Bregman do ponto  $(\bar{x}, \bar{y}) = (1, 0)$  sobre o conjunto  $\mathbb{K} = \{(x, y) \in \mathbb{R}^2; x = y\}$  é o ponto  $(\hat{x}, \hat{y}) = \left(\frac{1}{2}, \frac{1}{2}\right)$ . Com efeito, determinar a projeção de Bregman de  $(\bar{x}, \bar{y})$  sobre  $\mathbb{K}$  é equivalente a resolver o problema de minimização convexa restrito*

$$\min \frac{1}{2}[(x - 1)^2 + y^2]$$

$$x - y = 0$$

$$\langle (x - 1, y), (x - z, y - w) \rangle \leq 0, \quad \forall (z, w) \in \mathbb{K}.$$

Vemos que este problema possui uma única solução  $(\hat{x}, \hat{y})$ , pois a sua função objetivo é estritamente convexa e as suas restrições são convexas. Sendo  $(\hat{\lambda}, \hat{\mu}) \in \mathbb{R} \times \mathbb{R}_+$  o multiplicador associado a  $(\hat{x}, \hat{y})$ , temos que  $((\hat{x}, \hat{y}), (\hat{\lambda}, \hat{\mu}))$  é ponto estacionário do sistema

$$(x - 1, y) + \lambda(1, -1) + \mu[(2x - z - 1, 2y - w)] = (0, 0) \tag{1.1}$$

$$x - y = 0 \tag{1.2}$$

$$\mu \geq 0 \tag{1.3}$$

$$(x - 1)(x - z) + y(y - w) \leq 0 \tag{1.4}$$

$$\mu[(x - 1)(x - z) + y(y - w)] = 0. \tag{1.5}$$

Como  $(\hat{x}, \hat{y}) \in \mathbb{K}$  e  $(z, w) \in \mathbb{K}$ , então  $\hat{x} = \hat{y}$ ,  $z = w$  e o sistema (1.1) – (1.5) fica escrito na forma

$$\hat{x} + \hat{\lambda} + \hat{\mu}(2\hat{x} - z - 1) - 1 = 0$$

$$\hat{x} - \hat{\lambda} + \hat{\mu}(2\hat{x} - z) = 0$$

$$\hat{\mu} \geq 0$$

$$(\hat{x} - 1)(\hat{x} - z) + \hat{x}(\hat{x} - z) \leq 0$$

$$\hat{\mu}[(\hat{x} - 1)(\hat{x} - z) + \hat{x}(\hat{x} - z)] = 0.$$

Se supormos que  $\hat{\mu} = 0$ , obtemos o sistema  $\begin{cases} \hat{x} + \hat{\lambda} - 1 = 0 \\ \hat{x} - \hat{\lambda} = 0 \end{cases}$  cuja solução é  $\hat{\lambda} = \hat{x} = \frac{1}{2}$ .

Sendo assim  $(\hat{x}, \hat{y}) = \left(\frac{1}{2}, \frac{1}{2}\right)$ . Se supormos que  $\hat{\mu} > 0$  obtemos, do sistema (1.1) – (1.5), a equação  $(2\hat{x} - 1)(w - z) = 0$  que possui infinitas soluções  $\hat{x} \in \mathbb{R}$ , pois  $w = z$ . Logo, a única solução viável é  $(\hat{x}, \hat{y}) = \left(\frac{1}{2}, \frac{1}{2}\right)$ .

## 1.4 O Problema de Desigualdade Variacional (PDV)

A primeira definição a ser apresentada nesta seção, refere-se ao problema de interesse principal nesta dissertação, a saber, o Problema de Desigualdade Variacional (PDV). Por uma questão de notação, sempre que nos referirmos a um problema (PDV) com estrutura específica determinada por um operador  $F$  sobre um subconjunto  $\mathbb{K}$  de  $\mathbb{R}^n$  escreveremos PDV( $F, \mathbb{K}$ ).

**Definição 1.17** (Problema de Desigualdade Variacional). *Seja  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  um operador contínuo e  $\mathbb{K} \subset \mathbb{R}^n$  um conjunto não-vazio, fechado e convexo. O problema PDV( $F, \mathbb{K}$ ) é definido como:*

$$\text{Encontrar } \bar{x} \in \mathbb{K} \text{ tal que } \langle F(\bar{x}), y - \bar{x} \rangle \geq 0 \text{ para todo } y \in \mathbb{K}.$$

O conjunto de soluções do problema PDV( $F, \mathbb{K}$ ) será denotado por  $S(F, \mathbb{K})$ .

O problema (PDV) pode ser tratado como um caso particular do Problema de Equilíbrio Clássico (PEC) definido a seguir:

**Definição 1.18** (Problema de Equilíbrio Clássico). *Seja  $f : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  uma função e  $\mathbb{K} \subset \mathbb{R}^n$  um conjunto não-vazio, convexo e fechado. O problema (PEC) consiste em:*

$$\text{Encontrar } \bar{x} \in \mathbb{K} \text{ tal que } f(\bar{x}, y) \geq 0, \text{ para todo } y \in \mathbb{K},$$

em que  $f$  satisfaz as seguintes propriedades:

$$(P1) \quad f(x, x) = 0 \text{ para cada } x \in \mathbb{K};$$

$$(P2) \quad f(x, \cdot) : \mathbb{K} \rightarrow \mathbb{R} \text{ é convexa e semicontínua inferiormente para todo } x \in \mathbb{K};$$

$$(P3) \quad f(\cdot, y) : \mathbb{K} \rightarrow \mathbb{R} \text{ é semicontínua superiormente para todo } y \in \mathbb{K}.$$

Uma observação interessante é o fato do problema (PEC) ser uma generalização do problema de minimização de uma função  $q : \mathbb{K} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Ou seja, encontrar  $\bar{x} \in \mathbb{K}$  tal que  $q(\bar{x}) \leq q(x)$  para todo  $x \in \mathbb{K}$  é equivalente a encontrar  $\bar{x} \in \mathbb{K}$  tal que  $f(\bar{x}, y) = q(y) - q(\bar{x}) \geq 0, \forall y \in \mathbb{K}$ .

Neste trabalho, tratamos de um problema de equilíbrio caracterizado especificamente por uma função  $f$  e por um subconjunto  $\mathbb{K}$  de  $\mathbb{R}^n$ . Simbolizamos nesta situação o problema de equilíbrio por  $PE(f, \mathbb{K})$ , e seu conjunto de soluções por  $S(f, \mathbb{K})$ .

Quando  $f$  tem como domínio um conjunto obtido por um produto cartesiano de dois conjuntos, diremos que  $f$  é uma "bifunção". No problema  $PE(f, \mathbb{K})$ , a bifunção  $f$  terá uma característica importante, a saber, a monotonicidade.

**Definição 1.19.** *Uma bifunção  $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  é dita monótona se*

$$f(x, y) + f(y, x) \leq 0, \quad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n. \quad (1.6)$$

Vemos adiante que a monotonicidade de uma bifunção é importante para garantir que a sequência gerada pelo algoritmo IPPEM, a ser apresentado no Capítulo 2, aplicado ao problema  $PE(f, \mathbb{K})$ , converge para alguma solução deste problema.

Um método clássico para resolver o  $PE(f, \mathbb{K})$  é o chamado Método do Ponto Proximal (MPP). No trabalho de Iusem (em [5]), assim como nesta dissertação, o método (MPP) que resolve  $PE(f, \mathbb{K})$  é uma ferramenta essencial na análise de convergência do algoritmo do Método Extragradiante Lagrangeano Aumentado Inexato (IALEM) apresentado no capítulo 3. A fim de assegurar a convergência do método (MPP) a uma solução do problema  $PE(f, \mathbb{K})$ , desde que ela exista, a bifunção  $f$  deve satisfazer, além de (P1)–(P3), as seguintes condições:

(P4) existe  $\theta \geq 0$  tal que

$$f(x, y) + f(y, x) \leq \theta \langle \nabla g(x) - \nabla g(y), x - y \rangle \quad \forall (x, y) \in \mathbb{K} \times \mathbb{K} \quad (1.7)$$

onde  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função de Bregman;

(P5) existe um  $x^* \in S(f, \mathbb{K})$  tal que  $f(y, x^*) \leq 0$  para todo  $y \in \mathbb{K}$ .

Nosso foco principal não será o método (MPP), mas sim um caso particular deste método, o chamado Método do Lagrangeano Aumentado (MLA). Os métodos (MPP) e (MLA) são apresentados na Seção 1.5 e 1.6, respectivamente.

Como afirmamos anteriormente, um problema de desigualdade variacional está relacionado a um problema de equilíbrio, ou seja, o problema  $PDV(F, \mathbb{K})$  pode ser relacionado ao problema  $PE(f, \mathbb{K})$ . De fato, defina  $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  como

$$f(x, y) = \langle F(x), y - x \rangle \quad \forall (x, y) \in \mathbb{K} \times \mathbb{K}. \quad (1.8)$$



Se o problema  $\text{PDV}(F, \mathbb{K})$  tem soluções e  $F$  é monótono e contínuo, então o problema  $\text{PE}(f, \mathbb{K})$ , com  $f$  definida por (1.8), satisfaz as propriedades (P1) – (P5). Com efeito,

(P1) Trivial.

(P2) Para cada  $x_0 \in \mathbb{R}^n$  fixo, temos que  $f(x_0, y) = \langle F(x_0), y - x_0 \rangle$  é uma função afim na variável  $y$ . Logo,  $f$  é convexa e semicontínua inferiormente para cada  $x_0 \in \mathbb{R}^n$ .

(P3) Para cada  $y_0 \in \mathbb{R}^n$  fixo, temos que  $f(x, y_0) = \langle F(x), y_0 - x \rangle$ . Como  $F$  é contínuo, segue que  $f$  é contínua, em particular,  $f$  é semicontínua superiormente para todo  $y_0 \in \mathbb{R}^n$ .

(P4) Como  $F$  é monótono, então

$$f(x, y) + f(y, x) = \langle F(x), y - x \rangle + \langle F(y), x - y \rangle = -\langle F(x) - F(y), x - y \rangle \leq 0$$

Logo,  $f$  é monótona. Como  $g$  é convexa e diferenciável então, pelo Teorema 1.3, concluímos que existe  $\theta \geq 0$  tal que

$$f(x, y) + f(y, x) \leq 0 \leq \theta \langle \nabla g(x) - \nabla g(y), x - y \rangle$$

para todo  $(x, y) \in \mathbb{K} \times \mathbb{K}$ .

(P5) Como o problema  $\text{PDV}(F, \mathbb{K})$  tem soluções, então  $S(F, \mathbb{K})$  é não-vazio. Assim, existe  $x^* \in S(F, \mathbb{K})$ . Agora, note que  $x^* \in S(F, \mathbb{K})$  se, e somente se,  $x^* \in S(f, \mathbb{K})$ , ou seja,  $S(F, \mathbb{K}) = S(f, \mathbb{K})$ . Como  $f$  é monótona, então para cada  $x^* \in S(f, \mathbb{K})$ , tem-se

$$f(y, x^*) \leq -f(x^*, y) \leq 0, \quad \forall y \in \mathbb{K}.$$

**Observação 1.3.** Se  $f$  é monótona e satisfaz as condições (P1) – (P3) do  $\text{PE}(f, \mathbb{K})$ , então  $f$  satisfaz (P5). Por outro lado, a validade de (P5) não implica na monotonicidade de  $f$ .

De fato, considere por exemplo  $\mathbb{K} = \left[ \frac{1}{2}, 1 \right]$  e defina  $f : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  como

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

Notamos que  $f(x, y) + f(y, x) = x(x - y) + y(y - x) = (x - y)^2 \geq 0$ , logo  $f$  não é monótona. No entanto, é fácil verificar que  $f$  satisfaz (P1) – (P3) e (P5). De fato, (P1) – (P3) segue diretamente da convexidade e da continuidade de  $f$ . Com respeito a (P5), observe que se  $f(x^*, y) \geq 0$  para todo  $y \in \mathbb{K}$ , isto é,  $x^* \in S(f, \mathbb{K})$ , então  $(x^* - y) \geq 0$ . Assim, para  $y \in \mathbb{K}$ , tem-se  $f(y, x^*) = y(y - x^*) \leq 0$ .

Assim, com estas suposições sobre  $F$  e com  $f$  definida por (1.8), temos que os problemas  $\text{PDV}(F, \mathbb{K})$  e  $\text{PE}(f, \mathbb{K})$  têm o mesmo conjunto solução. Logo, todos os métodos que resolvem  $\text{PE}(f, \mathbb{K})$  também resolvem  $\text{PDV}(F, \mathbb{K})$ . Daqui em diante, nos métodos apresentados neste trabalho sempre iremos nos referir ao problema  $\text{PE}(f, \mathbb{K})$  ao invés do  $\text{PDV}(F, \mathbb{K})$ , e assumimos que  $f$  é sempre da forma dada em (1.8).

Terminamos esta seção apresentando uma propriedade que garante uma condição necessária e suficiente à existência de soluções do problema  $\text{PE}(f, \mathbb{K})$ , desde que  $f$  seja provida de propriedades de monotonicidade. Esta propriedade, que chamaremos de (P6), é tida na literatura como uma espécie de coercividade.

**Proposição 1.5.** ([6], p.7) *Suponha que  $f$  satisfaz (P1) – (P3) e (P5). Então,  $\text{PE}(f, \mathbb{K})$  tem soluções se, e somente se,  $f$  satisfaz a propriedade*

(P6) *Para qualquer sequência  $\{x^k\} \subseteq \mathbb{K}$  satisfazendo  $\lim_{k \rightarrow \infty} \|x^k\| = +\infty$ , existe  $u \in \mathbb{K}$  e  $k_0 \in \mathbb{N}$  tal que  $f(x^k, u) \leq 0$  para todo  $k \geq k_0$ .*

*Demonstração.* Ver Teorema 4.3 em [4]. ■

**Observação 1.4.** *A propriedade (P6) é reescrita no caso em que  $\mathbb{K}$  é compacto da seguinte forma:*

(P6\*) *Para qualquer sequência  $\{x^k\} \subseteq \mathbb{K}$ , existe  $u \in \mathbb{K}$  e  $k_0 \in \mathbb{N}$  tal que  $f(x^k, u) \leq 0$  para todo  $k \geq k_0$ .*

**Afirmção:** *A Proposição 1.5 é satisfeita quando substituímos (P6) por (P6\*).*

*Com efeito, se não fosse assim, existiria  $\{x^k\} \subseteq \mathbb{K}$  tal que, para qualquer  $u \in \mathbb{K}$ , tem-se  $f(x^k, u) > 0$ , para todo  $k \in \mathbb{N}$ . Como  $\mathbb{K}$  é compacto, então existe uma subsequência  $\{x^{k_j}\} \subset \mathbb{K}$  tal que  $x^{k_j} \rightarrow \bar{x} \in \mathbb{K}$ . Daí, por (P3), tem-se*

$$f(\bar{x}, u) \geq \limsup_{k \rightarrow \infty} f(x^{k_j}, u) > 0, \quad \forall u \in \mathbb{K}.$$

*Em particular, para  $u = \bar{x}$ , tem-se, por (P1),*

$$0 = f(\bar{x}, \bar{x}) = f(\bar{x}, u) > 0,$$

*contradição. Portanto, quando  $\mathbb{K}$  é compacto e  $f$  satisfaz (P1) – (P3) e (P5),  $\text{PE}(f, \mathbb{K})$  tem soluções se, e somente se,  $f$  satisfaz (P6\*).*

## 1.5 O Método do Ponto Proximal

Apresentamos nesta seção o chamado Método do Ponto Proximal (MPP) usado para resolver um problema de minimização irrestrita de uma função. Este método consiste em resolver o problema original através de uma sequência de subproblemas obtidos a partir dele,

e estes subproblemas, em geral, devem ter as suas funções objetivo mais fáceis de minimizar do que a do problema original.

Iniciamos a seção apresentando a classe de problemas que motivou a criação deste método. As informações apresentadas a seguir foram obtidas em ([2], p.p.2 – 3).

**Definição 1.20.** *Sejam  $X$  um espaço de funções e  $L : X \rightarrow X$  um operador diferenciável (ou integro-diferenciável). Um problema da forma*

$$L(f) = 0 \tag{1.9}$$

*em que  $f$  é um elemento do conjunto  $X$ , é dito mal-colocado se não tem soluções ou tem mais que uma solução ou tem uma única solução, porém esta solução não depende de forma contínua sobre alguns parâmetros do operador  $L$ . Quando o problema (1.9) não é mal-colocado o chamamos de bem-colocado.*

O processo de transformação de um problema mal-colocado para um bem-colocado é chamado *regularização*. A idéia deste processo é causar uma pequena perturbação no operador  $L$  de modo que, se o problema (1.9) é mal-colocado, podemos substituí-lo por um operador  $L + \alpha M$ ,  $\alpha \in \mathbb{R}$ , em que  $M : X \rightarrow X$  é tal que o problema

$$L(f) + \alpha M(f) = (L + \alpha M)(f) = 0 \tag{1.10}$$

seja bem-colocado para qualquer  $\alpha > 0$ . Associado a cada  $\alpha$ , (1.10) tem uma única solução  $f_\alpha$  da qual se espera algum tipo de aproximação da solução de (1.9) quando  $\alpha$  se aproximar de zero. Assim, a regularização consiste em causar uma perturbação em um problema mal-colocado a fim de que se obtenha um problema bem-colocado que forneça soluções aproximadas. O número real  $\alpha > 0$  é denominado *parâmetro de regularização*.

O conceito de regularização pode ser aplicado aos problemas de Otimização. De fato, tomando  $X = \mathbb{R}^n$  e  $L = \nabla f$ , onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função convexa diferenciável, (1.9) se torna  $\nabla f(x) = 0$  ou equivalentemente

$$\min_{x \in \mathbb{R}^n} f(x). \tag{1.11}$$

Se assumirmos que  $f$  é limitada inferiormente e tomarmos  $p : \mathbb{R}^n \rightarrow \mathbb{R}$  estritamente convexa tal que  $\lim_{\|x\| \rightarrow \infty} p(x) = +\infty$ , então mesmo que (1.11) não tenha solução ou tenha mais que uma solução, o problema regularizado

$$\min_{x \in \mathbb{R}^n} f(x) + \alpha p(x) \tag{1.12}$$

terá uma única solução para cada  $\alpha > 0$ . Com efeito, o fato de  $f$  ser limitada inferiormente implica que  $\lim_{\|x\| \rightarrow \infty} f(x) + \alpha p(x) = +\infty$ . Como  $f + \alpha p$  é contínua em  $\mathbb{R}^n$ , garante-se a existência de soluções para (1.12). Por fim, note que  $p$  ser estritamente convexa implica que  $f + \alpha p$  é

estrinamente convexa, logo a solução de (1.12) é única.

