SMNE Simpósio de Métodos Numéricos em Engenharia
25 a 27 de outubro, 2017
UFPR

LAGRANGEANO AUMENTADO QUADRÁTICO E EXPONENCIAL APLICADO A PROBLEMAS DE EQUILÍBRIO

Elvis Manuel Rodríguez Torrealba[†]
Luiz Carlos Matioli[*]
Programa de Pós-Graduação em Matemática - PPGM
Universidade Federal do Paraná, UFPR
Curitiba, Brasil
[†]Relvis.375@gmail.com
[*]matioli@ufpr.br

Romulo Alberto Castillo
Centro Tecnológico de Joinville - CTJ
Universidade Federal de Santa Catarina, UFSC
Joinville, Brasil
romulo.castillo@ufsc.br

***Resumo*—Neste trabalho propomos uma reformulação dos métodos de Lagrangeano aumentado para resolver problemas de equilíbrio gerais proposto em [3]. A ideia principal é combinar as condições dadas em [3], [4] e [5] para introduzir métodos de Lagrangeano aumentado com outras penalidades além da penalidade clássica de Powell-Hestenes-Rockafellar (PHR) dada em [3]. Veremos que a teoria de convergência é construída via ponto proximal, a partir da relação de dualidade, utilizando distâncias de Bregman.**

***Palavras-chave*—Problema de Equilíbrio, Ponto proximal, Lagrangeano aumentado.**

I. INTRODUÇÃO

Pelo que consta na literatura o problema de equilíbrio foi originalmente introduzido em [6], [7] com o nome de problema de desigualdade de Ky Fan e retomado formalmente por Blum e Oettli em [2], em cujo artigo, os autores apresentam o problema na forma como é conhecido até hoje. Mundialmente, este problema tem recebido muita atenção por parte dos pesquisadores. Por um lado, devido a aplicações em áreas como Física, Química, Engenharia, Economia e Matemática. Neste último, se deve principalmente, a sua formulação matemática simples e unificada que permite incluir diversos problemas como casos particulares, dentre os quais podemos mencionar: problema de minimização convexa, problemas de ponto fixo, problemas de complementaridade, problemas de equilíbrio de Nash, problemas de desigualdade variacional e problemas de minimização vetorial (ver, [1], [4], [8]). A maioria dos autores têm estudado condições para encontrar soluções do problema de equilíbrio em dimensão finita e infinita, por exemplo [1], [4], [10], e estas condições são muito ricas no caso de dimensão finita, pois permitem construir diversos algoritmos, dentre os quais destacamos [1], [3]–[5], [9].

Recentemente, em [3] foi introduzido o método de Lagrangeano aumentado para resolver problemas de equilíbrio em dimensão finita, para o caso em que o conjunto viável é formado por restrições de desigualdade convexas. Este método foi introduzido utilizando a função de penalidade clássica de Powell-Hestenes-Rockafellar (PHR). Na estrutura deste método de Lagrangeano aumentado para problemas de equilíbrio apresentado em [3], os autores mostram uma conexão com os métodos de ponto proximal para problemas de equilíbrio definido em [4], o qual, no caso de dimensão finita permite mostrar que ambos algoritmos (Lagrangeano aumentado e ponto proximal) geram uma mesma sequência e assim aproveitar os resultados obtidos em [4]. Seguindo a ideia do método de Lagrangeano aumentado para problemas de otimização, a nossa proposta é reconstruir os Métodos de Lagrangeano aumentado para problemas de equilíbrio definidos em [3] e mostrar que combinando as hipóteses dos métodos de

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -(CAPES).

ponto proximal dados em [4], [5] é possível considerar outros tipos de penalidades, como a penalidade exponencial. Colocando assim, penalidades que são diferenciáveis que não estavam presentes na clássica (PHR) dada em [3].

II. Preliminares

Ao longo do texto, vamos considerar algumas hipóteses, definições e resultados importantes que utilizaremos na definição e análise de convergência dos algoritmos.

Dados um conjunto $C \subset \mathbb{R}^n$ não vazio, convexo e fechado, e uma bifunção $f : C \times C \to \mathbb{R}$, que satisfaz:

P1: $f(x, x) = 0$ para todo $x \in C$,
P2: $f(x, \cdot) : C \to \mathbb{R}$ é convexa e semicontínua inferior para todo $x \in C$,
P3: $f(\cdot, y) : C \to \mathbb{R}$ é semicontínua superior para todo $y \in C$.