**Observação 1.5.** *Uma função  $p : \mathbb{R}^n \rightarrow \mathbb{R}$  que satisfaz a condição  $\lim_{\|x\| \rightarrow \infty} p(x) = +\infty$  é na literatura chamada de coerciva. O Corolário 1.2.3 em [8] garante que o problema de minimizar  $p$  em  $\mathbb{R}^n$  sempre possui solução global, desde que  $p$  seja contínua e coerciva.*

**Observação 1.6.** *O fato da convexidade estrita de  $f + \alpha p$  implicar na unicidade da solução de (1.12) advém do Teorema de Minimização Convexa (Teorema 3.1.5) apresentado em ([8], p.77).*

**Observação 1.7.** *Seja  $x(\alpha)$  a única solução de (1.12). Sob algumas hipóteses razoáveis (incluindo a existência de soluções de (1.11)), pode ser provado que  $\lim_{\alpha \rightarrow 0^+} x(\alpha)$  existe e resolve (1.11).*

Em geral, o problema na regularização dada em (1.12) é que para  $\alpha > 0$  muito pequeno, a função  $f + \alpha p$  será numericamente tão mal comportada quanto  $f$ , ou seja, se o sistema  $\nabla f(x) = 0$  é mal-condicionado, então  $\nabla f(x) + \alpha \nabla p(x) = 0$  também será mal-condicionado quando  $\alpha$  se aproximar de 0, apesar do fato de (1.12) ter uma única solução para todo  $\alpha > 0$ . A fim de superarmos tal dificuldade, usamos uma regularização para (1.11) que não exija a regularização do parâmetro  $\alpha$  quando este se aproximar de 0 (uma possibilidade é considerar  $\alpha$  constante). Esta regularização para (1.11) será feita através do chamado Método do Ponto Proximal.

O Método do Ponto Proximal gera uma sequência  $\{x^k\} \subset \mathbb{R}^n$  da seguinte maneira: dado  $x^0 \in \mathbb{R}^n$ , a atualização de  $x^k$  é dada por

$$x^{k+1} = \operatorname{argmin}_{x \in \mathbb{R}^n} \{f(x) + \alpha_k \|x - x^k\|^2\} \quad (1.13)$$

em que  $\alpha_k$  é um número real tal que  $\alpha_k \in (0, \tilde{\alpha}]$  para algum  $\tilde{\alpha} > 0$ . Sob algumas hipóteses razoáveis, prova-se que a sequência gerada pelo método (MPP) converge para uma solução de (1.11). De fato, o seguinte teorema nos fornece tal resultado de convergência.

**Teorema 1.4.** ([2], p.4) *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa e continuamente diferenciável. Assuma que o conjunto  $U$  de minimizadores de  $f$  em  $\mathbb{R}^n$  seja não-vazio. Então a sequência  $\{x^k\}$  gerada por (1.13) converge para um ponto  $x^* \in U$ .*

*Demonstração.* Ver Teorema 2.1 em [2]. ■

Para concluir esta seção, devemos fazer uma importante observação: em Otimização o método (MPP) é visto como um esquema conceitual e não como um algoritmo implementável. Isto se deve ao fato de que os subproblemas

$$\min_{x \in \mathbb{R}^n} \{f(x) + \alpha_k \|x^k - x\|^2\}$$

requerem, em cada iteração do método (MPP), algum tipo de procedimento numérico para resolvê-los. Em muitos problemas, isto pode se tornar computacionalmente muito caro, ou mesmo nem valer a pena, quando os subproblemas são mais difíceis que o problema original.

## 1.6 Métodos de Lagrangeano Aumentado

Os métodos de Lagrangeano aumentado foram motivados para evitar os casos de descontinuidade da função Lagrangeana, apresentada no Lagrangeano clássico, vista como uma função do multiplicador de Lagrange associada a restrição de desigualdade. Estes métodos são tidos na literatura como os métodos mais eficazes para resolver problemas de Programação Não-Linear, eles combinam a função Lagrangeana clássica com métodos de penalidade e de dualidade local para compensar as desvantagens associadas a qualquer método sozinho. O leitor interessado pode obter mais informações a respeito destes métodos em [9] e [10].

Iniciamos a seção apresentando o Lagrangeano clássico, o qual é usado para resolver problemas de minimização de uma função diferenciável em um subconjunto de  $\mathbb{R}^n$  definido por restrições de igualdade e/ou desigualdade. Em seguida, apresenta-se os métodos de Lagrangeano aumentado obtidos a partir da regularização do problema de minimização da Lagrangeana clássica e, por fim, que estes métodos são um caso particular do Método do Ponto Proximal.

Considere o problema

$$\min f(x) \tag{1.14}$$

$$\text{s.a. } h_i(x) \leq 0 \quad (1 \leq i \leq m) \tag{1.15}$$

em que  $f, h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  são funções convexas e diferenciáveis. A Lagrangeana clássica para este problema é a função  $L : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  definida por

$$L(x, \lambda) = \begin{cases} f(x) + \sum_{i=1}^m \lambda_i h_i(x), & \text{se } \lambda_i \geq 0; \\ +\infty, & \text{c.c.} \end{cases} \tag{1.16}$$

Um método para resolver o problema (1.14) – (1.15) é gerar duas seqüências  $\{x^k\} \subset \mathbb{R}^n$  e  $\{\lambda^k\} \subset \mathbb{R}^m$ , de modo que, dado o multiplicador  $\lambda^k \geq 0$ , devemos encontrar  $x^k$  que resolve o problema

$$\min_{x \in \mathbb{R}^n} L(x, \lambda^k)$$

isto é,

$$x^k = \operatorname{argmin}_{x \in \mathbb{R}^n} L(x, \lambda^k). \tag{1.17}$$

Encontrado  $x^k$ , o multiplicador  $\lambda^k$  deverá ser atualizado de maneira apropriada. Aqui surge um problema no aspecto computacional, a função  $L(x, \cdot)$  geralmente não é contínua. Isto se deve ao fato de que cada atualização  $x^k$  nem sempre pode estar bem definida, ou seja,

para cada  $\lambda^k$  dado, o problema (1.17) pode não ter solução alguma ou a sequência gerada não converge a uma solução do problema. Assim é necessário realizar uma modificação nesta função Lagrangeana, de tal modo que se obtenha uma função na qual se apresente propriedades que eliminem essa patologia. Esta função, chamada de Lagrangeana aumentada, será a base do método do Lagrangeano aumentado (MLA).

Para definirmos uma função Lagrangeana aumentada para o problema (1.14) – (1.15), usamos uma função Lagrangeana aumentada para o seguinte problema com restrição de igualdade

$$\min f(x) \quad (1.18)$$

$$\text{s.a. } p_j(x) = 0 \quad (1 \leq j \leq l) \quad (1.19)$$

dada em ([10], p.451) por

$$L_\alpha(x, \lambda) = f(x) + \sum_{j=1}^l \lambda_j p_j(x) + \alpha \sum_{j=1}^l (p_j(x))^2 \quad (1.20)$$

em que  $p_j : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função diferenciável, para cada  $j$ , e  $\alpha > 0$ . Reescrevendo (1.14) – (1.15) como um problema com restrição de igualdade, obtemos o problema equivalente:

$$\min f(x) \quad (1.21)$$

$$\text{s.a. } h_i(x) + z_i^2 = 0 \quad (1 \leq i \leq m) \quad (1.22)$$

Definindo, para cada  $i = 1, \dots, m$ ,  $p_i(x) = h_i(x) + z_i^2$  e usando (1.20), obtemos uma função  $\bar{L}_\alpha : (\mathbb{R}^n \times \mathbb{R}^m) \times \mathbb{R}^m \rightarrow \mathbb{R}$  da forma

$$\bar{L}_\alpha((x, \lambda), z) = f(x) + \sum_{i=1}^m \lambda_i (h_i(x) + z_i^2) + \alpha \sum_{i=1}^m (h_i(x) + z_i^2)^2. \quad (1.23)$$

Agora observe que na variável  $z = (z_1, \dots, z_m)$ , para cada  $\lambda \in \mathbb{R}^m$  dado e  $x \in \mathbb{R}^n$  fixo, o mínimo global do problema

$$\min_{z \in \mathbb{R}^m} \bar{L}_\alpha((x, \lambda), z)$$

é atingido no ponto  $z_i^*$ , onde

$$z_i^* = \begin{cases} 0, & \text{se } h_i(x) + \frac{\lambda_i}{2\alpha} > 0; \\ \pm \sqrt{-(h_i(x) + \frac{\lambda_i}{2\alpha})}, & \text{c.c.} \end{cases}$$

para  $i = 1, \dots, m$ . Ignorando o sinal das coordenadas de  $z_i^*$ , este minimizador é único, pois  $\bar{L}_\alpha$  é estritamente convexa na variável  $z$  e, além disso, vê-se que

$$h_i(x) + (z_i^*)^2 = h_i(x) + \max \left\{ 0, - \left( h_i(x) + \frac{\lambda_i}{2\alpha} \right) \right\} = \max \left\{ h_i(x), \frac{-\lambda_i}{2\alpha} \right\}.$$

Substituindo esta última expressão em (1.23), obtemos

$$\begin{aligned}\bar{L}_\alpha((x, \lambda), z^*) &= f(x) + \sum_{i=1}^m \lambda_i \left( \max \left\{ h_i(x), \frac{-\lambda_i}{2\alpha} \right\} \right) + \alpha \sum_{i=1}^m \left( \max \left\{ h_i(x), \frac{-\lambda_i}{2\alpha} \right\} \right)^2 \\ &= f(x) + \frac{1}{4\alpha} \sum_{i=1}^m [(\max\{0, \lambda_i + 2\alpha h_i(x)\})^2 - (\lambda_i)^2]\end{aligned}$$

o que motiva a seguinte definição.

**Definição 1.21.** *Seja  $x \in \mathbb{R}^n$ . Tome  $\alpha > 0$ . A função  $\bar{L}_\alpha : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  definida por*

$$\bar{L}_\alpha(x, \lambda) = f(x) + \frac{1}{4\alpha} \sum_{i=1}^m [(\max\{0, \lambda_i + 2\alpha h_i(x)\})^2 - (\lambda_i)^2] \quad (1.24)$$

é denominada função Lagrangeana aumentada associada ao problema (1.14) – (1.15).

Agora, podemos descrever o método do Lagrangeano aumentado da seguinte forma:

**Algoritmo do Método do Lagrangeano Aumentado.**

**Passo 1.** Escolher um multiplicador inicial  $\lambda^0 \in \mathbb{R}_+^m$ .

**Passo 2.** Encontrar  $x^k \in \mathbb{R}^n$  tal que

$$x^k = \operatorname{argmin}_{x \in \mathbb{R}^n} \bar{L}_\alpha(x, \lambda^k). \quad (1.25)$$

**Passo 3.** Atualizar o multiplicador  $\lambda^k \in \mathbb{R}^m$  através da fórmula

$$\lambda_i^{k+1} = \max\{0, \lambda_i^k + 2\alpha h_i(x^k)\}. \quad (1.26)$$

**Observação 1.8.** *Pelo Teorema 7.1 de [2], a sequência  $\{\lambda^k\}$  gerada pelo método (MLA) será a mesma sequência gerada pelo método (MPP) aplicado a função  $-\varphi : \mathbb{R}^m \rightarrow \mathbb{R}$  definida por*

$$\varphi(\lambda) = \min_{x \in \mathbb{R}^n} L(x, \lambda) \quad (1.27)$$

com  $L$  dada por (1.16). Observe que  $-\varphi$  é a função objetivo de um problema de minimização convexa irrestrito

$$\min_{\lambda \in \mathbb{R}_+^m} -\varphi(\lambda) \quad (1.28)$$

e que este problema também é o problema dual associado a (1.14) – (1.15). Supondo que (1.14) – (1.15) e (1.28) têm soluções e que a sequência  $\{x^k\}$  gerada pelo método (MLA) está bem definida, isto é, que cada subproblema irrestrito,

$$\min_{x \in \mathbb{R}^n} L_\alpha(x, \lambda^k)$$

com  $L_\alpha$  dada por (1.24), tem solução então, pelo Teorema 4 de [3], a sequência  $\{\lambda^k\}$  converge para uma solução de (1.28) e os pontos de acumulação da sequência  $\{x^k\}$ , caso existam, serão

soluções de (1.14) – (1.15). Assim, concluímos que o método (MLA) é um caso particular do método (MPP).

A partir de agora apresentamos extensões associadas as funções Lagrangeanas apresentadas em (1.16) e (1.24) para o escopo problema  $PE(f, \mathbb{K})$ , e que são os importantes aos algoritmos apresentados nos capítulos 2 e 3. Consideremos inicialmente a extensão associada a função (1.16), que é a bifunção

$$\mathcal{L} : (\mathbb{R}^n \times \mathbb{R}^m) \times (\mathbb{R}^n \times \mathbb{R}^m) \rightarrow \mathbb{R}$$

definida por

$$\mathcal{L}((x, \lambda), (y, \mu)) = f(x, y) + \sum_{i=1}^m \lambda_i h_i(y) - \sum_{i=1}^m \mu_i h_i(x) \quad (1.29)$$

em que  $f$  é dada no problema  $PE(f, \mathbb{K})$ , com  $\mathbb{K} = \{x \in \mathbb{R}^n; h_i(x) \leq 0, i = 1, \dots, m\}$ , onde cada  $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função convexa. Quando consideramos o problema de minimização

$$\min q(x)$$

$$\text{s.a. } h_i(x) \leq 0 \quad (1 \leq i \leq m)$$

como um caso particular do problema  $PE(f, \mathbb{K})$ , com  $f$  dada por  $f(x, y) = q(y) - q(x)$ , a expressão em (1.29) se reduz a

$$\begin{aligned} \mathcal{L}((x, \lambda), (y, \mu)) &= q(y) - q(x) + \sum_{i=1}^m \lambda_i h_i(y) - \sum_{i=1}^m \mu_i h_i(x) \\ &= (q(y) + \sum_{i=1}^m \lambda_i h_i(y)) - (q(x) + \sum_{i=1}^m \mu_i h_i(x)) \\ &= L(y, \lambda) - L(x, \mu) \end{aligned}$$

em que  $L$  é a Lagrangeana clássica definida em (1.16). A bifunção dada em (1.29) é chamada *Lagrangeano Exato*. Uma extensão associada a (1.29), obtida através de (1.24), é o chamado *Lagrangeano Aumentado Proximal*. E a extensão a (1.29), obtida através da sua linearização na variável  $y$ , é o chamado *Lagrangeano Aumentado Linearizado* (ver [5]).

## 1.7 Condições de otimalidade de Karush-Kuhn-Tucker(KKT) para o problema de equilíbrio

Nesta seção apresentamos as condições de otimalidade para os problemas de equilíbrio  $PE(f, \mathbb{K})$ , com  $f$  definida por  $f(x, y) = \langle F(x), y - x \rangle$ , onde  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  é um operador monótono contínuo, e

$$\mathbb{K} = \{x \in \mathbb{R}^n; h_i(x) \leq 0\} \quad (1.30)$$



em que  $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  é convexa para cada  $i = 1, \dots, m$ . Relembramos as condições clássicas de Karush-Kuhn-Tucker(KKT) para o problema de minimização de uma função em um subconjunto não-vazio de  $\mathbb{R}^n$  e, em seguida, estendemos estas condições ao formato mais geral do problema de minimização através do subdiferencial das restrições que definem o conjunto viável  $\mathbb{K}$ .

Inicialmente, façamos uma distinção entre as restrições de desigualdade que compõem o problema de minimização (1.14) – (1.15).

**Definição 1.22.** ([8], p.193) *Seja  $\bar{x} \in \mathbb{K}$ , onde  $\mathbb{K}$  é definido por (1.30). Dizemos que a restrição de desigualdade que corresponde ao índice  $i \in \{1, \dots, m\}$  é ativa no ponto  $\bar{x}$  quando  $h_i(\bar{x}) = 0$ , i.e., quando a desigualdade é satisfeita como igualdade. O conjunto de índice das restrições ativas no ponto  $\bar{x}$  será denotado por*

$$I(\bar{x}) = \{1 \leq i \leq m; h_i(\bar{x}) = 0\}.$$

Agora caracterizemos os pontos estacionários do problema (1.14) – (1.15). Para isto, peçamos que as restrições  $h_i$ , com  $i = 1, \dots, m$ , satisfaçam alguma condição de qualificação de restrição (ver capítulo 4 em [8]), entre estas destacamos a condição de Slater.

**Condição de Slater (CS).** *Se  $\tilde{I}$  é o conjunto de índices  $i$  tal que a função  $h_i$  é afim, então existe  $w \in \mathbb{R}^n$  tal que*

$$\begin{cases} h_i(w) \leq 0, & i \in \tilde{I} \\ h_i(w) < 0, & i \notin \tilde{I} \end{cases}.$$

**Observação 1.9.** *Se o problema (1.14) – (1.15) não apresentar restrições afins então, pela condição (CS), o interior do conjunto viável será não-vazio. Caso contrário, a fronteira do conjunto viável será não-vazia.*

A seguir, destacamos algumas propriedades obtidas sobre o conjunto de multiplicadores associado às restrições do problema.

**Teorema 1.5.** ([11], p.312) *Uma condição necessária e suficiente para que o conjunto de multiplicadores associado a solução  $x^*$  do problema (1.14) – (1.15) seja não-vazio, compacto e convexo é que valha a condição de Slater.*

*Demonstração.* Ver Teorema 2.3.2 em [11]. ■

**Teorema 1.6.** *Sejam  $f, h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  funções diferenciáveis em  $x^* \in \mathbb{R}^n$ . Suponha que  $x^*$  seja um minimizador local do problema (1.14) – (1.15). Então, sob a condição de Slater, tem-se que existe  $\lambda^* \in \mathbb{R}_+^m$  tal que*

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0 \tag{1.31}$$

$$\lambda_i^* h_i(x^*) = 0, i = 1, \dots, m. \tag{1.32}$$

Se o conjunto  $\{\nabla h_i(x^*); i \in I(x^*)\}$  é linearmente independente, então  $\lambda^*$  que satisfaz (1.31)–(1.32) é único.

*Demonstração.* Ver Teorema 4.2.1 em [8]. ■

**Definição 1.23.** Dizemos que  $x^* \in \mathbb{R}^n$  é um ponto estacionário ao problema (1.14) – (1.15) (ou ainda que é um ponto de KKT), quando  $x^* \in \mathbb{K}$  e existe  $\lambda^* \in \mathbb{R}_+^m$  tal que as condições (1.31) – (1.32) são satisfeitas.

**Observação 1.10.** Uma observação que deve ser feita, a respeito dos pontos estacionários e os pontos que são soluções de (1.14) – (1.15), é que sob as condições dadas pelo Teorema 1.6, toda solução local do problema (1.14) – (1.15) é um ponto estacionário do problema. A recíproca é verdadeira somente se a função objetivo e as restrições do problema forem convexas. (Ver Teorema 4.2.2 em [8]).

Podemos resumir a caracterização dos pontos estacionários ao problema (1.14) – (1.15) feita no Teorema 1.6, através do seguinte sistema de equações e inequações nas variáveis  $(x, \lambda)$ :

$$\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla h_i(x) = 0 \quad (1.33)$$

$$\lambda_i \geq 0 \quad (1.34)$$

$$h_i(x) \leq 0 \quad (1.35)$$

$$\lambda_i h_i(x) = 0 \quad (1.36)$$

em que  $1 \leq i \leq m$ . A caracterização dada por (1.33) – (1.36) pode ser estendida ao problema PE( $f, \mathbb{K}$ ), para  $f$  definida por (1.8) e  $h_i$  convexa não necessariamente diferenciável. Nesse caso, os pontos estacionários ao problema PE( $f, \mathbb{K}$ ) serão as soluções do sistema:

$$0 \in F(x) + \sum_{i=1}^m \lambda_i \partial h_i(x) \quad (1.37)$$

$$\lambda_i \geq 0$$

$$h_i(x) \leq 0$$

$$\lambda_i h_i(x) = 0$$

em que  $1 \leq i \leq m$ . Veja que as condições (1.34) – (1.36) permaneceram inalteradas por se tratarem de condições de viabilidade e complementaridade do problema. Porém, a condição (1.33) foi substituída por uma mais adequada ao problema de equilíbrio, de fato a condição dada por (1.37) é bem natural a esta classe de problemas. Para vermos isto claramente, considere o subdiferencial de  $h_i$ , no ponto  $x \in \mathbb{R}^n$ ,

$$\partial h_i(x) = \{y^i \in \mathbb{R}^n; h_i(z) \geq h_i(x) + \langle y^i, z - x \rangle, \forall z \in \mathbb{R}^n\}.$$

Temos que (1.37) é equivalente a

$$0 = F(x) + \sum_{i=1}^m \lambda_i y^i \quad (1.38)$$

para pelo menos um  $y^i \in \partial h_i(x)$ ,  $i = 1, \dots, m$ , onde  $F$  é expressa através de  $f$  em (1.8). Para cada  $z \in \mathbb{R}^n$ , tomando o produto interno em (1.38) por  $z - x$  obtemos,

$$\begin{aligned} 0 &= \langle F(x), z - x \rangle + \sum_{i=1}^m \lambda_i \langle y^i, z - x \rangle \\ &= f(x, z) + \sum_{i=1}^m \lambda_i \langle y^i, z - x \rangle \\ &\leq f(x, z) + \sum_{i=1}^m \lambda_i (h_i(z) - h_i(x)). \end{aligned}$$

Se  $\bar{x}$  é um ponto estacionário ao problema  $PE(f, \mathbb{K})$ , com multiplicador  $\bar{\lambda} \in \mathbb{R}_+^m$  associado, então (1.34) – (1.37) são satisfeitas e portanto, para todo  $z \in \mathbb{K}$ ,

$$0 \leq f(\bar{x}, z) + \sum_{i=1}^m \bar{\lambda}_i (h_i(z) - h_i(\bar{x})) = f(\bar{x}, z) + \sum_{i=1}^m \bar{\lambda}_i h_i(z) - \sum_{i=1}^m \bar{\lambda}_i h_i(\bar{x}) \leq f(\bar{x}, z).$$

Ou seja, a condição (1.37) combinada com (1.34) – (1.36) fornecerá soluções ao problema  $PE(f, \mathbb{K})$ .

**Observação 1.11.** *O sistema (1.34) – (1.37) é a forma não diferenciável do Teorema 1.6 aplicado ao problema  $PE(f, \mathbb{K})$ . Em outras palavras, este sistema descreve as condições de KKT para o problema  $PE(f, \mathbb{K})$  no caso em que  $f(x, \cdot)$  e  $h_i$ , ( $1 \leq i \leq m$ ), são convexas e não necessariamente diferenciáveis. Em particular, a convexidade da função  $f(x, \cdot)$  e do conjunto  $\mathbb{K}$  tornam as condições de otimalidade do Teorema 1.6, aplicado a  $PE(f, \mathbb{K})$ , suficientes.*

**Observação 1.12.** *Seja  $x^*$  uma solução do problema  $PE(f, \mathbb{K})$ . Temos que  $x^*$  é uma solução do problema de minimização convexa*

$$\min_{y \in \mathbb{K}} f(x^*, y)$$

pois  $0 = f(x^*, x^*) \leq f(x^*, y)$  para todo  $y \in \mathbb{K}$ . Assim, se vale a condição de Slater então, pelo Teorema 1.6 aplicado ao caso não diferenciável, tem-se que existe  $\lambda^* \in \mathbb{R}_+^m$  tal que

$$0 \in \partial_y f(x^*, x^*) + \sum_{i=1}^m \lambda_i^* \partial_y h_i(x^*)$$

$$\lambda_i^* h_i(x^*) = 0.$$

Como  $f(x^*, y) = \langle F(x^*), y - x^* \rangle$  é uma função diferenciável na variável  $y$ , então

$$\partial_y f(x^*, y) = \{\nabla_y \langle F(x^*), y - x^* \rangle\} = \{F(x^*)\}.$$

Logo, podemos reescrever as condições dadas acima da seguinte forma:

$$0 \in F(x^*) + \sum_{i=1}^m \lambda_i^* \partial h_i(x^*)$$

$$\lambda_i^* h_i(x^*) = 0.$$

Para terminarmos esta seção, vamos definir a partir (1.34) – (1.37) o conceito de par ótimo para o problema  $PE(f, \mathbb{K})$ .

**Definição 1.24.** ([5], p.15) Dizemos que  $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$  é um par ótimo para  $PE(f, \mathbb{K})$  se

$$0 \in F(x^*) + \sum_{i=1}^m \lambda_i^* \partial h_i(x^*) \tag{1.39}$$

$$\lambda_i^* \geq 0 \tag{1.40}$$

$$h_i(x^*) \leq 0 \tag{1.41}$$

$$\lambda_i^* h_i(x^*) = 0 \tag{1.42}$$

em que  $1 \leq i \leq m$  e  $F$  é dada em (1.8).

**Observação 1.13.** Se  $(x^*, \lambda^*)$  é um par ótimo para o problema  $PE(f, \mathbb{K})$ , então  $x^*$  é um ponto estacionário ao  $PE(f, \mathbb{K})$ .

# Capítulo 2

## Análise de Convergência do Algoritmo IPPEM

Neste capítulo apresentamos um dos algoritmos mais importantes deste trabalho, o chamado Método Extragradiante Ponto Proximal Inexato, ou abreviadamente IPPEM. Este método, quando aplicado ao problema de equilíbrio  $\text{PE}(f, \mathbb{K})$ , gera uma sequência de pontos em  $\mathbb{R}^n$  que converge a uma solução do  $\text{PE}(f, \mathbb{K})$ , desde que  $f$  seja monótona e o problema tenha solução. Ressaltamos, antes de apresentar o esquema do algoritmo, que a proposta não é resolver o problema  $\text{PE}(f, \mathbb{K})$  através de IPPEM, mas utilizar seus resultados de convergência a fim de obter informações a respeito da convergência do algoritmo do Método Extragradiante Lagrangeano Aumentado Inexato (IALEM), o qual será apresentado no capítulo 3.

### 2.1 Funções Regularizadoras e o Método do Ponto Proximal para $\text{PE}(f, \mathbb{K})$

Iniciamos esta seção recordando que o problema de equilíbrio  $\text{PE}(f, \mathbb{K})$  consiste em encontrar  $\bar{x} \in \mathbb{K}$  tal que  $f(\bar{x}, y) \geq 0$  para todo  $y \in \mathbb{K}$ , onde  $f : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  é uma função monótona que satisfaz as propriedades (P1) – (P3) apresentadas na Seção 1.4. No algoritmo IPPEM introduzimos uma função, chamada de *função regularizadora*, que regulariza  $\text{PE}(f, \mathbb{K})$  no mesmo sentido que o método (MPP) regularizava o problema (1.11). Uma forma de regularização do problema de equilíbrio, chamada de *forma exata de regularização*, é feita em [7] da seguinte forma:

Fixe um parâmetro de regularização  $\gamma \in \mathbb{R}_+$  e considere o vetor  $z \in \mathbb{R}^n$ . Para qualquer  $f$  que satisfaça (P1) – (P3), iremos associar outra bifunção  $\tilde{f} : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  que é chamada uma *regularização* de  $f$ . Ela é definida como

$$\tilde{f}(x, y) = f(x, y) + \gamma \langle \nabla g(x) - \nabla g(z), y - x \rangle,$$

onde  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função de Bregman. O problema de equilíbrio  $\text{PE}(\tilde{f}, \mathbb{K})$  é denomi-

nado *forma exata de regularização* do problema PE( $f, \mathbb{K}$ ) ou simplesmente *forma regularizada*.

No algoritmo IPPEM usamos a chamada *forma inexata de regularização* para o problema PE( $f, \mathbb{K}$ ). Seja  $e \in \mathbb{R}^n$  um vetor chamado *vetor de erro*, o qual possui norma pequena. A regularização inexata de  $f$  é a bifunção  $\tilde{f}^e : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  dada por

$$\tilde{f}^e(x, y) = f(x, y) + \gamma \langle \nabla g(x) - \nabla g(z), y - x \rangle + \langle e, y - x \rangle.$$

O problema de equilíbrio PE( $\tilde{f}^e, \mathbb{K}$ ) é denominado *forma inexata de regularização* do problema PE( $f, \mathbb{K}$ ) ou problema perturbado.

Com estas noções de regularização em mãos, podemos estender o método (MPP) apresentado na Seção 1.5, quando aplicado a (1.11), ao problema PE( $f, \mathbb{K}$ ), lembrando que

(P4) existe  $\theta \geq 0$  tal que

$$f(x, y) + f(y, x) \leq \theta \langle \nabla g(x) - \nabla g(y), x - y \rangle \quad \forall (x, y) \in \mathbb{K} \times \mathbb{K} \quad (2.1)$$

onde  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função de Bregman.

**Método do Ponto Proximal para PE( $f, \mathbb{K}$ ).** *Considere uma sequência de parâmetros de regularização  $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ , para algum  $\bar{\gamma} > \theta$ , em que  $\theta \geq 0$  é dado em (P4).*

**Passo 1.** Escolha  $x^0 \in \mathbb{K}$ .

**Passo 2.** Dado  $\{x^k\} \subset \mathbb{K}$ , o iterando  $x^{k+1}$  é a única solução do problema PE( $\tilde{f}_k, \mathbb{K}$ ), em que  $\tilde{f}_k : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  é definida como

$$\tilde{f}_k(x, y) = f(x, y) + \gamma_k \langle \nabla g(x) - \nabla g(x^k), y - x \rangle. \quad (2.2)$$

A função dada em (2.2) é a forma exata da regularização de  $f$  para o problema PE( $f, \mathbb{K}$ ).

**Método do Ponto Proximal Inexato para PE( $f, \mathbb{K}$ ).** *Considere uma sequência de parâmetros de regularização  $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ , para algum  $\bar{\gamma} > \theta$ , em que  $\theta \geq 0$  é dado em (P4).*

**Passo 1.** Escolha  $x^0 \in \mathbb{K}$ .

**Passo 2.** Dado  $\{x^k\} \subset \mathbb{K}$ , o iterando  $x^{k+1}$  é a única solução do problema perturbado PE( $\tilde{f}_k^e, \mathbb{K}$ ), em que  $\tilde{f}_k^e : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  é dada por

$$\tilde{f}_k^e(x, y) = f(x, y) + \gamma_k \langle \nabla g(x) - \nabla g(x^k), y - x \rangle - \langle e^k, y - x \rangle. \quad (2.3)$$

A expressão em (2.3) é a forma inexata da regularização de  $f$  para o problema PE( $f, \mathbb{K}$ ).

Os métodos de ponto proximal, aplicados ao PE( $f, \mathbb{K}$ ), são importantes na análise de convergência dos métodos de Lagrangeno aumentado para problemas de desigualdade variacional. Não apresentamos a análise de convergência do método (MPP) aplicado ao PE( $f, \mathbb{K}$ ), por que isto foge ao objetivo deste trabalho, porém, ao leitor interessado, tal análise pode ser encontrada em ([6], p.p.12 – 16).

Antes de apresentarmos o algoritmo IPPEM na Seção 2.2, vale a pena fazer a seguinte observação a respeito do vetor de erro apresentado na regularização inexata do problema  $PE(f, \mathbb{K})$ .

**Observação 2.1.** Quando dizemos que o vetor de erro tem norma "pequena", estamos nos referindo ao sentido que  $e^k$  é limitado por uma função, determinada através de dados previamente informados, em cada iteração  $k$ , a saber o parâmetro de regularização  $\gamma_k$ , a iteração  $x^k$  do (MPP) aplicado ao  $PE(f, \mathbb{K})$  e o vetor  $\tilde{x}^k \in S(\tilde{f}_k^e, \mathbb{K})$ .

## 2.2 O Método Extragradiante Ponto Proximal Inexato

Iniciamos esta seção descrevendo o esquema geral do Método Extragradiante Ponto Proximal Inexato (IPPEM). Em seguida, apresentamos um estudo detalhado da análise de convergência dada em [6].

**Algoritmo IPPEM para  $PE(f, \mathbb{K})$ .** Considere uma sequência exógena limitada  $\{\gamma_k\} \subset \mathbb{R}_+$  de parâmetros de regularização e uma tolerância de erro relativo  $\sigma \in (0, 1)$ .

**Passo 1.** Iniciar o algoritmo com  $x^0 \in \mathbb{K}$ .

**Passo 2.** Dado  $x^k$ , encontrar  $\tilde{x}^k \in \mathbb{R}^n$  tal que  $\tilde{x}^k$  resolve  $PE(\tilde{f}_k^e, \mathbb{K})$ , com  $\tilde{f}_k^e$  dada por

$$\tilde{f}_k^e(x, y) = f(x, y) + \gamma_k \langle \nabla g(x) - \nabla g(x^k), y - x \rangle - \langle e^k, y - x \rangle,$$

onde  $e^k$  satisfaz:

$$D_g \left( \tilde{x}^k, (\nabla g)^{-1} \left( \nabla g(\tilde{x}^k) - \frac{1}{\gamma_k} e^k \right) \right) \leq \sigma D_g(\tilde{x}^k, x^k) \quad (2.4)$$

**Passo 3.** Se  $\tilde{x}^k = x^k$ , então PARE!. Caso contrário, faça

$$x^{k+1} = (\nabla g)^{-1} \left( \nabla g(\tilde{x}^k) - \frac{1}{\gamma_k} e^k \right), \quad (2.5)$$

e volte para o Passo 2.

**Observação 2.2.** Vamos interpretar geometricamente o algoritmo IPPEM no caso em que  $g(x) = \frac{1}{2} \|x\|^2$ . Nesta situação, a equação (2.5) é

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

O iterando  $\tilde{x}^k$  define, através do vetor de erro  $e^k$ , um hiperplano  $H_k$  que separa  $x^k$  do conjunto de soluções do problema  $PE(f, \mathbb{K})$ . A atualização  $x^{k+1}$  é obtida tomando um passo a partir de  $x^k$ , na direção de  $H_k$ , com respeito a métrica induzida pela distância de Bregman  $D_g$ . Assim,  $x^{k+1}$  se aproximaria, a cada iteração  $x^k$  dada, do conjunto  $S(f, \mathbb{K})$ , conforme a Figura 2.1.

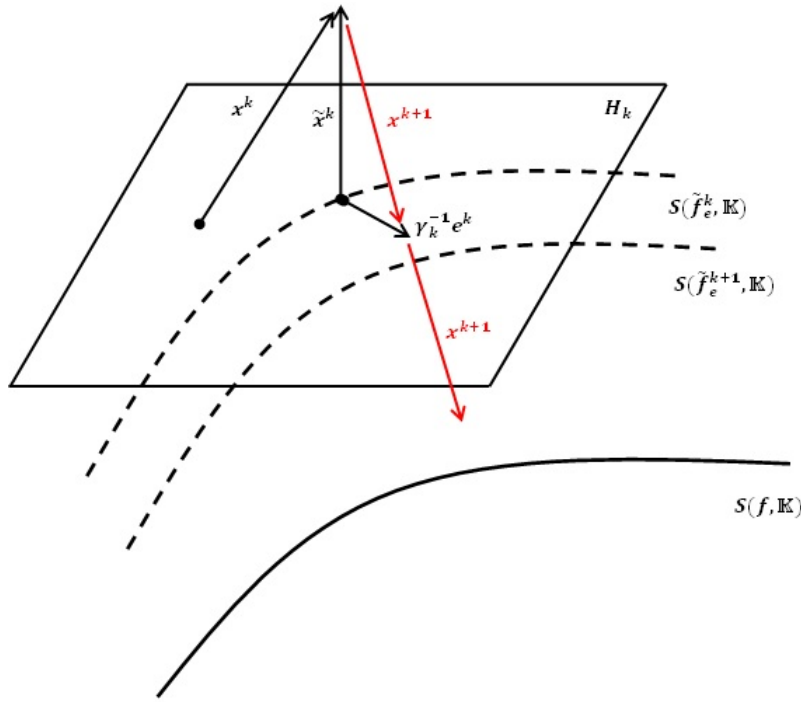


Figura 2.1: Iteração do algoritmo IPPEM

**Observação 2.3.** A expressão exógena se refere a dados obtidos através de informações externas ao problema, por exemplo, os dados obtidos através do estudo empírico de um fenômeno associado ao problema.

Mostramos a seguir os resultados de convergência do algoritmo IPPEM. Devemos ressaltar que o resultado mais importante dessa análise, o Teorema 2.1, garante que, sob algumas hipóteses para  $f$ , a sequência gerada pelo algoritmo IPPEM converge a uma solução do problema  $PE(f, \mathbb{K})$ .

O primeiro resultado é uma propriedade da sequência gerada pelo algoritmo IPPEM, que mostra a relação, em termos da distância de Bregman, entre os seus iterandos e um ponto qualquer do conjunto de soluções do problema  $PE(f, \mathbb{K})$ .

**Lema 2.1.** Considere o problema  $PE(f, \mathbb{K})$  onde  $f$  satisfaz (P1) – (P3) e (P5). Tome  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  satisfazendo (H1) – (H5). Sejam  $\{x^k\}, \{\tilde{x}^k\}, \{\gamma_k\}$  e  $\sigma$  dados pelo algoritmo IPPEM. Então, para qualquer  $x^* \in S(f, \mathbb{K})$ , tem-se

i)

$$[D_g(x^*, x^{k+1}) - D_g(x^*, x^k)] - [D_g(\tilde{x}^k, x^{k+1}) - D_g(\tilde{x}^k, x^k)] = \langle \nabla g(x^k) - \nabla g(x^{k+1}), x^* - \tilde{x}^k \rangle \quad (2.6)$$

ii)

$$D_g(x^*, x^{k+1}) \leq D_g(x^*, x^k) - \frac{1}{\gamma_k} \langle v^k, \tilde{x}^k - x^* \rangle - (1 - \sigma) D_g(\tilde{x}^k, x^k) \leq D_g(x^*, x^k) \quad (2.7)$$



com

$$v^k = \gamma_k(\nabla g(x^k) - \nabla g(\tilde{x}^k)) + e^k. \quad (2.8)$$

Em particular, (2.7) vale se  $f$  é monótona e satisfaz (P1) – (P3).

*Demonstração.* *i)* Seja

$$M = D_g(x^*, x^k) + \langle \nabla g(x^k) - \nabla g(x^{k+1}), x^* - \tilde{x}^k \rangle + D_g(\tilde{x}^k, x^{k+1}) - D_g(\tilde{x}^k, x^k).$$

Pela definição de distância de Bregman  $D_g$  dada na Definição 1.13, temos que

$$\begin{aligned} M &= [g(x^*) - g(x^k) - \langle \nabla g(x^k), x^* - x^k \rangle] + \langle \nabla g(x^k) - \nabla g(x^{k+1}), x^* - \tilde{x}^k \rangle \\ &\quad + [g(\tilde{x}^k) - g(x^{k+1}) - \langle \nabla g(x^{k+1}), \tilde{x}^k - x^{k+1} \rangle] - [g(\tilde{x}^k) - g(x^k) - \langle \nabla g(x^k), \tilde{x}^k - x^k \rangle] \\ &= g(x^*) - g(x^{k+1}) - \langle \nabla g(x^{k+1}), x^* - x^{k+1} \rangle \\ &= D_g(x^*, x^{k+1}), \end{aligned}$$

donde obtemos (2.6).