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

$$\text{Encontrar } x^* \in C \text{ tal que } f(x^*, y) \geq 0, \text{ para todo } y \in C. \quad (1)$$

O conjunto de soluções para este problema, será denotado como $S(f, C)$.

Dado que a análise de convergência será realizada utilizando os métodos de ponto proximal então precisamos lembrar o conceito de distância de Bregman, assim como algumas propriedades importantes que utilizaremos ao longo do texto.

A. Distâncias de Bregman e suas propriedades

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

$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y \rangle, \quad (2)$$

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

A seguinte proposição estabelece algumas propriedades básicas satisfeitas pela função D_ψ definida em (2). A demonstração é omitida pois decorre diretamente da definição.

Proposição 1. *Sejam h e D_ψ dadas como acima, então*

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

Além das propriedades dadas na Proposição (1), vamos estabelecer outras hipóteses sobre a função ψ que define D_ψ.

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

Para definir a segunda hipótese precisaremos lembrar o conceito de convexidade total definido em [11]. Daqui em diante, denotaremos $\mathbb{R}^n_+$ o ortante não negativo dos números reais e $\mathbb{R}^n_{++} = \mathbb{R}^n_+ \setminus \{0\}$ o ortante positivo.

Definição 2. *[Convexidade total] Definimos o módulo de convexidade de ψ, como sendo a função $\nu_\psi : int\mathcal{D} \times \mathbb{R}_+ \to \mathbb{R}_+$ dada por*

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

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

Este conceito é estudado em [11] e entre as propriedades mais importantes em dimensão finita é estabelecido na seguinte proposição.

Proposição 3. *Seja $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função diferenciável, convexa, própria e semincontínua inferior. Então vale que:*

(i) *Se ψ é totalmente convexa então, é estritamente convexa.*
(ii) *Seja $\mathcal{D}$ fechado. Se ψ é continua e estritamente convexa em $\mathcal{D}$ então, ψ é totalmente convexa.*

Assim, dizer que ψ é totalmente convexa equivale a dizer que ψ é estritamente convexa quando $\mathcal{D}$ é fechado. Com isto claro, definimos a segunda hipótese, sobre conjuntos limitados.

H2: ψ é totalmente convexa em conjuntos limitados, isto é, $\inf_{x \in C} \nu_\psi(x, t) > 0$, para todo conjunto limitado $C \subset int\mathcal{D}$ e $t > 0$.

As seguintes hipóteses exigidas serão estabelecidas sobre $\nabla\psi$. A primeira exige que $\nabla\psi$ seja uniformemente contínua em conjuntos limitados e a última exige sobrejetividade em $int\mathcal{D}$. Mais precisamente:

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

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

As três primeiras hipóteses foram definidas de maneira independente tanto em [4] como em [5], para definir algoritmos de ponto proximal que resolvem o problema de equilíbrio (1). Mencionamos que em [4] foi considerada além de H4, outras hipóteses para mostrar convergência e a boa definição de algoritmos inexatos.

Uma propriedade importante que garante a boa definição dos algoritmos inexatos é dada pela projeção com distâncias de Bregman sobre um conjunto convexo e fechado.

Proposição 4. *Suponha que ψ satisfaz H1-H2 e $C \subset \mathcal{D}$ é convexo, fechado e não vazio, então para $\bar{x} \in int\mathcal{D}$ o problema,*

$$\begin{array}{ll} \min & D_\psi(x, \bar{x}) \\ s.a & x \in C, \end{array} \tag{3}$$

admite solução única. Em particular, Se $C \subset int\mathcal{D}$ então vale que $\langle \nabla\psi(\bar{x}) - \nabla\psi(\hat{x}), y - \hat{x}\rangle \leq 0$, para todo $y \in C$ e $\hat{x}$ solução de (3).

Demonstração. Veja [11] pág. 70. □

A seguir, veremos que fazendo uma pequena adaptação em uma das hipóteses dadas em [5], podemos obter os mesmos resultados de convergência para os algoritmos apresentados em [4].

B. Algoritmos de ponto proximal exatos e inexatos para problemas de equilíbrio.

Para obter uma boa definição e convergência dos algoritmos de ponto proximal dados em [4], [5], vamos precisar adicionar as seguintes hipóteses sobre a bifunção f do problema (1):

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

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

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

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

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

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

P5: Para toda sequência $(x^n)_{n\in\mathbb{N}} \subset C$ satisfazendo $\lim\limits_{n\to\infty} \|x^n\| = +\infty$, existe $u \in C$ e $n_0 \in \mathbb{N}$ tal que $f(x^n, u) \leq 0$ para todo $n \geq n_0$.