*ii)* Aplicando o operador  $\nabla g$  na igualdade (2.5), obtemos

$$\nabla g(\tilde{x}^k) = \nabla g(x^{k+1}) + \frac{1}{\gamma_k} e^k. \quad (2.9)$$

Substituindo (2.9) em (2.8), ficamos com

$$v^k = \gamma_k(\nabla g(x^k) - \nabla g(x^{k+1})). \quad (2.10)$$

Agora, por (2.5), verifica-se que (2.4) é reescrita como

$$D_g(\tilde{x}^k, x^{k+1}) \leq \sigma D_g(\tilde{x}^k, x^k). \quad (2.11)$$

Usando a relação obtida em *i)*, combinada com (2.10) e (2.11), obtemos

$$D_g(x^*, x^{k+1}) = D_g(x^*, x^k) + \frac{1}{\gamma_k} \langle v^k, x^* - \tilde{x}^k \rangle + D_g(\tilde{x}^k, x^{k+1}) - D_g(\tilde{x}^k, x^k) \quad (2.12)$$

$$\leq D_g(x^*, x^k) - \frac{1}{\gamma_k} \langle v^k, \tilde{x}^k - x^* \rangle - (1 - \sigma) D_g(\tilde{x}^k, x^k). \quad (2.13)$$

Como  $D_g$  é uma aplicação não-negativa e  $\sigma \in [0, 1)$ , então

$$(1 - \sigma) D_g(\tilde{x}^k, x^k) \geq 0. \quad (2.14)$$

Agora veja que, por (2.3) e (2.8),  $\tilde{x}^k \in S(\tilde{f}_k^e, \mathbb{K})$  implica que

$$\begin{aligned} 0 \leq \tilde{f}_k^e(\tilde{x}^k, y) &= f(\tilde{x}^k, y) + \gamma_k \langle \nabla g(\tilde{x}^k) - \nabla g(x^k), y - \tilde{x}^k \rangle - \langle e^k, y - \tilde{x}^k \rangle \\ &= f(\tilde{x}^k, y) + \gamma_k \langle -(\nabla g(x^k) - \nabla g(\tilde{x}^k) + e^k), y - \tilde{x}^k \rangle \\ &= f(\tilde{x}^k, y) + \langle -v^k, y - \tilde{x}^k \rangle \end{aligned}$$

para todo  $y \in \mathbb{K}$ . Em particular, para  $y = x^*$ , temos

$$f(\tilde{x}^k, x^*) \geq \langle v^k, x^* - \tilde{x}^k \rangle.$$

Como  $x^* \in S(f, \mathbb{K})$ , então  $f(x^*, \tilde{x}^k) \geq 0$ . Daí, por (P5) vale que  $f(\tilde{x}^k, x^*) \leq 0$ , logo

$$\langle v^k, \tilde{x}^k - x^* \rangle \geq 0. \quad (2.15)$$

Portanto, de (2.14) e (2.15) aplicado a (2.13), concluímos que

$$D_g(x^*, x^{k+1}) \leq D_g(x^*, x^k) - \left[ \frac{1}{\gamma_k} \langle v^k, \tilde{x}^k - x^* \rangle + (1 - \sigma) D_g(\tilde{x}^k, x^k) \right] \leq D_g(x^*, x^k)$$

■

Antes de prosseguirmos com os resultados de convergência para o algoritmo IPPEM, apresentamos um resultado de caráter técnico que será essencial na prova da Proposição 2.1.

**Lema 2.2.** ([6], p.6) *Assuma que  $g$  satisfaz a propriedade:*

(H2) *Para todo conjunto  $\mathbb{L} \subset \mathbb{R}^n$  limitado, tem-se  $\inf_{x \in \mathbb{L}} \nu_g(x, t) > 0, \forall t > 0$ ,*

onde  $\nu_g$  é dado pela Definição 1.14. Sejam  $\{x^k\}, \{y^k\} \subset \mathbb{R}^n$  duas seqüências tais que pelo menos uma delas é limitada. Se  $\lim_{k \rightarrow \infty} D_g(y^k, x^k) = 0$ , então  $\lim_{k \rightarrow \infty} \|x^k - y^k\| = 0$ .

*Demonstração.* Suponha, por absurdo, que existam duas seqüências  $\{x^k\}$  e  $\{y^k\}$  em  $\mathbb{R}^n$  tais que  $\{x^k\}$  é limitada e  $\{D_g(y^k, x^k)\} \rightarrow 0$ , mas  $\{\|y^k - x^k\|\}$  não converge para zero, quando  $k \rightarrow \infty$ . Temos que existem  $\alpha > 0$  e subsequências  $\{x^{k_i}\} \subset \{x^k\}$  e  $\{y^{k_i}\} \subset \{y^k\}$  tais que  $\alpha \leq \|y^{k_i} - x^{k_i}\|$ , para todo  $i \in \mathbb{N}$ . Sendo a seqüência  $\{x^k\}$  limitada, então o conjunto

$$\mathbb{L} = \{x \in \mathbb{R}^n; x \in \{x^k\}\}$$

é limitado. Assim,

$$D_g(y^{k_i}, x^{k_i}) \geq \nu_g(x^{k_i}, \|y^{k_i} - x^{k_i}\|) \geq \nu_g(x^{k_i}, \alpha) \geq \inf_{x \in \mathbb{L}} \nu_g(x, \alpha)$$

para todo  $i \in \mathbb{N}$ . Daí, tomando o limite quando  $k \rightarrow \infty$ , obtemos

$$0 = \lim_{k \rightarrow \infty} D_g(y^{k_i}, x^{k_i}) \geq \inf_{x \in \mathbb{L}} \nu_g(x, \alpha) \geq 0$$

isto é,  $\inf_{x \in \mathbb{L}} \nu_g(x, \alpha) = 0$  para todo  $\alpha > 0$ , o que é uma contradição por  $g$  satisfazer (H2). ■

O próximo resultado que apresentamos, mostra as propriedades de convergência das seqüências  $\{x^k\}$  e  $\{\tilde{x}^k\}$ , dadas pelo algoritmo IPPEM, e da seqüência  $\{v^k\}$ , dada no Lema 2.1.

**Proposição 2.1.** *Considere o problema  $PE(f, \mathbb{K})$ . Assuma que  $f$  satisfaz (P1)–(P5). Tome uma função de Bregman  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  qualquer satisfazendo (H1) – (H5), e uma seqüência exógena  $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ , onde  $\theta$  é uma constante dada em (P4). Seja  $\{x^k\}$  a seqüência gerada pelo Algoritmo IPPEM. Se o problema  $PE(f, \mathbb{K})$  tem soluções, então:*

- (i) Para todo  $x^* \in S(f, \mathbb{K})$ ,  $\{D_g(x^*, x^k)\}$  é não crescente e convergente;
- (ii)  $\{x^k\}$  é limitada;
- (iii)  $\sum_{k=0}^{\infty} \frac{1}{\gamma_k} \langle v^k, \tilde{x}^k - x^* \rangle < \infty$  em que  $v^k = \gamma_k(\nabla g(x^k) - \nabla g(\tilde{x}^k)) + e^k$ ;
- (iv)  $\sum_{k=0}^{\infty} D_g(\tilde{x}^k, x^k) < \infty$ ;
- (v)  $\sum_{k=0}^{\infty} D_g(\tilde{x}^k, x^{k+1}) < \infty$ ;
- (vi)  $\{\tilde{x}^k - x^k\}$  converge para 0, e conseqüentemente  $\{\tilde{x}^k\}$  é limitada;
- (vii)  $\{x^{k+1} - x^k\}$  converge para 0;
- (viii)  $\{v^k\}$  converge para 0.

*Demonstração.* (i) Seja  $x^* \in S(f, \mathbb{K})$ . Pelo Lema 2.1, temos que  $D_g(x^*, x^{k+1}) \leq D_g(x^*, x^k)$  para qualquer  $k \geq 0$ , isto é,  $\{D_g(x^*, x^k)\}$  é não-crescente. Sendo  $\{D_g(x^*, x^k)\}$  não-negativa, vem que, para todo  $k \in \mathbb{N}$ ,

$$0 \leq \dots \leq D_g(x^*, x^{k+1}) \leq D_g(x^*, x^k) \leq \dots \leq D_g(x^*, x^1) \leq D_g(x^*, x^0).$$

Portanto,  $\{D_g(x^*, x^k)\}$  é uma seqüência de números reais monótona e limitada, logo é convergente.

(ii) Como  $D_g(x^*, x^k) \leq c = D_g(x^*, x^0)$  para todo  $k \geq 1$ , então  $\{x^k\}$  está contido no conjunto de nível  $c$  de  $D_g(x^*, \cdot)$ , o qual é limitado por (H1). Daí, concluímos que  $\{x^k\}$  é limitada.

(iii) – (iv) Pela relação (2.7) do Lema 2.1, temos que

$$\frac{1}{\gamma_k} \langle v^k, \tilde{x}^k - x^* \rangle + (1 - \sigma) D_g(\tilde{x}^k, x^k) \leq D_g(x^*, x^k) - D_g(x^*, x^{k+1})$$

Daí,

$$\sum_{k=0}^{\infty} \gamma_k^{-1} \langle v^k, \tilde{x}^k - x^* \rangle + (1-\sigma) \sum_{k=0}^{\infty} D_g(\tilde{x}^k, x^k) \leq \sum_{k=0}^{\infty} D_g(x^*, x^k) - \sum_{k=0}^{\infty} D_g(x^*, x^{k+1}) = D_g(x^*, x^0) < \infty$$

Portanto,  $\sum_{k=0}^{\infty} \frac{1}{\gamma_k} \langle v^k, \tilde{x}^k - x^* \rangle < \infty$  e  $\sum_{k=0}^{\infty} D_g(\tilde{x}^k, x^k) < \infty$ .

(v) Pela desigualdade dada em (2.11), e por  $\sigma$  pertencer ao intervalo  $(0, 1)$ , segue que

$$\sum_{k=0}^{\infty} D_g(\tilde{x}^k, x^{k+1}) \leq \sigma \sum_{k=0}^{\infty} D_g(\tilde{x}^k, x^k) \leq \sum_{k=0}^{\infty} D_g(\tilde{x}^k, x^k) < \infty.$$

(vi) Por (iv),

$$\sum_{k=0}^{\infty} D_g(\tilde{x}^k, x^k) < \infty \Rightarrow \lim_{k \rightarrow \infty} D_g(\tilde{x}^k, x^k) = 0.$$

Em (ii) vimos que  $\{x^k\}$  limitada, então pelo Lema 2.2,  $\lim_{k \rightarrow \infty} \|x^k - \tilde{x}^k\| = 0$ . Logo,

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

Em particular,  $\{\tilde{x}^k - x^k\} \rightarrow 0$  implica que  $\{\tilde{x}^k - x^k\}$  é limitada, isto é, existe  $M_1 > 0$  tal que  $\|\tilde{x}^k - x^k\| \leq M_1$ . Novamente por (ii), temos que  $\|x^k\| \leq M_2$  para algum  $M_2 > 0$ . Assim,

$$\|\tilde{x}^k\| = \|(\tilde{x}^k - x^k) + x^k\| \leq \|(\tilde{x}^k - x^k)\| + \|x^k\| \leq M_1 + M_2, \quad \forall k \geq 0.$$

Portanto,  $\{\tilde{x}^k\}$  é limitada.

(vii) Observe que combinando (ii), (iv) e (v) com o Lema 2.2, obtemos

$$\begin{aligned} \lim_{k \rightarrow \infty} \|x^k - x^{k+1}\| &= \lim_{k \rightarrow \infty} \|(x^k - \tilde{x}^k) - (x^{k+1} - \tilde{x}^k)\| \\ &\leq \lim_{k \rightarrow \infty} \|(x^k - \tilde{x}^k)\| + \lim_{k \rightarrow \infty} \|(x^{k+1} - \tilde{x}^k)\| = 0. \end{aligned}$$

Segue daí que  $\lim_{k \rightarrow \infty} (x^k - x^{k+1}) = 0$ .

(viii) Como  $\gamma_k \subset (\theta, \bar{\gamma}]$  e  $\nabla g$  é uniformemente contínuo, devido a (H3), então, por (vii),

$$\lim_{k \rightarrow \infty} \gamma_k [\nabla g(x^k) - \nabla g(x^{k+1})] = 0,$$

isto é,

$$\lim_{k \rightarrow \infty} v^k = 0. \quad \blacksquare$$

Vamos ao resultado de convergência mais importante desta seção, entretanto, antes de enunciá-lo, precisamos da seguinte definição.

**Definição 2.1.** Uma seqüência  $\{z^k\} \subset \mathbb{K}$  é uma seqüência que resolve assintoticamente  $PE(f, \mathbb{K})$  se

$$\liminf_{k \rightarrow \infty} f(z^k, y) \geq 0$$

para todo  $y \in \mathbb{K}$ .

**Teorema 2.1.** Considere o problema  $PE(f, \mathbb{K})$ . Assuma que  $f$  satisfaz (P1) – (P5). Tome uma função de Bregman  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  qualquer satisfazendo (H1) – (H5), e uma seqüência exógena  $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ , onde  $\theta$  é uma constante dada em (P4). Seja  $\{x^k\}$  a seqüência gerada pelo Algoritmo IPPEM. Se  $PE(f, \mathbb{K})$  tem soluções, então

- (i)  $\{\tilde{x}^k\}$  é uma seqüência que resolve assintoticamente  $PE(f, \mathbb{K})$ ;
- (ii) todos os pontos de acumulação de  $\{x^k\}$  resolvem  $PE(f, \mathbb{K})$ .
- (iii) Se, além disso,  $g$  satisfaz (H6), então a seqüência  $\{x^k\}$  converge para alguma solução  $x^*$  do  $PE(f, \mathbb{K})$ . Isto é, a seqüência  $\{x^k\}$  gerada por IPPEM possui um único ponto de acumulação.

*Demonstração.* (i) Fixe  $y \in \mathbb{K}$ . Como  $\tilde{x}^k$  resolve  $PE(\tilde{f}_k^e, \mathbb{K})$  então, pela definição de  $\tilde{f}_k^e$  e através da desigualdade de Cauchy-Schwartz, obtemos

$$\begin{aligned} 0 \leq \tilde{f}_k^e(\tilde{x}^k, y) &= f(\tilde{x}^k, y) + \gamma_k \langle \nabla g(\tilde{x}^k) - \nabla g(x^k), y - \tilde{x}^k \rangle - \langle e^k, y - \tilde{x}^k \rangle \\ &= f(\tilde{x}^k, y) + \langle -v^k, y - \tilde{x}^k \rangle \\ &\leq f(\tilde{x}^k, y) + \|v^k\| \|y - \tilde{x}^k\|. \end{aligned}$$

Pela Proposição 2.1, itens (ii) e (vi), temos que  $\{\tilde{x}^k\}$  é limitada e  $\{y - \tilde{x}^k\}$  é limitada para todo  $y \in \mathbb{K}$ . Logo, tomando o limite na expressão acima, quando  $k \rightarrow \infty$ , e usando (viii) obtemos

$$0 \leq \liminf_{k \rightarrow \infty} f(\tilde{x}^k, y), \quad \forall y \in \mathbb{K}.$$

(ii) Pela Proposição 2.1 itens (ii) e (vi), tem-se que  $\{x^k\}$  tem pontos de acumulação, todos os quais também são pontos de acumulação de  $\{\tilde{x}^k\}$ . Além disso, estes pontos pertencem a  $\mathbb{K}$ , pois ele é fechado. Seja  $\tilde{x} = \lim_{k \rightarrow \infty} \tilde{x}^{l_k}$  em que  $\{\tilde{x}^{l_k}\}$  é uma subsequência de  $\{\tilde{x}^k\}$ . Como  $f(\cdot, y)$  é semicontínua superiormente, então

$$0 \leq \limsup_{k \rightarrow \infty} f(\tilde{x}^{l_k}, y) \leq f(\tilde{x}, y), \quad \forall y \in \mathbb{K},$$

e portanto  $\tilde{x} \in S(f, \mathbb{K})$ .

(iii) Resta provar que  $\{x^k\}$  possui um único ponto de acumulação. Suponha que  $\bar{x}$  é outro ponto de acumulação de  $\{x^k\}$ , isto é,  $\bar{x} = \lim_{k \rightarrow \infty} x^{j_k}$  em que  $\{x^{j_k}\} \subset \{x^k\}$ . Como  $\tilde{x}, \bar{x} \in S(f, \mathbb{K})$  então, pela Proposição 2.1 item (i),  $\{D_g(\bar{x}, x^k)\}$  e  $\{D_g(\tilde{x}, x^k)\}$  convergem, digamos para  $\eta \geq 0$

e  $\mu \geq 0$ , respectivamente. Através da relação (2.6), tem-se

$$\langle \nabla g(x^{l_k}) - \nabla g(x^{j_k}), \bar{x} - \tilde{x} \rangle = D_g(\bar{x}, x^{j_k}) - D_g(\bar{x}, x^{l_k}) + D_g(\tilde{x}, x^{l_k}) - D_g(\tilde{x}, x^{j_k}).$$

Logo,

$$|\langle \nabla g(x^{l_k}) - \nabla g(x^{j_k}), \bar{x} - \tilde{x} \rangle| \leq |D_g(\bar{x}, x^{j_k}) - D_g(\bar{x}, x^{l_k})| + |D_g(\tilde{x}, x^{l_k}) - D_g(\tilde{x}, x^{j_k})|$$

Tomando o limite nesta expressão, quando  $k \rightarrow \infty$ , obtemos

$$\liminf_{k \rightarrow \infty} |\langle \nabla g(x^{l_k}) - \nabla g(x^{j_k}), \bar{x} - \tilde{x} \rangle| \leq |\eta - \eta| + |\mu - \mu| = 0$$

o que contradiz a hipótese (H6). Portanto,  $\tilde{x} = \bar{x}$ . ■

**Observação 2.4.** No Teorema 2.1 poderíamos reescrever a hipótese "f satisfaz (P1) – (P5)" por "f é monótona", desde que f seja definida por

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

onde  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  é um operador monótono.

## 2.3 Aplicações dos resultados de convergência de IPPEM

Vemos nesta seção que os resultados de convergência obtidos na Seção 2.2 são facilmente aplicados ao algoritmo IPPEM quando usado na resolução do problema  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , devido a monotonicidade de  $\mathcal{L}$  dada por (1.29). Ressaltamos que o desenvolvimento desta seção será equivalente ao feito por Iusem em [5], ou seja, trataremos o caso particular no qual a distância de Bregman é associada a função  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $g(x) = \frac{1}{2}\|x\|^2$ .

Começamos esta seção fazendo as seguintes considerações: no problema  $\text{PE}(f, \mathbb{K})$  assumimos, a partir daqui, que

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

e

$$\mathbb{K} = \{x \in \mathbb{R}^n; h_i(x) \leq 0\}$$

onde  $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  é convexa para cada  $i = 1, \dots, m$ . A função de Bregman  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  será dada por  $g(x) = \frac{1}{2}\|x\|^2$  e estará associada a distância de Bregman  $D_g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  definida por

$$D_g(x, y) = \frac{1}{2}\|x - y\|^2.$$

A expressão definida para  $g$  fará com que  $\nabla g \equiv I$ , logo nas fórmulas de regularização, assim como no algoritmo IPPEM, o gradiente de  $g$  será identificado com o operador identidade

$I : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Por fim, convecionamos, quando for necessário, para o vetor de erro que

$$e^k \equiv (e^k, 0) \in \mathbb{R}^n \times \mathbb{R}^m.$$

A primeira proposição mostra que a função Lagrangeana associada ao problema  $\text{PE}(f, \mathbb{K})$  é uma bifunção monótona. De fato,

**Proposição 2.2.** ([5], p.14) *Sejam  $f : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  uma bifunção monótona e  $\mathbb{K} = \{x \in \mathbb{R}^n; h_i(x) \leq 0\}$ , em que  $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  é convexa para cada  $i = 1, \dots, m$ . Então,  $\mathcal{L} : (\mathbb{R}^n \times \mathbb{R}^m) \times (\mathbb{R}^n \times \mathbb{R}^m) \rightarrow \mathbb{R}$  dada por*

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

é monótona.

*Demonstração.* Como  $f$  é monótona, vê-se facilmente que

$$\mathcal{L}((x, \lambda), (y, \mu)) + \mathcal{L}((y, \mu), (x, \lambda)) = f(x, y) + f(y, x) \leq 0.$$

■

Munidos da monotonicidade de  $\mathcal{L}$ , aplicamos o algoritmo IPPEM para resolver o problema  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . Primeiro, necessitamos exibir a forma regularizada inexata de  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  na iteração  $k$ . A fórmula de regularização inexata de  $f$ , dada pelo problema  $\text{PE}(f, \mathbb{K})$ , é uma bifunção  $\widehat{f} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  dada por

$$\widehat{f}_k^e(x, y) = f(x, y) + \gamma_k \langle \nabla g(x) - \nabla g(x^k), y - x \rangle - \langle e^k, y - x \rangle.$$

Assim, identificando  $f$  com  $\mathcal{L}$  e levando em conta as considerações feitas acima, temos que

$$\begin{aligned} \widehat{\mathcal{L}}_k^e((x, \lambda), (y, \mu)) &= \mathcal{L}((x, \lambda), (y, \mu)) + \gamma_k \langle (x, \lambda) - (x^k, \lambda^k), (y, \mu) - (x, \lambda) \rangle \\ &\quad - \langle (e^k, 0), (y, \mu) - (x, \lambda) \rangle \\ &= \mathcal{L}((x, \lambda), (y, \mu)) + \gamma_k \langle (x - x^k, \lambda - \lambda^k), (y - x, \mu - \lambda) \rangle \\ &\quad - \langle (e^k, 0), (y - x, \mu - \lambda) \rangle \\ &= \mathcal{L}((x, \lambda), (y, \mu)) + \gamma_k \langle x - x^k, y - x \rangle + \gamma_k \langle \lambda - \lambda^k, \mu - \lambda \rangle - \langle e^k, y - x \rangle. \end{aligned}$$

Logo, obtemos que a forma regularizada do problema  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  é o problema perturbado  $\text{PE}(\widehat{\mathcal{L}}_k^e, \mathbb{R}^n \times \mathbb{R}_+^m)$  onde

$$\widehat{\mathcal{L}}_k^e((x, \lambda), (y, \mu)) = \mathcal{L}((x, \lambda), (y, \mu)) + \gamma_k \langle x - x^k, y - x \rangle + \gamma_k \langle \lambda - \lambda^k, \mu - \lambda \rangle - \langle e^k, y - x \rangle. \quad (2.16)$$

Portanto, o Algoritmo IPPEM, com as devidas adaptações, tem o seguinte esquema quando aplicado ao problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ :

**Algoritmo IPPEM para  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ .** Considere uma sequência exógena limitada  $\{\gamma_k\} \subset \mathbb{R}_+$  de parâmetros de regularização e uma tolerância de erro relativo  $\sigma \in (0, 1)$ .

**Passo 1.** Iniciar o algoritmo com  $(\bar{x}^0, \bar{\lambda}^0) \in \mathbb{R}^n \times \mathbb{R}_+^m$ .

**Passo 2.** Dado  $(\bar{x}^k, \bar{\lambda}^k)$ , encontrar

$$(\hat{x}^k, \hat{\lambda}^k) \in \mathbb{R}^n \times \mathbb{R}^m$$

que resolve o problema  $PE(\hat{\mathcal{L}}_k^e, \mathbb{R}^n \times \mathbb{R}_+^m)$ , com  $\hat{\mathcal{L}}_k^e$  definida por (2.16), onde  $(e^k, 0) \in \mathbb{R}^n \times \mathbb{R}^m$  satisfaz a condição:

$$\|(e^k, 0)\| = \|e^k\| \leq \sigma \gamma_k \|(\hat{x}^k - \bar{x}^k, \hat{\lambda}^k - \bar{\lambda}^k)\|. \quad (2.17)$$

**Passo 3.** Se  $(\hat{x}^k, \hat{\lambda}^k) = (\bar{x}^k, \bar{\lambda}^k)$ , então PARE!. Caso contrário, faça

$$\bar{x}^{k+1} = \hat{x}^k - \frac{1}{\gamma_k} e^k \quad (2.18)$$

$$\bar{\lambda}^{k+1} = \hat{\lambda}^k, \quad (2.19)$$

e volte para o Passo 2.

Um resultado de convergência para o algoritmo descrito acima é dado a seguir.

**Corolário 2.1.** ([5], p.14) Considere o problema  $PE(f, \mathbb{K})$  com  $\mathbb{K}$  dado por (1.30) e  $f$  monótona. Tome  $\{\gamma_k\} \subset (0, \bar{\gamma}]$  para algum  $\bar{\gamma} > 0$ . Seja  $\{(x^k, \lambda^k)\}$  a sequência gerada pelo algoritmo IPPEM aplicado ao  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . Se o problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  tem soluções, então  $\{(x^k, \lambda^k)\}$  converge para algum par  $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ .

*Demonstração.* A Proposição 2.2 garante que  $\mathcal{L}$  é monótona. Assim, pela Observação 2.4, podemos aplicar o Teorema 2.1 ao problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , pois pela Proposição 1.3, a função de Bregman  $g(x) = \frac{1}{2}\|x\|^2$  satisfaz as condições do Teorema 2.1. Logo, a sequência  $\{(x^k, \lambda^k)\}$  converge para alguma solução  $(x^*, \lambda^*)$  de  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . ■



# Capítulo 3

## Análise de Convergência do Algoritmo IALEM

Neste capítulo apresentamos a análise de convergência do principal algoritmo desta dissertação, o chamado Método Extragradiante Lagrangeano Aumentado Inexato, ou abreviadamente IALEM, apresentado em [5]. Estudamos a convergência do algoritmo para este método através da análise de convergência feita para o algoritmo IPPEM no capítulo 2. Vemos que IPPEM aplicado ao problema  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  gera a mesma sequência determinada pelo algoritmo IALEM aplicado ao problema  $\text{PE}(f, \mathbb{K})$ , e concluímos que a convergência do algoritmo IPPEM a uma solução do problema  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  implicará na convergência do algoritmo IALEM a uma solução do problema  $\text{PE}(f, \mathbb{K})$ .

Assumimos que no problema de equilíbrio  $\text{PE}(f, \mathbb{K})$ , o conjunto  $\mathbb{K} \subset \mathbb{R}^n$  é não-vazio, fechado e convexo, definido por

$$\mathbb{K} = \{x \in \mathbb{R}^n; h_i(x) \leq 0\} \quad (3.1)$$

em que cada  $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  é convexa e satisfaz a condição de Slater para cada  $i = 1, \dots, m$ . Além disso,  $f$  é monótona e definida por

$$f(x, y) = \langle F(x), y - x \rangle \quad (3.2)$$

assim satisfazendo as propriedades (P1) – (P5) exigidas na funcionalidade do algoritmo IPPEM aplicado ao problema  $\text{PE}(f, \mathbb{K})$ .

### 3.1 Relações entre as soluções dos problemas $\text{PE}(f, \mathbb{K})$ e $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$

Nesta seção relacionamos as soluções dos problemas  $\text{PE}(f, \mathbb{K})$  e  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  aos pontos ótimos do problema  $\text{PE}(f, \mathbb{K})$ . Vemos que os pares ótimos do  $\text{PE}(f, \mathbb{K})$  são soluções

do  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , em particular, resolver este último problema é suficiente para obter uma solução do  $PE(f, \mathbb{K})$ . Assumimos nestes resultados que a função Lagrangeana  $\mathcal{L}$  dada em (1.29) não é necessariamente diferenciável.

**Proposição 3.1.** ([5], p.15) *Considere o problema  $PE(f, \mathbb{K})$ . As seguintes afirmações são equivalentes:*

(i)  $(x^*, \lambda^*) \in \mathbb{K} \times \mathbb{R}_+^m$  é um par ótimo para  $PE(f, \mathbb{K})$ ;

(ii)  $(x^*, \lambda^*)$  resolve o problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ .

*Demonstração.* (ii)  $\Rightarrow$  (i) Para cada  $(x^*, \lambda^*) \in \mathbb{K} \times \mathbb{R}_+^m$ , defina o problema de minimização restrito

$$\min \mathcal{L}((x^*, \lambda^*), (x, \lambda)) \tag{3.3}$$

$$\text{s.a. } (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_+^m. \tag{3.4}$$

Suponha que  $(x^*, \lambda^*)$  resolve o problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . Observe que  $(x^*, \lambda^*)$  é solução de (3.3) – (3.4), pois

$$\mathcal{L}((x^*, \lambda^*), (x, \lambda)) \geq 0 = \mathcal{L}((x^*, \lambda^*), (x^*, \lambda^*))$$

para todo  $(x, \lambda) \in \mathbb{K} \times \mathbb{R}_+^m$ . Sendo (3.4) determinado por restrições afins, a condição de Slater valerá e daí, pelo Teorema 1.6 (aplicado ao caso não diferenciável), existe um vetor de multiplicadores  $\eta^* = (\eta_1^*, \dots, \eta_m^*) \in \mathbb{R}^m$  associado a  $(x^*, \lambda^*)$  tal que

$$0 \in \partial_\lambda \left( \mathcal{L}((x^*, \lambda^*), (x, \lambda)) + \sum_{i=1}^m \eta_i^* (-\lambda_i) \right) \Big|_{(x, \lambda) = (x^*, \lambda^*)} \tag{3.5}$$

$$0 \in \partial_x \left( \mathcal{L}((x^*, \lambda^*), (x, \lambda)) + \sum_{i=1}^m \eta_i^* (-\lambda_i) \right) \Big|_{(x, \lambda) = (x^*, \lambda^*)} \tag{3.6}$$

$$-\lambda_i^* \leq 0 \tag{3.7}$$

$$\eta_i^* \geq 0 \tag{3.8}$$

$$\eta_i^* \lambda_i^* = 0 \tag{3.9}$$

Observe que (3.5) e (3.6) podem ser simplificadas. De fato, de (3.6) tem-se

$$0 \in \partial_x \mathcal{L}((x^*, \lambda^*), (x, \lambda)) \Big|_{(x, \lambda) = (x^*, \lambda^*)} + \partial_x \left( \sum_{i=1}^m \eta_i^* (-\lambda_i) \right) \Big|_{(x, \lambda) = (x^*, \lambda^*)}$$

o que implica em

$$0 \in \partial_x f(x^*, x) \Big|_{x=x^*} + \sum_{i=1}^m \lambda_i^* \partial_x h_i(x) \Big|_{x=x^*} - \partial_x \left( \sum_{i=1}^m \lambda_i h_i(x^*) \right) \Big|_{\lambda=\lambda^*}.$$

Sendo  $f(x^*, x) = \langle F(x^*), x - x^* \rangle$ , obtemos

$$0 \in F(x^*) + \sum_{i=1}^m \lambda_i^* \partial_x h_i(x^*).$$

Agora, veja que (3.5) nos fornece

$$0 \in \partial_\lambda f(x^*, x)|_{x=x^*} + \partial_\lambda \left( \sum_{i=1}^m \lambda_i^* h_i(x) \right) \Big|_{x=x^*} - \sum_{i=1}^m \partial_\lambda \lambda_i h_i(x^*) \Big|_{\lambda=\lambda^*} + \sum_{i=1}^m \eta_i^* \partial_\lambda (-\lambda_i) \Big|_{\lambda=\lambda^*}$$

Isto é,  $h_i(x^*) + \eta_i^* = 0$  para cada  $i = 1, \dots, m$ . Assim, podemos resumir (3.5) – (3.9) no seguinte sistema de equações

$$0 \in F(x^*) + \sum_{i=1}^m \lambda_i^* \partial h_i(x^*)$$

$$h_i(x^*) + \eta_i^* = 0$$

$$\lambda_i^* \geq 0$$

$$\eta_i^* \geq 0$$

$$\lambda_i^* \eta_i^* = 0$$

em que  $1 \leq i \leq m$  e  $F$  é dado em (3.2). Deste sistema concluimos, pela Definição 1.24, que  $(x^*, \lambda^*)$  é um par ótimo para o problema  $\text{PE}(f, \mathbb{K})$ .

(i)  $\Rightarrow$  (ii) Veja que o problema (3.3) – (3.4) é de minimização convexa. Com efeito, considere, para cada  $((x^1, \lambda^1), (x^2, \lambda^2)) \in (\mathbb{K} \times \mathbb{R}_+^m) \times (\mathbb{K} \times \mathbb{R}_+^m)$  e  $\beta \in [0, 1]$ , a expressão

$$M = \mathcal{L}((x^*, \lambda^*), (\beta x^1 + (1 - \beta)x^2, \beta \lambda^1 + (1 - \beta)\lambda^2)).$$

Temos que se  $f(x, \cdot)$  e  $h_i$ ,  $1 \leq i \leq m$ , são convexas, então

$$\begin{aligned} M &= f(x^*, \beta x^1 + (1 - \beta)x^2) + \sum_{i=1}^m \lambda_i^* h_i(\beta x^1 + (1 - \beta)x^2) - \sum_{i=1}^m (\beta \lambda_i^1 + (1 - \beta)\lambda_i^2) h_i(x^*) \\ &\leq \beta f(x^*, x^1) + (1 - \beta)f(x^*, x^2) + \beta \sum_{i=1}^m \lambda_i^* h_i(x^1) + (1 - \beta) \sum_{i=1}^m \lambda_i^* h_i(x^2) - \beta \sum_{i=1}^m \lambda_i^1 h_i(x^*) \\ &\quad - (1 - \beta) \sum_{i=1}^m \lambda_i^2 h_i(x^*) \\ &= \beta \left[ f(x^*, x^1) + \sum_{i=1}^m \lambda_i^* h_i(x^1) - \sum_{i=1}^m \lambda_i^1 h_i(x^*) \right] + (1 - \beta) \left[ f(x^*, x^2) + \sum_{i=1}^m \lambda_i^* h_i(x^2) \right. \\ &\quad \left. - \sum_{i=1}^m \lambda_i^2 h_i(x^*) \right] = \beta \mathcal{L}((x^*, \lambda^*), (x^1, \lambda^1)) + (1 - \beta) \mathcal{L}((x^*, \lambda^*), (x^2, \lambda^2)). \end{aligned}$$

Logo, a função objetivo de (3.3) – (3.4) é convexa. Por fim, note que a restrição  $-\lambda_i \leq 0$  deste problema é trivialmente convexa para todo  $i = 1, \dots, m$ . Portanto, as condições de KKT deste problema são suficientes. Como  $(x^*, \lambda^*)$  é um par ótimo para  $PE(f, \mathbb{K})$ , então  $(x^*, \lambda^*)$  é um ponto de KKT para (3.3) – (3.4), associado ao multiplicador

$$\eta^* = (-h_1(x^*), \dots, -h_m(x^*)),$$

que satisfaz (3.5) – (3.9). Sendo as condições de KKT deste problema suficientes, segue que  $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . ■

**Proposição 3.2.** ([5], p.15) *Considere o problema  $PE(f, \mathbb{K})$ . Se  $x^* \in S(f, \mathbb{K})$  e a condição (CS) dada na Seção 1.7 é válida para toda  $h_i$ ,  $1 \leq i \leq m$ , que define o conjunto viável  $\mathbb{K}$ , então existe  $\lambda^* \in \mathbb{R}_+^m$  tal que  $(x^*, \lambda^*)$  é um par ótimo para  $PE(f, \mathbb{K})$ . Reciprocamente, se  $(x^*, \lambda^*)$  é um par ótimo para  $PE(f, \mathbb{K})$ , então  $x^* \in S(f, \mathbb{K})$ .*

*Demonstração.* Como vale a condição (CS) então, pelo Teorema 1.6 (aplicado ao caso não diferenciável), existe um vetor  $\lambda^* \in \mathbb{R}^m$  associado a  $x^* \in S(f, \mathbb{K})$  tal que

$$\begin{aligned} 0 &\in F(x^*) + \sum_{i=1}^m \lambda_i^* \partial h_i(x^*) \\ \lambda_i^* &\geq 0 \\ h_i(x^*) &\leq 0 \\ \lambda_i^* h_i(x^*) &= 0 \end{aligned}$$

em que  $1 \leq i \leq m$  e  $F$  é dada por (3.2). Segue da Definição 1.24 que  $(x^*, \lambda^*)$  é um par ótimo para  $PE(f, \mathbb{K})$ .

Reciprocamente, se  $(x^*, \lambda^*)$  é um par ótimo para  $PE(f, \mathbb{K})$ , então (1.39) – (1.42) são satisfeitas, isto é,  $x^*$  é um ponto de KKT para o problema

$$\min_{x \in \mathbb{K}} f(x^*, x). \tag{3.10}$$

Como  $f(x^*, \cdot)$  é uma função convexa e  $\mathbb{K}$  é um conjunto convexo, então todo ponto de KKT de (3.10) é solução do respectivo problema. Assim, sendo

$$0 = f(x^*, x^*) = \min_{x \in \mathbb{K}} f(x^*, x),$$

isto é,  $f(x^*, x) \geq 0$  para todo  $x \in \mathbb{K}$ , conclui-se que  $x^* \in S(f, \mathbb{K})$ . ■

O resultado mais importante desta seção é dado a seguir, ele estabelece que resolver o problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  é suficiente para resolver o problema  $PE(f, \mathbb{K})$ .

**Corolário 3.1.** ([5], p.15) *Considere o problema  $PE(f, \mathbb{K})$ . Se  $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , então  $x^* \in S(f, \mathbb{K})$ . Reciprocamente, se  $x^* \in S(f, \mathbb{K})$  e vale a condição (CS), então existe um vetor  $\lambda^* \in \mathbb{R}_+^m$  tal que  $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ .*

*Demonstração.* Veja que  $(x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m) \Rightarrow (x^*, \lambda^*) \in \mathbb{K} \times \mathbb{R}_+^m$  é um par ótimo para o  $PE(f, \mathbb{K}) \Rightarrow x^* \in S(f, \mathbb{K})$ .