Observação 5. *Estas hipóteses são as mais comuns encontradas na literatura para abordar existência de soluções do problema de equilíbrio. As hipóteses P4 e P4* são conhecidas como $\theta-submontona$ e Pseudomonótona, respectivamente. A hipótese clássica que podemos encontrar na literatura é dada no caso particular em que $\theta = 0$, na hipótese P4. Neste caso, note que a expressão se reduz a $f(x, y) + f(y, x) \leq 0$ para todo $x, y \in C$ e estes tipos de bifunções são conhecidas como monótonas. Outro caso particular para hipótese P4 é dado quando $\psi = \frac{1}{2}\|\cdot\|^2$. Neste caso, note que hipótese P4 é reescrita como $f(x, y) + f(y, x) \leq \theta\|x - y\|^2$ para todo $x, y \in C$, a qual é usada em [3].*

Adicionalmente, mencionamos que em [10], os autores realizaram um estudo detalhado, mostrando as relações e diferenças entres estas hipóteses e como estas influenciam na existência de soluções do problema de equilíbrio definido em (1).

Voltando ao objetivo desta seção de resolver o problema de equilíbrio (1) utilizando algoritmos de ponto proximal exatos ou inexatos, vamos introduzir, como em [4], o conceito de regularização de uma bifunção f dada como em (1).

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

$$\tilde{f}^e(x, y) = f(x, y) + \gamma\langle \nabla\psi(x) - \nabla\psi(\bar{x}), y - x\rangle - \langle e, y - x\rangle \tag{4}$$

O seguinte resultado é uma adaptação da Proposição 3.1 dada em [5] e estabelece algumas propriedades básicas para a função $\tilde{f}^e$.

Proposição 6. *Seja $\bar{x} \in C \subset int\mathcal{D}$. Suponha que f satisfaz as condições P1−P4 com $\theta < \gamma$, em que θ é a constante de submonotonicidade não necessariamente nula. Então a função $\tilde{f}^e : C \times C \to \mathbb{R}$ dada como em (4), satisfaz P1−P4 com $\theta = 0$ (isto é, $\tilde{f}^e$ é monótona). Além disso, se*

P6: *Para cada $(x^k)_{k\in\mathbb{N}} \subset C$ tal que $\lim\limits_{k\to\infty} \|x^k\| = \infty$, satisfaz que*

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

então $\tilde{f}^e$ satisfaz P5.

Demonstração. Note que $\tilde{f}^e$ satisfaz P1 por definição. Para mostrar P2, basta observar que como f satisfaz P2 e $y \to \langle \nabla\psi(x) - \nabla\psi(\bar{x}) - e, y - x\rangle$ é contínua e convexa então, $\tilde{f}^e$ também satisfaz esta hipótese. Analogamente a função $x \to \langle \nabla\psi(x) - \nabla\psi(\bar{x}) - e, y - x\rangle$ é contínua o que mostra P3. Agora como f satisfaz P4 então temos que

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

Ou seja, $\tilde{f}^e$ satisfaz P4 com $\theta = 0$. Para finalizar, mostremos P5. Seja $(x^k)_{k\in\mathbb{N}} \subset C$ satisfazendo $\|x^k\| \to \infty$. Utilizando novamente o fato que f satisfaz P4 temos que para cada $k \in \mathbb{N}$

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

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

A seguinte proposição, mostrada em [5] estabelece que a solução do problema regularizado EP($\tilde{f}^e, C$) é única.

Proposição 7. *Seja $\bar{x} \in C \subset int\mathcal{D}$. Suponha que f satisfaz as condições P1−P4 com $\theta < \gamma$. Então as seguintes afirmações são válidas:*

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

Uma condição suficiente para garantir que a hipótese P6 é satisfeita ocorre quando ψ é uma função coerciva, isto é, $\lim_{\|x\|\to\infty} \frac{\psi(x)}{\|x\|} = +\infty$. O qual é mostrado na seguinte proposição.

Proposição 8. *Suponha que f satisfaz P1-P4. Se ψ é coerciva e admite derivada direcional diferenciável em C, então P6 é satisfeita e consequêntemente, $S(\tilde{f}^e, C) \neq \emptyset$.*

Demonstração. Veja Corolário 3.3 de [5]. □

Tendo em vista uma classe de funções para as quais garantimos solução do problema regularizado com perturbação e, podemos considerar os algoritmos de ponto proximal exato e inexatos definidos em [4].