Reciprocamente, se vale a condição (CS) e  $x^* \in S(f, \mathbb{K})$ , então existe  $\lambda^* \in \mathbb{R}_+^m$  tal que  $(x^*, \lambda^*)$  é um par ótimo para  $PE(f, \mathbb{K}) \Rightarrow (x^*, \lambda^*) \in S(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . ■

## 3.2 O Método Extragradiante Lagrangeano Aumentado Inexato

Descrevemos nesta seção o algoritmo IALEM e apresentamos sua análise de convergência a partir dos resultados obtidos para o algoritmo IPPEM na Seção 2.2. Destacamos que a principal motivação ao desenvolvimento desta seção é o fato de que o Algoritmo IPPEM aplicado ao problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , onde  $\mathcal{L}$  é o Lagrangeano exato dado em (1.24), gera a mesma sequência determinada pelo Algoritmo IALEM aplicado ao problema  $PE(f, \mathbb{K})$ , quando  $f$  é monótona.

Iniciamos esta seção apresentando uma extensão a função Lagrangeana aumentada dada na Definição 1.21. Observamos que no desenvolvimento desta seção a função de Bregman  $g$  usada é da forma  $g(x) = \frac{1}{2}\|x\|^2$ .

**Definição 3.1.** *Considere o problema  $PE(f, \mathbb{K})$ . Tome um parâmetro de regularização  $\gamma = \frac{1}{2\alpha} > 0$ . Defina uma função  $s_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$  por*

$$s_i(x, y, \lambda, \gamma) = \frac{\gamma}{2} \left[ \left( \max \left\{ 0, \lambda_i + \frac{h_i(y)}{\gamma} \right\} \right)^2 - \left( \max \left\{ 0, \lambda_i + \frac{h_i(x)}{\gamma} \right\} \right)^2 \right]. \quad (3.11)$$

A função  $\tilde{\mathcal{L}} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$  definida por

$$\tilde{\mathcal{L}}_\gamma(x, y, \lambda, z) = f(x, y) + \gamma \langle x - z, y - x \rangle + \sum_{i=1}^m s_i(x, y, \lambda, \gamma) \quad (3.12)$$

é denominada função Lagrangeana aumentada proximal para o problema  $PE(f, \mathbb{K})$ .

Convecionamos, a partir daqui, que a função Lagrangeana  $L$  dada depende diretamente do parâmetro  $\alpha$ , e em vez de escrevermos  $L_\alpha(x, \lambda)$ , colocar-se-á  $L(x, \lambda, \alpha)$ . Assim, em (3.12) escrevemos  $\tilde{\mathcal{L}}(x, y, \lambda, z, \gamma)$  em vez de  $\tilde{\mathcal{L}}_\gamma(x, y, \lambda, z)$ .

Observe que (3.12) é a forma regularizada exata do Lagrangeano aumentado associado ao problema  $PE(f, \mathbb{K})$ . De fato, uma Lagrangeana aumentada para  $PE(f, \mathbb{K})$ , com parâmetro

$\theta > 0$ , é uma função  $\mathcal{L}' : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$  dada por

$$\begin{aligned} \mathcal{L}'(x, y, \lambda, \theta) &= f(x, y) + \frac{1}{4\theta} \sum_{i=1}^m [(\max\{0, \lambda_i + 2\theta h_i(y)\})^2 - \lambda_i^2] \\ &\quad - \frac{1}{4\theta} \sum_{i=1}^m [(\max\{0, \lambda_i + 2\theta h_i(x)\})^2 - \lambda_i^2] \\ &= f(x, y) + \frac{1}{4\theta} \sum_{i=1}^m [(\max\{0, \lambda_i + 2\theta h_i(y)\})^2 - (\max\{0, \lambda_i + 2\theta h_i(x)\})^2]. \end{aligned}$$

Fixando um parâmetro de regularização  $\gamma > 0$ , a forma regularizada exata de  $\mathcal{L}'$  torna-se uma função  $\tilde{\mathcal{L}}$  definida por

$$\tilde{\mathcal{L}}(x, y, \lambda, z, \theta) = \mathcal{L}'(x, y, \lambda, \theta) + \gamma \langle x - z, y - x \rangle.$$

Assumindo que  $\gamma = \frac{1}{2\theta}$ , obtemos

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

No algoritmo IALEM, assim como fora feito para o algoritmo IPPEM, usa-se a forma inexata da regularização (3.12) para descrevê-lo. Considere  $e \in \mathbb{R}^n$  o vetor de erro. A forma inexata de (3.12) é dada por

$$\tilde{\mathcal{L}}^e(x, y, \lambda, z, \gamma) = f(x, y) + \gamma \langle x - z, y - x \rangle + \sum_{i=1}^m s_i(x, y, \lambda, \gamma) - \langle e, y - x \rangle \quad (3.13)$$

onde  $\gamma > 0$ . A partir de (3.13), podemos descrever o algoritmo IALEM aplicado ao problema  $PE(f, \mathbb{K})$ .

**Algoritmo IALEM para  $PE(f, \mathbb{K})$ .** Considere uma sequência exógena limitada  $\{\gamma_k\} \subset \mathbb{R}_+$  de parâmetros de regularização e uma tolerância de erro relativo  $\sigma \in (0, 1)$ .

**Passo 1.** Iniciar o algoritmo com  $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}_+^m$ .

**Passo 2.** Dado  $(x^k, \lambda^k)$ , encontrar  $\tilde{x}^k \in \mathbb{R}^n$  que resolve o problema perturbado  $PE(\tilde{\mathcal{L}}_k^e, \mathbb{R}^n)$ , em que  $\tilde{\mathcal{L}}_k^e$  é dada por

$$\tilde{\mathcal{L}}_k^e(x, y, \lambda^k, x^k, \gamma_k) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle + \sum_{i=1}^m s_i(x, y, \lambda^k, \gamma_k) - \langle e^k, y - x \rangle \quad (3.14)$$

com  $s_i$  dado por (3.11), e  $e^k \in \mathbb{R}^n$  que satisfaz a condição:

$$\|e^k\| \leq \sigma \gamma_k \|(\tilde{x}^k - x^k, \lambda^{k+1} - \lambda^k)\| \quad (3.15)$$

onde  $\lambda^{k+1} = (\lambda_1^{k+1}, \dots, \lambda_m^{k+1})$  é atualizado no próximo passo.

**Passo 3.** Definir a próxima iteração de  $\lambda^k$  como

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k} \right\} \quad (3.16)$$

para cada  $i = 1, \dots, m$ .

**Passo 4.** Se  $(x^k, \lambda^k) = (\tilde{x}^k, \lambda^{k+1})$ , então PARE!. Caso contrário, faça

$$x^{k+1} = \tilde{x}^k - \frac{1}{\gamma_k} e^k, \quad (3.17)$$

e volte para o Passo 2.

**Observação 3.1.** Em vez de escrevermos  $\tilde{\mathcal{L}}_k^e(x, y, \lambda^k, x^k, \gamma_k)$ , colocamos apenas  $\tilde{\mathcal{L}}_k^e(x, y)$ , por questão apenas de simplicidade.

Um dos principais resultados desta dissertação é o resultado de convergência para o algoritmo IALEM dado no Teorema 3.1. Antes de enunciá-lo, precisamos de um importante resultado de existência e unicidade de solução à forma regularizada inexata do problema  $PE(f, \mathbb{K})$ . Tal resultado é descrito a seguir.

**Lema 3.1.** Considere o problema  $PE(f, \mathbb{K})$ . Suponha que  $f$  é monótona. Fixe um vetor de erro  $e \in \mathbb{R}^n$ , um vetor  $z \in \mathbb{R}^n$  arbitrário e um parâmetro de regularização  $\gamma > 0$ . Se  $\tilde{f} : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$  é definida como

$$\tilde{f}(x, y) = f(x, y) + \gamma \langle x - z, y - x \rangle - \langle e, y - x \rangle$$

então  $PE(\tilde{f}, \mathbb{K})$  tem uma única solução.

*Demonstração.* Vamos provar inicialmente que  $\tilde{f}$  é monótona e satisfaz as propriedades (P1) – (P3) e (P6) dadas na Seção 1.4 e, em seguida, aplicar a Proposição 1.5 para garantir a existência de solução para o problema  $PE(\tilde{f}, \mathbb{K})$ . Observe que, para todo  $(x, y) \in \mathbb{K} \times \mathbb{K}$ , a monotonicidade de  $f$  implica que

$$\begin{aligned} \tilde{f}(x, y) + \tilde{f}(y, x) &= f(x, y) + f(y, x) + \gamma[\langle x - z, y - x \rangle + \langle y - z, x - y \rangle] - \langle e, y - x \rangle \\ &\quad - \langle e, x - y \rangle \\ &= f(x, y) + f(y, x) + \gamma \langle x - y, y - x \rangle \\ &\leq -\gamma \|x - y\|^2 \leq 0. \end{aligned}$$

Isto é,  $\tilde{f}$  é monótona. Afirmamos que  $\tilde{f}$  satisfaz (P1) – (P3) e (P6). Com efeito,

(P1) É imediatamente verificada;

(P2) Para todo  $x_0 \in \mathbb{K}$  fixo, e toda sequência  $\{y^k\} \subset \mathbb{K}$  tal que  $\{y^k\} \rightarrow \bar{y} \in \mathbb{K}$ , tem-se

$$\begin{aligned} \liminf_{k \rightarrow \infty} \tilde{f}(x_0, y^k) &= \liminf_{k \rightarrow \infty} [f(x_0, y^k) + \gamma \langle x_0 - z, y^k - x_0 \rangle - \langle e, y^k - x_0 \rangle] \\ &\geq f(x_0, \bar{y}) + \gamma \langle x_0 - z, \bar{y} - x_0 \rangle - \langle e, \bar{y} - x_0 \rangle = \tilde{f}(x_0, \bar{y}). \end{aligned}$$

Assim,  $\tilde{f}(x, \cdot)$  é semicontínua inferiormente para todo  $x \in \mathbb{K}$ . Além disso, considerando a função  $p : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $p(y) = \gamma \langle x - z, y - x \rangle - \langle e, y - x \rangle$ , temos que  $p$  é afim e como  $f(x, \cdot)$  é convexa, segue que  $\tilde{f}(x, \cdot)$  é convexa.

(P3) Para todo  $y_0 \in \mathbb{K}$  fixo, e toda sequência  $\{x^k\} \subset \mathbb{K}$  tal que  $\{x^k\} \rightarrow \bar{x} \in \mathbb{K}$ , tem-se

$$\begin{aligned} \limsup_{k \rightarrow \infty} \tilde{f}(x^k, y_0) &= \limsup_{k \rightarrow \infty} [f(x^k, y_0) + \gamma \langle x^k - z, y_0 - x^k \rangle - \langle e, y_0 - x^k \rangle] \\ &\leq f(\bar{x}, y_0) + \gamma \langle \bar{x} - z, y_0 - \bar{x} \rangle - \langle e, y_0 - \bar{x} \rangle = \tilde{f}(\bar{x}, y_0). \end{aligned}$$

Assim,  $\tilde{f}(\cdot, y)$  é semicontínua superiormente para todo  $y \in \mathbb{K}$ .

(P6) Tome uma sequência  $\{x^k\} \subseteq \mathbb{K}$ , para todo  $k \in \mathbb{N}$ , tal que  $\lim_{k \rightarrow \infty} \|x^k\| = +\infty$ . Afirmamos que esta propriedade vale para cada  $u$  que é a projeção de Bregman de  $z$  sobre  $\mathbb{K}$ . De fato, note que

$$\begin{aligned} \tilde{f}(x^k, u) &= f(x^k, u) - \gamma \langle x^k - z, x^k - u \rangle - \langle e, u - x^k \rangle \\ &= f(x^k, u) - \gamma [\langle x^k, x^k \rangle - \langle x^k, u \rangle - \langle z, x^k \rangle + \langle z, u \rangle] - \langle e, u - x^k \rangle \\ &= f(x^k, u) - \gamma [(\langle x^k, x^k \rangle - \langle x^k, u \rangle - \langle z, x^k \rangle + \langle z, u \rangle) - (\langle u, u \rangle - \langle u, x^k \rangle) \\ &\quad - (\langle u, x^k \rangle - \langle u, u \rangle)] - \langle e, u - x^k \rangle \\ &= f(x^k, u) - \gamma [\langle u - z, x^k - u \rangle - \langle x^k - u, x^k - u \rangle] - \langle e, u - x^k \rangle \\ &\leq -f(u, x^k) + \gamma \langle z - u, x^k - u \rangle - \gamma \|x^k - u\|^2 - \langle e, u - x^k \rangle \\ &\leq -f(u, x^k) - \gamma \|x^k - u\|^2 - \langle e, u - x^k \rangle. \end{aligned}$$

Agora observe que sendo  $f(x, \cdot)$  convexa, o conjunto  $\partial_y f(x, \cdot)$  é não vazio. Além disso,  $\mathbb{K}$  convexo e fechado implica que  $\text{int } \mathbb{K}$  é não vazio. Assim, para  $w \in \text{int } \mathbb{K}$  e  $v \in \partial_y f(u, w)$ , tem-se

$$f(u, y) \geq f(u, w) + \langle v, y - w \rangle, \quad \forall y \in \mathbb{K}.$$

Como  $\{x^k\} \subset \mathbb{K}$ , podemos tomar  $y = x^k \in \mathbb{K}$ , obtendo assim

$$f(u, x^k) - f(u, w) \geq \langle v, x^k - w \rangle.$$

Daí, um limitante superior para  $-f(u, x^k)$  é

$$\begin{aligned} -f(u, x^k) &\leq \langle v, w - x^k \rangle - f(u, w) \\ &\leq \|v\| \|w - x^k\| - f(u, w) \\ &= \|v\| \|(w - u) + (u - x^k)\| - f(u, w) \\ &\leq \|v\| \|w - u\| + \|v\| \|u - x^k\| - f(u, w). \end{aligned}$$



Prosseguindo, obtemos um limitante superior para o termo  $-\|x^k - u\|^2$ . Observamos que

$$\begin{aligned}\|x^k - u\|^2 &= 2 \left( \frac{1}{2} \|x^k\|^2 - \langle x^k, u \rangle + \frac{1}{2} \|u\|^2 \right) \\ &\geq \frac{1}{2} \|x^k\|^2 - \langle x^k, u \rangle - \frac{1}{2} \|u\|^2 + \langle u, u \rangle \\ &= \frac{1}{2} \|x^k\|^2 - \frac{1}{2} \|u\|^2 - \langle u, x^k - u \rangle \\ &\geq \frac{1}{2} \|x^k\|^2 - \frac{1}{2} \|u\|^2 - \|u\| \|x^k - u\|.\end{aligned}$$

Assim, com os limitantes acima obtidos, verificamos que

$$\begin{aligned}\tilde{f}(x^k, u) &\leq \|v\|(\|w - u\| + \|x^k - u\|) - f(u, w) - \gamma \left( \frac{1}{2} \|x^k\|^2 - \frac{1}{2} \|u\|^2 - \|u\| \|x^k - u\| \right) \\ &\quad + \|e\| \|x^k - u\| \\ &= -\gamma \left[ \frac{1}{2} \|x^k\|^2 - \frac{\|v\| + \|e\| + \gamma \|u\|}{\gamma} \|x^k - u\| \right] + \frac{\gamma}{2} \|u\|^2 \|v\| \|w - u\| - f(u, w).\end{aligned}$$

Como  $\lim_{k \rightarrow \infty} \|x^k\| = +\infty$  e  $\frac{\|v\| + \|e\| + \gamma \|u\|}{\gamma} > 0$ , então por (H5) dada na Proposição 1.3, segue que  $\lim_{k \rightarrow \infty} \tilde{f}(x^k, u) \leq -\infty$ , isto é,

$$\lim_{k \rightarrow \infty} \tilde{f}(x^k, u) = -\infty.$$

Logo, para algum  $k_0 \in \mathbb{N}$ , tem-se  $\tilde{f}(x^k, u) \leq 0$  para todo  $k \geq k_0$ . Portanto, pela Proposição 1.5, segue que  $\text{PE}(\tilde{f}, \mathbb{K})$  tem soluções.

Para finalizarmos a demonstração deste lema, vamos provar a unicidade da solução do problema  $\text{PE}(\tilde{f}, \mathbb{K})$ . Suponha que  $\tilde{x}'$  e  $\tilde{x}''$  sejam duas soluções de  $\text{PE}(\tilde{f}, \mathbb{K})$ . Temos que:

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

Como  $f$  é monótona e  $\gamma > 0$ , então

$$\begin{aligned}0 \leq \tilde{f}(\tilde{x}', \tilde{x}'') + \tilde{f}(\tilde{x}'', \tilde{x}') &= f(\tilde{x}', \tilde{x}'') + f(\tilde{x}'', \tilde{x}') - \gamma [\|\tilde{x}'\|^2 - 2 \langle \tilde{x}', \tilde{x}'' \rangle + \|\tilde{x}''\|^2] \\ &= f(\tilde{x}', \tilde{x}'') + f(\tilde{x}'', \tilde{x}') - \gamma \|\tilde{x}' - \tilde{x}''\|^2 \\ &\leq -\gamma \|\tilde{x}' - \tilde{x}''\|^2 \leq 0.\end{aligned}$$

Segue daí que  $\tilde{x}' = \tilde{x}''$  e, portanto o problema  $\text{PE}(\tilde{f}, \mathbb{K})$  possui uma única solução. ■

**Observação 3.2.** Seja  $\mathcal{L}'_k : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$  dada por

$$\mathcal{L}'_k(x, y, \lambda^k, \gamma_k) = f(x, y) + \sum_{i=1}^m s_i(x, y, \lambda^k, \gamma_k).$$

Considere o problema  $PE(\mathcal{L}'_k, \mathbb{R}^n)$ . Se  $\mathcal{L}'_k$  é monótona, então o problema  $PE(\tilde{\mathcal{L}}_k^e, \mathbb{R}^n)$ , com  $\tilde{\mathcal{L}}_k^e$  dada por

$$\tilde{\mathcal{L}}_k^e(x, y) = \mathcal{L}'_k(x, y, \lambda^k, \gamma_k) + \gamma_k \langle x - x^k, y - x \rangle - \langle e^k, y - x \rangle,$$

tem uma única solução. Em outras palavras, o Passo 2 de IALEM gera uma sequência  $\{(\tilde{x}^k, e^k)\}$  que está bem definida para toda iteração  $k$ . Com efeito, para quaisquer  $(x, y) \in \mathbb{K} \times \mathbb{K}$  verifica-se facilmente que

$$\mathcal{L}'_k(x, y, \lambda^k, \gamma_k) + \mathcal{L}'_k(y, x, \lambda^k, \gamma_k) \leq 0, \quad \forall k \in \mathbb{N}.$$

Logo, aplicando o Lema 3.1, obtém-se o resultado desejado.

O Teorema 3.1, dado a seguir, estabelece que a sequência gerada pelo algoritmo IALEM, aplicado ao problema  $PE(f, \mathbb{K})$ , com  $f$  monótona, coincide com a sequência gerada pelo algoritmo IPPEM, aplicado ao problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , onde  $\mathcal{L}$  é a função Lagrangeana associada a  $PE(f, \mathbb{K})$ .

**Teorema 3.1.** ([5]) Considere o problema  $PE(f, \mathbb{K})$ . Assuma que  $f$  é monótona. Fixe uma sequência  $\{\gamma_k\} \subset \mathbb{R}_+$  e uma tolerância de erro relativo  $\sigma \in (0, 1)$ . Sejam  $\{(x^k, \lambda^k)\}$  a sequência gerada pelo algoritmo IALEM aplicado ao problema  $PE(f, \mathbb{K})$ , com vetor de erro  $e^k \in \mathbb{R}^n$  associado, e  $\{\bar{x}^k, \bar{\lambda}^k\}$  a sequência gerada pelo algoritmo IPPEM aplicado ao problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , com vetor de erro  $(e^k, 0) \in \mathbb{R}^n \times \mathbb{R}^m$  associado. Suponha que o parâmetro de regularização  $\gamma$  e o erro relativo  $\sigma$  sejam os mesmos em ambos os algoritmos. Se  $(x^0, \lambda^0) = (\bar{x}^0, \bar{\lambda}^0)$ , então  $(x^k, \lambda^k) = (\bar{x}^k, \bar{\lambda}^k)$  para todo  $k \in \mathbb{N}$ .

*Demonstração.* Vamos provar o teorema por indução sobre  $k$ . Por hipótese, o resultado é válido para  $k = 0$ . Suponhamos que o resultado vale para  $k \geq 0$ . Seja  $(\hat{x}^k, \hat{\lambda}^k)$  a solução do problema  $PE(\hat{\mathcal{L}}_k^e, \mathbb{R}^n \times \mathbb{R}_+^m)$ , com  $\hat{\mathcal{L}}_k^e$  definida por (2.16). Pelo Lema 3.1 esta solução existe e é única, devido a monotonicidade de  $\mathcal{L}$ . Além disso, pela Proposição 3.1,  $(\hat{x}^k, \hat{\lambda}^k)$  é uma solução do problema de minimização convexa

$$\min \hat{\mathcal{L}}_k^e((\hat{x}^k, \hat{\lambda}^k), (x, \lambda)) \tag{3.18}$$

$$\text{s.a. } (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_+^m \tag{3.19}$$

Como (3.18)–(3.19) apresenta apenas restrições afins, então a condição de Slater é satisfeita,

assim existe um vetor  $\eta^k \in \mathbb{R}^m$  de KKT tal que

$$0 \in \partial_x \left[ \widehat{\mathcal{L}}_k^e((\widehat{x}^k, \widehat{\lambda}^k), (x, \lambda)) + \sum_{i=1}^m \eta_i^k(-\lambda_i) \right] \quad (3.20)$$

$$0 \in \partial_\lambda \left[ \widehat{\mathcal{L}}_k^e((\widehat{x}^k, \widehat{\lambda}^k), (x, \lambda)) + \sum_{i=1}^m \eta_i^k(-\lambda_i) \right] \quad (3.21)$$

$$\widehat{\lambda}_i^k \geq 0$$

$$\eta_i^k \geq 0$$

$$\widehat{\lambda}_i^k \eta_i^k = 0.$$

em que  $1 \leq i \leq m$ . A partir de (3.20) tem-se que

$$\begin{aligned} 0 \in & \partial_x \mathcal{L}_k((\widehat{x}^k, \widehat{\lambda}^k), (x, \lambda)) + \gamma_k \left[ \partial_x \langle \widehat{x}^k - x^k, x - \widehat{x}^k \rangle + \partial_x \langle \widehat{\lambda}^k - \lambda^k, \lambda - \widehat{\lambda}^k \rangle \right] - \partial_x \langle e^k, x - \widehat{x}^k \rangle \\ & + \partial_x \left[ \sum_{i=1}^m \eta_i^k(-\lambda_i) \right] \end{aligned} \quad (3.22)$$

$$\in F(\widehat{x}^k) + \sum_{i=1}^m \widehat{\lambda}_i^k \partial h_i(x) + \gamma_k(\widehat{x}^k - x^k) - e^k \quad (3.23)$$

onde (3.23), no ponto  $(x, \lambda) = (\widehat{x}^k, \widehat{\lambda}^k)$ , é equivalente a

$$\gamma_k(x^k - \widehat{x}^k) + e^k \in F(\widehat{x}^k) + \sum_{i=1}^m \widehat{\lambda}_i^k \partial h_i(\widehat{x}^k).$$

E de (3.21), vê-se que

$$\begin{aligned} 0 \in & \partial_\lambda \mathcal{L}_k((\widehat{x}^k, \widehat{\lambda}^k), (x, \lambda)) + \gamma_k [\partial_\lambda \langle \widehat{x}^k - x^k, x - \widehat{x}^k \rangle + \partial_\lambda \langle \widehat{\lambda}^k - \lambda^k, \lambda - \widehat{\lambda}^k \rangle] - \partial_\lambda \langle e^k, x - \widehat{x}^k \rangle \\ & + \sum_{i=1}^m \eta_i^k \partial_\lambda (-\lambda_i) = (h_1(\widehat{x}^k), \dots, h_m(\widehat{x}^k)) + \gamma_k(\widehat{\lambda}^k - \lambda^k) - (\eta_1^k, \dots, \eta_m^k), \end{aligned}$$

isto é,

$$\eta_i^k = -h_i(\widehat{x}^k) + \gamma_k(\widehat{\lambda}_i^k - \lambda_i^k)$$

para cada  $i = 1, \dots, m$ . Assim, usando a hipótese de indução e as equivalências acima encontradas, obtemos o sistema

$$\gamma_k(\widehat{x}^k - \widehat{x}^k) + e^k \in F(\widehat{x}^k) + \sum_{i=1}^m \widehat{\lambda}_i^k \partial h_i(\widehat{x}^k) \quad (3.24)$$

$$-h_i(\widehat{x}^k) + \gamma_k(\widehat{\lambda}_i^k - \lambda_i^k) = \eta_i^k \quad (3.25)$$

$$\widehat{\lambda}_i^k \geq 0 \quad (3.26)$$

$$\eta_i^k \geq 0 \quad (3.27)$$

$$\widehat{\lambda}_i^k \eta_i^k = 0 \quad (3.28)$$

em que  $F$  é dada por (3.2) e  $1 \leq i \leq m$ . Podemos simplificar o sistema (3.24) – (3.28) de modo que seja expresso por única relação de pertinência. Com efeito, por (3.25) e (3.27), temos que

$$-h_i(\widehat{x}^k) + \gamma_k(\widehat{\lambda}_i^k - \bar{\lambda}_i^k) \geq 0$$

isto é,  $\widehat{\lambda}_i^k \geq \frac{1}{\gamma_k} h_i(\widehat{x}^k) + \bar{\lambda}_i^k$ . Veja por (3.25) e (3.28) que se  $\widehat{\lambda}_i^k > 0$ , então  $\eta_i^k = 0$  e assim

$$\widehat{\lambda}_i^k = \frac{1}{\gamma_k} h_i(\widehat{x}^k) + \bar{\lambda}_i^k.$$

E que se  $\eta_i^k > 0$ , teremos  $\widehat{\lambda}_i^k = 0$ . Logo,

$$\widehat{\lambda}_i^k = \max \left\{ 0, \bar{\lambda}_i^k + \frac{h_i(\widehat{x}^k)}{\gamma_k} \right\} \quad (1 \leq i \leq m).$$

Portanto, (3.24) – (3.28) pode ser reescrito como

$$\gamma_k(\bar{x}^k - \widehat{x}^k) + e^k \in F(\widehat{x}^k) + \sum_{i=1}^m \widehat{\lambda}_i^k \partial h_i(\widehat{x}^k) \quad (3.29)$$

$$\widehat{\lambda}_i^k = \max \left\{ 0, \bar{\lambda}_i^k + \frac{h_i(\widehat{x}^k)}{\gamma_k} \right\} \quad (1 \leq i \leq m). \quad (3.30)$$

Se substituirmos (3.30) em (3.29), ficamos finalmente com uma única expressão

$$\gamma_k(\bar{x}^k - \widehat{x}^k) + e^k \in F(\widehat{x}^k) + \left[ \sum_{i=1}^m \left( \max \left\{ 0, \bar{\lambda}_i^k + \frac{h_i(\widehat{x}^k)}{\gamma_k} \right\} \right) \partial h_i(\widehat{x}^k) \right] \quad (3.31)$$

Vamos obter agora uma expressão análoga a (3.31) através do Passo 2 do algoritmo IALEM. Seja  $\tilde{x}^k \in \mathbb{R}^n$  a solução do problema PE( $\tilde{\mathcal{L}}_k^e, \mathbb{R}^n$ ) obtida no Passo 2 de IALEM, em que  $\tilde{\mathcal{L}}_k^e$  é dada por (3.14). Observamos que se  $\tilde{x}^k$  é solução deste problema, isto equivale a  $\tilde{x}^k$  ser solução do problema de minimização convexa

$$\min_{y \in \mathbb{R}^n} \tilde{\mathcal{L}}_k^e(\tilde{x}^k, y)$$

De fato,  $\tilde{\mathcal{L}}_k^e(\tilde{x}^k, \tilde{x}^k) = 0 \leq \tilde{\mathcal{L}}_k^e(\tilde{x}^k, y)$  para todo  $y \in \mathbb{R}^n$ . Assim,  $\tilde{x}^k \in S(\tilde{\mathcal{L}}_k^e, \mathbb{R}^n)$  se, e somente se,

$$0 \in \partial_y \tilde{\mathcal{L}}_k^e(\tilde{x}^k, \tilde{x}^k) \quad (3.32)$$

isto é,

$$0 \in \left. \partial_y f(\tilde{x}^k, y)|_{y=\tilde{x}^k} + \gamma_k \partial_y (\langle \tilde{x}^k - x^k, y - \tilde{x}^k \rangle)|_{y=\tilde{x}^k} + \frac{\gamma_k}{2} \sum_{i=1}^m \left[ \partial_y \left( \max \left\{ 0, \lambda_i^k + \frac{h_i(y)}{\gamma_k} \right\} \right)^2 \right] \right|_{y=\tilde{x}^k} - \left. \partial_y \left( \max \left\{ 0, \lambda_i + \frac{h_i(\tilde{x}^k)}{\gamma} \right\} \right)^2 \right] \right|_{y=\tilde{x}^k} - \partial_y (\langle e^k, y - \tilde{x}^k \rangle)|_{y=\tilde{x}^k}$$

ou seja,

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

Logo, (3.32) é equivalente a

$$\gamma_k(x^k - \tilde{x}^k) + e^k \in F(\tilde{x}^k) + \left[ \sum_{i=1}^m \left( \max \left\{ 0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k} \right\} \right) \partial h_i(\tilde{x}^k) \right]. \quad (3.33)$$

Aplicando a hipótese de indução em (3.33), concluímos de (3.31) que  $\hat{x}^k$  também satisfaz (3.33). Como a condição de otimalidade de (3.32) é suficiente (por ser um problema convexo), então  $\hat{x}^k \in S(\tilde{\mathcal{L}}_k^e, \mathbb{R}^n)$ . Além disso, pelo Lema 3.1,  $\text{PE}(\tilde{\mathcal{L}}_k^e, \mathbb{R}^n)$  tem uma única solução (conforme Observação 3.2). Logo,

$$\hat{x}^k = \tilde{x}^k$$

e daí, através das atualizações (3.17) e (2.18), dadas no Passo 4 de IALEM e Passo 3 de IPPEM (aplicado ao problema  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ ), respectivamente, concluímos que

$$x^{k+1} = \bar{x}^{k+1}.$$

Por fim, a atualização das variáveis duais (2.19) do algoritmo IPPEM, aplicado ao problema  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , e (3.30) nos fornecem

$$\bar{\lambda}_i^{k+1} = \hat{\lambda}_i^k = \max \left\{ 0, \bar{\lambda}_i^k + \frac{h_i(\hat{x}^k)}{\gamma_k} \right\} \quad (1 \leq i \leq m). \quad (3.34)$$

Comparando (3.34) com a atualização (3.16) das variáveis duais no Passo 3 de IALEM, e usando o fato de que  $\hat{x}^k = \tilde{x}^k$  e  $\lambda_i^k = \bar{\lambda}_i^k$ , finalmente obtemos

$$\lambda^{k+1} = \bar{\lambda}^{k+1}.$$

Segue daí que  $(x^k, \lambda^k) = (\bar{x}^k, \bar{\lambda}^k)$  para todo  $k \in \mathbb{N}$ , como queríamos demonstrar. ■

O segundo principal resultado deste trabalho garante que se o problema  $\text{PE}(f, \mathbb{K})$  tem soluções, então a sequência gerada pelo algoritmo IALEM para resolvê-lo converge para

algum dos seus pares ótimos, em particular, há convergência para uma solução do problema  $PE(f, \mathbb{K})$ . De maneira precisa, tal resultado é assim enunciado:

**Teorema 3.2.** ([5]) *Considere o problema  $PE(f, \mathbb{K})$ . Assuma que*

- (i)  *$f$  é monótona;*
- (ii)  *$\mathbb{K} = \{x \in \mathbb{R}^n; h_i(x) \leq 0\}$ ;*
- (iii) *a condição de qualificação (CS) da restrição  $h_i$  vale para cada  $i = 1, \dots, m$  que define  $\mathbb{K}$ ;*
- (iv)  *$\{\gamma_k\} \subset (0, \bar{\gamma}]$  para algum  $\bar{\gamma} > 0$ .*

*Seja  $\{(x^k, \lambda^k)\}$  a sequência gerada pelo algoritmo IALEM. Se o problema  $PE(f, \mathbb{K})$  tem soluções, então a sequência  $\{(x^k, \lambda^k)\}$  converge para algum par ótimo  $(x^*, \lambda^*)$  para  $PE(f, \mathbb{K})$ , e conseqüentemente  $x^* \in S(f, \mathbb{K})$ .*

*Demonstração.* Pelo Teorema 3.1, a sequência  $\{(x^k, \lambda^k)\}$  coincide com a sequência gerada por IPPEM aplicada ao  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . Como  $PE(f, \mathbb{K})$  tem soluções e vale a condição de Slater então, pelo Corolário 3.1,  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  tem soluções. Assim, pelo Corolário 2.1, a sequência  $\{(x^k, \lambda^k)\}$  converge a uma solução  $(x^*, \lambda^*)$  do  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ . Aplicando novamente o Corolário 3.1, conclui-se que  $x^* \in S(f, \mathbb{K})$ . ■

Na análise de convergência de qualquer algoritmo, uma característica importante que deve ser explorada é a questão da terminação finita do algoritmo. É natural esperar que se o algoritmo para num número finito de interações  $k$ , então o iterando  $x^k$ , da sequência determinada por este algoritmo, é uma solução do problema em que o algoritmo fora aplicado. Isto também se aplica ao algoritmo IALEM quando usado para resolver o problema  $PE(f, \mathbb{K})$ . De fato, temos a seguinte proposição a respeito da terminação finita do algoritmo IALEM.

**Proposição 3.3.** ([5]) *Suponha que o Algoritmo IALEM pare na iteração  $k$ . Então, o vetor  $x^k \in \mathbb{R}^n$  gerado por este algoritmo é uma solução do problema  $PE(f, \mathbb{K})$ .*

*Demonstração.* Se o algoritmo IALEM termina na  $k$ -ésima iteração, então pelo Passo 4,  $(x^k, \lambda^k) = (\tilde{x}^k, \lambda^{k+1})$ . Daí, por (3.15), obtemos  $e^k = 0$ . Agora, veja que se  $e^k = 0$  em (3.14), então

$$\tilde{\mathcal{L}}_k^0(x, y) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle + \sum_{i=1}^m s_i(x, y, \lambda^k, \gamma_k).$$

Como  $\tilde{x}^k = x^k$ , temos que

$$\tilde{\mathcal{L}}_k^0(\tilde{x}^k, y) = f(\tilde{x}^k, y) + \sum_{i=1}^m s_i(\tilde{x}^k, y, \lambda^k, \gamma_k).$$

Sendo a função  $\tilde{\mathcal{L}}_k^0(\tilde{x}^k, \cdot)$  convexa, então o problema de minimização irrestrito

$$\min_{y \in \mathbb{R}^n} \tilde{\mathcal{L}}_k^0(\tilde{x}^k, y)$$

é convexo. Além disso,  $\tilde{x}^k$  é uma solução deste problema. De fato, sendo  $\tilde{x}^k$  uma solução do problema  $\text{PE}(\tilde{\mathcal{L}}_k^0, \mathbb{R}^n)$ , tem-se

$$0 = \tilde{\mathcal{L}}_k^0(\tilde{x}^k, \tilde{x}^k) \leq \tilde{\mathcal{L}}_k^0(\tilde{x}^k, y), \quad \forall y \in \mathbb{K}.$$

Assim, pelo Teorema 1.2,

$$0 \in \partial \tilde{\mathcal{L}}_k^0(\tilde{x}^k, \tilde{x}^k) = F(\tilde{x}^k) + \left[ \sum_{i=1}^m \left( \max \left\{ 0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k} \right\} \right) \partial h_i(\tilde{x}^k) \right], \quad (3.35)$$

ou seja,

$$0 \in F(\tilde{x}^k) + \sum_{i=1}^m \lambda_i^k \partial h_i(\tilde{x}^k) \quad (3.36)$$

em que

$$\lambda_i^{k+1} = \lambda_i^k = \max \left\{ 0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k} \right\}. \quad (3.37)$$

Segue de (3.37) que se  $\lambda_i^k = 0$ , então  $h_i(\tilde{x}^k) \leq 0$ . E caso  $\lambda_i^k = \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k}$ , então  $h_i(\tilde{x}^k) = 0$  qualquer que seja  $\lambda_i^k \geq 0$ , isto é,  $\lambda_i^k h_i(\tilde{x}^k) = 0$ . Logo, as condições (3.36) – (3.37) implicam que  $(\tilde{x}^k, \tilde{\lambda}^k)$  é um par ótimo para o problema  $\text{PE}(f, \mathbb{K})$  e, portanto, pela Proposição 3.2,  $\tilde{x}^k \in S(f, \mathbb{K})$ . ■

### 3.3 Variantes do algoritmo IALEM

Nesta seção apresentamos a forma exata do algoritmo IALEM, que chamamos Método do Lagrangeano Aumentado Exato ou simplesmente EALM. Este algoritmo é um caso particular de IALEM, assim sua análise de convergência é análoga a que fora feita na seção anterior. Também apresentamos uma variante do algoritmo IALEM, o chamado Método Extragradiante Lagrangeano Aumentado Inexato Linearizado, que abreviamos por LIALEM, que na literatura é eventualmente tido como um método mais adequado a computação atual.

Iniciamos descrevendo o algoritmo EALM para resolver o problema de equilíbrio  $\text{PE}(f, \mathbb{K})$ .

**Algoritmo EALM para  $\text{PE}(f, \mathbb{K})$ .** *Considere uma sequência de parâmetros  $\{\gamma_k\} \subset \mathbb{R}_+$  limitada.*

**Passo 1.** *Iniciar o algoritmo com  $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}_+^m$ .*

**Passo 2.** *Na iteração  $k$ , a atualização  $x^{k+1}$  será a única solução do problema de desigualdade*

variacional irrestrito  $PE(\tilde{\mathcal{L}}_k^0, \mathbb{R}^n)$ , com  $\tilde{\mathcal{L}}_k^0$  definida a partir de (3.14).

**Passo 3.** Atualizar as variáveis duais através da fórmula

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

**Observação 3.3.** No Passo 2 do algoritmo EALM, o iterando  $x^{k+1}$  é único pelo Lema 3.1 aplicado a  $e^k = 0$  para todo inteiro  $k \geq 0$ .

A análise de convergência de EALM é análoga àquela feita para IALEM, basta tomarmos  $e^k = 0$ , para cada  $k \geq 0$ , em cada resultado de convergência apresentado na seção anterior.

Observe que enquanto a diferença entre os algoritmos IALEM e EALM é apenas a questão do vetor erro, a diferença entre os algoritmos LIALEM e IALEM será a função Lagrangeana a ser usada. Enquanto no algoritmo IALEM a função Lagrangeana usada foi (3.12), a qual foi obtida através da regularização inexata do Lagrangeano aumentado obtido a partir de (1.29), no algoritmo LIALEM usar-se uma Lagrangeana obtida através da linearização, na variável  $y$ , de (3.13). Para executarmos esta linearização assumimos que as restrições  $h_i$ , que determinam o conjunto viável  $\mathbb{K}$ , sejam continuamente diferenciáveis.

Seja  $\mathcal{L}$  a função Lagrangeana, para o problema  $PE(f, \mathbb{K})$ , dada por

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

Note que esta função é diferenciável na variável  $y$ , pois  $f(x, \cdot)$  é diferenciável e, para cada  $i \in \{1, \dots, m\}$ ,  $h_i$  é diferenciável. Assim, podemos definir uma aproximação de primeira ordem  $\bar{\mathcal{L}}$  para  $\mathcal{L}$ , em torno do ponto  $y = x$ , da seguinte forma:

$$\bar{\mathcal{L}}((x, \lambda), (y, \mu)) = f(x, y) + \sum_{i=1}^m \lambda_i [h_i(x) + \langle \nabla h_i(x), y - x \rangle] - \sum_{i=1}^m \mu_i h_i(x) \quad (3.38)$$

$$= \langle F(x), y - x \rangle + \sum_{i=1}^m \lambda_i \langle \nabla h_i(x), y - x \rangle + \sum_{i=1}^m (\lambda_i - \mu_i) h_i(x). \quad (3.39)$$

A expressão dada em (3.39) é denominada *Lagrangeano Linearizado*.

**Definição 3.2.** Seja  $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m$ . A função  $\bar{\mathcal{L}} : (\mathbb{R}^n \times \mathbb{R}^m) \times (\mathbb{R}^n \times \mathbb{R}^m) \rightarrow \mathbb{R}$  definida por

$$\bar{\mathcal{L}}((x, \lambda), (y, \mu)) = \langle F(x), y - x \rangle + \sum_{i=1}^m \lambda_i \langle \nabla h_i(x), y - x \rangle + \sum_{i=1}^m (\lambda_i - \mu_i) h_i(x) \quad (3.40)$$

é denominada *função Lagrangeana linearizada associada ao problema  $PE(f, \mathbb{K})$* .

**Observação 3.4.** Note que, na função Lagrangeana linearizada,  $f(x, y)$  não foi alterado, pois é uma função afim na variável  $y$ . Por conseguinte, veja que não há necessidade de linearizarmos a função  $\mathcal{L}$  na variável  $\mu$ , porque ela é afim como uma função dessa variável.



No algoritmo LIALEM usamos necessariamente uma Lagrangeana linearizada obtida de (3.13). Para determinarmos a expressão para esta Lagrangeana, primeiro devemos linearizar, na variável  $y$ , a expressão definida para  $s_i$  dada em (3.11).

Seja  $\bar{s}_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$  a linearização de  $s_i$  em torno do ponto  $y = x$ , temos que  $\bar{s}_i$  será dada por:

$$\begin{aligned}\bar{s}_i(x, y, \lambda, \gamma) &= s_i(x, x, \lambda, \gamma) + \langle \nabla s_i(x, x, \lambda, \gamma), y - x \rangle \\ &= \left\langle \frac{\gamma}{2} 2 \max \left\{ 0, \lambda_i + \frac{h_i(x)}{\gamma} \right\} \frac{1}{\gamma} \nabla h_i(x), y - x \right\rangle \\ &= \max \left\{ 0, \lambda_i + \frac{h_i(x)}{\gamma} \right\} \langle \nabla h_i(x), y - x \rangle.\end{aligned}$$

A partir da expressão dada acima, podemos escrever a Lagrangeana linearizada (na variável  $y$ ) para (3.13) da seguinte forma:

$$\bar{\mathcal{L}}(x, y) = \langle F(x), y - x \rangle + \gamma \langle x - z, y - x \rangle + \sum_{i=1}^m \bar{s}_i(x, y, \lambda, \gamma) - \langle e, y - x \rangle.$$

Conhecendo a linearização do Lagrangeano aumentado dado em (3.13), podemos descrever o Algoritmo LIALEM, para o problema  $PE(f, \mathbb{K})$ , assim como segue:

**Algoritmo LIALEM para  $PE(f, \mathbb{K})$ .** Considere uma sequência exógena limitada  $\{\gamma_k\} \subset \mathbb{R}_+$  de parâmetros de regularização e uma tolerância de erro relativo  $\sigma \in (0, 1)$ .

**Passo 1.** Iniciar o algoritmo com  $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}_+^m$ .

**Passo 2.** Dado  $(x^k, \lambda^k)$ , defina  $\bar{s}_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$  como

$$\bar{s}_i(x, y, \lambda, \mu) = \max \left\{ 0, \lambda_i + \frac{h_i(x)}{\gamma} \right\} \langle \nabla h_i(x), y - x \rangle, \quad (3.41)$$

encontre  $\tilde{x}^k \in \mathbb{R}^n$  que resolve o problema perturbado  $PE(\bar{\mathcal{L}}_k^e, \mathbb{R}^n)$ , em que  $\bar{\mathcal{L}}_k^e : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  é definida como

$$\bar{\mathcal{L}}_k^e(x, y) = \langle F(x), y - x \rangle + \gamma_k \langle x - x^k, y - x \rangle + \sum_{i=1}^m \bar{s}_i(x, y, \lambda^k, \gamma_k) - \langle e^k, y - x \rangle \quad (3.42)$$

com  $F$  dado por (3.2) e  $e^k \in \mathbb{R}^n$  que satisfaz a condição:

$$\|e^k\| \leq \sigma \gamma_k \|(\tilde{x}^k - x^k, \lambda^{k+1} - \lambda^k)\| \quad (3.43)$$

onde  $\lambda^{k+1} = (\lambda_1^{k+1}, \dots, \lambda_m^{k+1})$  é atualizado no próximo passo.

**Passo 3.** Definir a próxima iteração de  $\lambda^k$  como

$$\lambda_i^{k+1} = \max \left\{ 0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k} \right\} \quad (3.44)$$

para cada  $i = 1, \dots, m$ .

**Passo 4.** Se  $(x^k, \lambda^k) = (\tilde{x}^k, \lambda^{k+1})$ , então PARE!. Caso contrário, faça

$$x^{k+1} = \tilde{x}^k - \frac{1}{\gamma_k} e^k \quad (3.45)$$

e volte para o Passo 2.

**Observação 3.5.** A única diferença entre os algoritmos IALEM e LIALEM é a bifunção que define o subproblema de desigualdade variacional irrestrito dado no Passo 2. De fato, na  $k$ -ésima iteração, temos que IALEM resolve o problema  $PE(\tilde{\mathcal{L}}_k^e, \mathbb{R}^n)$ , com  $\tilde{\mathcal{L}}_k^e$  definida por (3.14), enquanto LIALEM resolve o problema  $PE(\bar{\mathcal{L}}_k^e, \mathbb{R}^n)$  com  $\bar{\mathcal{L}}_k^e$  dada por (3.42).

Uma pergunta natural consiste em indagar se os resultados de convergência do algoritmo IPPEM (aplicado ao problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ ) também se aplicam ao algoritmo LIALEM, assim como fora feito para IALEM na Seção 3.2. A resposta é afirmativa. Por exemplo, um resultado que o algoritmo IPPEM nos fornece a respeito de convergência é o Teorema 2.1, e este se aplica ao problema  $PE(\bar{\mathcal{L}}, \mathbb{R}^n)$ , devido a monotonicidade de  $\bar{\mathcal{L}}$ .

**Proposição 3.4.** Considere o problema  $PE(f, \mathbb{K})$ . Assuma que  $f$  é monótona. Então  $\bar{\mathcal{L}}$ , dada por (3.40), é monótona.

*Demonstração.* Como  $f$  é monótona, então

$$\begin{aligned} \bar{\mathcal{L}}((x, \lambda), (y, \mu)) + \bar{\mathcal{L}}((y, \mu), (x, \lambda)) &= f(x, y) + f(y, x) + \sum_{i=1}^m \lambda_i [h_i(x) + \langle \nabla h_i(x), y - x \rangle - h_i(y)] \\ &\quad + \sum_{i=1}^m \mu_i [h_i(y) + \langle \nabla h_i(y), x - y \rangle - h_i(x)] \\ &\leq \sum_{i=1}^m \lambda_i [h_i(x) + \langle \nabla h_i(x), y - x \rangle - h_i(y)] + \sum_{i=1}^m \mu_i [h_i(y) \\ &\quad + \langle \nabla h_i(y), x - y \rangle - h_i(x)]. \end{aligned}$$

Sendo as restrições  $h_i$  convexas e diferenciáveis, obtemos

$$h_i(y) \geq h_i(x) + \langle \nabla h_i(x), y - x \rangle \quad \forall y \in \mathbb{K};$$

$$h_i(x) \geq h_i(y) + \langle \nabla h_i(y), x - y \rangle \quad \forall x \in \mathbb{K}.$$

e, portanto

$$\bar{\mathcal{L}}((x, \lambda), (y, \mu)) + \bar{\mathcal{L}}((y, \mu), (x, \lambda)) \leq 0.$$

Isto é,  $\bar{\mathcal{L}}$  é monótona. ■

Aplicando o Teorema 2.1, concluímos que a sequência gerada pelo algoritmo IPPEM aplicado ao  $PE(\bar{\mathcal{L}}, \mathbb{R}^n \times \mathbb{R}_+^m)$  convergirá a uma solução deste problema. Outros resultados

de convergência que são facilmente verificados, ao substituir o problema  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$  pelo problema  $PE(\bar{\mathcal{L}}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , são as Proposições 3.1 e 3.2. De fato, basta substituir, na demonstração de cada proposição, a condição Lagrangeana

$$0 \in F(x^*) + \sum_{i=1}^m \lambda_i^* \partial h_i(x^*)$$

por

$$0 = F(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*)$$

que é equivalente devido a diferenciabilidade de  $h_i$ , para cada  $i \in \{1, \dots, m\}$ . Para outros resultados de convergência, como os Teoremas 3.1 e 3.2, isto também não é uma tarefa complicada. De fato, veja que o Teorema 3.1 reescrito para o algoritmo LIALEM aplicado ao problema  $PE(f, \mathbb{K})$  fica na seguinte forma:

**Teorema 3.3.** *Considere o problema  $PE(f, \mathbb{K})$ . Assuma que  $f$  é monótona. Fixe uma sequência  $\{\gamma_k\} \subset \mathbb{R}_+$  e uma tolerância de erro relativo  $\sigma \in (0, 1)$ . Sejam  $\{(x^k, \lambda^k)\}$  a sequência gerada pelo algoritmo LIALEM aplicado ao  $PE(f, \mathbb{K})$ , com vetor de erro  $e^k \in \mathbb{R}^n$  associado, e  $\{(\bar{x}^k, \bar{\lambda}^k)\}$  a sequência gerada pelo algoritmo IPPEM aplicado ao  $PE(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , com vetor de erro  $(e^k, 0) \in \mathbb{R}^n \times \mathbb{R}^m$  associado. Suponha que o parâmetro de regularização  $\gamma$  e o erro relativo  $\sigma$  sejam os mesmos em ambos os algoritmos. Se  $(x^0, \lambda^0) = (\bar{x}^0, \bar{\lambda}^0)$ , então  $(x^k, \lambda^k) = (\bar{x}^k, \bar{\lambda}^k)$  para todo  $k \in \mathbb{N}$ .*

*Demonstração.* A prova deste teorema é análoga a feita no Teorema 3.1, basta observar que sendo as restrições  $h_i$  diferenciáveis, então as funções Lagrangeanas  $\hat{\mathcal{L}}_k^e$  e  $\bar{\mathcal{L}}_k^e$  serão diferenciáveis na variável  $y$ , assim as condições (3.31) e (3.33) acabam reescritas, respectivamente, como

$$\begin{aligned} \gamma_k(\bar{x}^k - \hat{x}^k) + e^k &= F(\hat{x}^k) + \left[ \sum_{i=1}^m \left( \max \left\{ 0, \lambda_i^k + \frac{h_i(\hat{x}^k)}{\gamma_k} \right\} \right) \nabla h_i(\hat{x}^k) \right] \\ \gamma_k(\bar{x}^k - \tilde{x}^k) + e^k &= F(\tilde{x}^k) + \left[ \sum_{i=1}^m \left( \max \left\{ 0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k} \right\} \right) \nabla h_i(\tilde{x}^k) \right]. \end{aligned}$$

Daí, utilizando a hipótese de indução e as fórmulas obtidas para as atualizações das variáveis primais e duais (dos algoritmos IPPEM e LIALEM), obtemos o resultado desejado. ■

Quanto ao Teorema 3.2, temos que a sua versão para o algoritmo LIALEM segue diretamente do Teorema 3.3, e do Corolário 2.1 e Corolário 3.1.

# Considerações Finais

Apresentamos neste trabalho métodos para a resolução de um problema de desigualdade variacional definido na Seção 1.4. Os métodos IALEM e seus variantes são algoritmos que resolvem o problema  $\text{PDV}(F, \mathbb{K})$  seja no caso que se tenha as restrições, que determinam o conjunto viável  $\mathbb{K}$ , diferenciáveis (como em LIALEM) ou não (que é o caso de IALEM).

Estes algoritmos encontram soluções do  $\text{PDV}(F, \mathbb{K})$  graças as propriedades que definimos sobre o operador  $F$ , como a monotonicidade, e sobre as restrições, como a convexidade. Em especial, a monotonicidade teve um papel importante ao garantir as hipóteses necessárias (como no Lema 2.1) à convergência de IPPEM, aplicado ao problema  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ , a uma solução deste problema. Além disso, tal propriedade é que garante a conexão entre IPPEM e IALEM, em particular, a ausência da monotonicidade em IPPEM é uma questão a ser pensada para futuras pesquisas, conforme sugere Iusem em [5]. De fato, é razoável analisar o que aconteceria se substituíssemos a monotonicidade de  $F$  por uma característica mais geral, como a pseudomonotonicidade, quais hipóteses seriam necessárias para garantir a convergência do algoritmo IPPEM e como elas influenciariam na análise de convergência de IALEM. Isto é um questionamento que poderia ser sugerido como temas de futuros trabalhos.

Outro ponto importante, a respeito da monotonicidade de  $F$ , é que tal propriedade garante a existência e unicidade de solução ao problema  $\text{PE}(\tilde{f}^e, \mathbb{K})$ , o qual é a forma regularizada inexata do problema  $\text{PE}(f, \mathbb{K})$ , mesmo que este último não possua solução. Isto é realmente um resultado notável. De fato, este resultado, expresso pelo Lema 3.1, garantiu a boa definição do Passo 2 do algoritmo IALEM e do algoritmo IPPEM, aplicados, respectivamente, aos problemas  $\text{PE}(f, \mathbb{K})$  e  $\text{PE}(\mathcal{L}, \mathbb{R}^n \times \mathbb{R}_+^m)$ .

Terminamos este trabalho ressaltando a importância dos métodos aqui apresentados. Os algoritmos IALEM, LIALEM e EALM surgem como propostas de métodos que mostram boas propriedades de convergência sob hipóteses razoáveis, e que computacionalmente podem se mostrar mais vantajosos que outros métodos existentes para resolução de problemas de desigualdade variacional.

Claro que não está sendo comparada a sua eficiência em relação a outros métodos de Lagrangeano aumentado existentes, até por que a atual implementação destes métodos é ainda objeto de pesquisa. Com a implementação destes algoritmos, fará sentido a sua comparação, em termos de eficiência (velocidade, precisão de resultados), com outros métodos de Lagrangeano aumentado para problemas de desigualdade variacional.

# Referências Bibliográficas

- [1] BUTNARIU, D.; IUSEM A.N., *Totally convex functions for fixed points computation and infinite dimensional optimization*. Dordrecht: Kluwer Academic Publishers, 2000.
- [2] IUSEM, A.N., *Proximal Point Methods in Optimization*, 20º Colóquio Brasileiro de Matemática. Rio de Janeiro: IMPA, 1995. [7](#), [8](#), [17](#), [18](#), [21](#)
- [3] IUSEM, A.N., *Augmented Lagrangians Methods and Proximal Points Methods for Convex Optimization*, *Investigación Operativa*, **8**, p. 11-49, (1999). [21](#)
- [4] IUSEM, A.N.; KASSAY, G.; SOSA, W., *On certain conditions for the existence of solutions of equilibrium problems* [16](#)
- [5] IUSEM, A.N.; NASRI, M., *Augmented Lagrangian Methods for Variational Inequality Problems*, *RAIRO-Oper. Res.*, **44**, p.p. 5-25, 2010. [1](#), [14](#), [22](#), [26](#), [36](#), [37](#), [38](#), [39](#), [40](#), [42](#), [43](#), [48](#), [52](#), [58](#)
- [6] IUSEM, A.N.; NASRI, M., *Inexact proximal point methods for equilibrium problems in Banach spaces*, *Numerical Functional Analysis and Optimization*, , **28**(11-12), p.p. 1279-1308, 2007. [2](#), [8](#), [9](#), [10](#), [12](#), [16](#), [28](#), [29](#), [32](#)
- [7] IUSEM, A.N. SOSA, W., *A proximal point method for equilibrium problems in Hilbert spaces*, *Optimization: A Journal of Mat. Progr. and Oper. Res.*, **59**, 8, p.p. 1259-1274, 2010. [27](#)
- [8] IZMAILOV, A.; SOLODOV, I., *Otimização-volume 1: Condições de otimalidade, Elementos de Análise Conveza e de Dualidade*, 2.ed. Rio de Janeiro: IMPA, 2009. [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [9](#), [18](#), [23](#), [24](#)
- [9] IZMAILOV, A.; SOLODOV, I., *Otimização-volume 2: Métodos Computacionais*. Rio de Janeiro: IMPA, 2007. [19](#)
- [10] LUENBERGER, D.; YE, Y., *Linear and Nonlinear Programming*, 3.ed. Stanford: Springer, 2008. [19](#), [20](#)
- [11] URRUTY, J.B.H.; LEMARÉCHAL, C., *Convex Analysis and Minimization Algorithms I*, 2.ed. Springer: Berlin, 1996. [23](#)

- 
- [12] VELÁSQUEZ, M.A., *Trajectoria Central, Métodos de Ponto Proximal Generalizado e Trajetória de Cauchy em Variedades Riemannianas.*, Dissertação de Mestrado, Universidade Federal de Campina Grande, (2007). 8