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

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

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

$$\tilde{f}^e_j(x, y) = f(x, y) + \gamma_j\langle\nabla\psi(x) - \nabla\psi(x^j), y - x\rangle - \langle e^j, y - x\rangle, \quad (5)$$

e e^j satisfaz

$$\|e^j\| \leq \begin{cases} \sigma\,\gamma^j D_\psi(\tilde{x}^j, x^j) & se \quad \|x^j - \tilde{x}^j\| < 1 \\ \sigma\gamma^j\nu_\psi(x^j, 1) & se \quad \|x^j - \tilde{x}^j\| \geq 1. \end{cases} \quad (6)$$

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

$$x^{j+1} = \arg\min_{x\in H_j} D_\psi(x, x^j). \quad (7)$$

Algoritmo IPPEM: Método de ponto proximal inexato com passo extragradiente para EP(f, C).

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

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

$$\tilde{f}^e_j(x, y) = f(x, y) + \gamma_j\langle\nabla\psi(x) - \nabla\psi(x^j), y - x\rangle - \langle e^j, y - x\rangle, \quad (8)$$

e e^j satisfaz

$$D_\psi(\tilde{x}^j, (\nabla\psi)^{-1}[\nabla\psi(\tilde{x}^j) - \gamma_j^{-1}e^j]) \leq \sigma D_\psi(\tilde{x}^j, x^j), \quad (9)$$

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

$$x^{j+1} = (\nabla\psi)^{-1}[\nabla\psi(\tilde{x}^j) - \gamma_j^{-1}e^j]. \quad (10)$$

As diferenças entre os Algoritmos IPPBPM e IPPEM são evidentes nos passos 2 e 4. A ideia no passo 2 é praticamente a mesma para ambos algoritmos, diferenciando-se apenas na condição que e^j deve cumprir. Agora, o Passo 4 em cada um dos algoritmos se divide em duas etapas, a primeira consiste em verificar se o critério de parada é atingido e caso isto não seja satisfeito, cada algoritmo utiliza uma forma específica para atualizar a próxima iteração. Note que estes passos estão bem definidos para ψ satisfazendo as hipóteses H1-H4 e a Proposição 4.

Por outro lado, as diferenças presentes nos algoritmos IPPBPM e IPPEM são evidentemente eliminadas quando $e^j = 0$, para todo $j \in \mathbb{N}$, pois neste caso vemos que o Passo 2 de cada algoritmo se reduz a encontrar uma solução do subproblema de equilíbrio EP($\tilde{f}_j$,C), com $\tilde{f}_j$ dada como em (8), com $e^j = 0$. Analogamente, no Passo 4, podemos observar que se $e^j = 0$ para todo $j \in \mathbb{N}$, então de (7) e (10) obtemos que a próxima iteração é simplificada para a igualdade $x^{j+1} = \tilde{x}^j$, em que $\tilde{x}^j$ é a solução obtida previamente do Passo 2. Com estes argumentos podemos definir o método de ponto proximal Exato.

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

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

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

$$\tilde{f}_j(x, y) = f(x, y) + \gamma_j\langle\nabla\psi(x) - \nabla\psi(x^j), y - x\rangle, \quad (11)$$

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

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

A análise de convergência para cada algoritmo é feita com detalhe em [4]. Salientamos que neste artigo, os autores consideram outras hipóteses que aqui serão substituídas pela hipótese P6.

Teorema 9. *Considere EP(f, C). Suponha que f satisfaz P1-P4 e P6, e que algumas entre as hipóteses P4*,P4′,P4″ é válida. Tome $\psi : \mathbb{R}^n \to \mathbb{R}$ satisfazendo H1-H4 e uma sequência exógena $(\gamma_j)_{j\in\mathbb{N}} \subset (\theta, \bar{\gamma}]$, em que $\bar{\gamma} > 0$ e θ é constante de submonotonicidade de P4. Seja $(x^j)_{j\in\mathbb{N}}$ a sequência gerada pelo algoritmo IPPBMP (ou IPPEM). Se EP(f, C) tem solução então*

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

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

Demonstração. Veja Teoremas 5.5 e 5.8 de [4]. $\square$

Na próxima seção, vamos estabelecer a relação dos algoritmos de ponto proximal, com os algoritmos de Lagrangeano aumentado para resolver problemas de equilíbrio.

III. MÉTODO DE LAGRANGEANO AUMENTADO QUADRÁTICO E EXPONENCIAL PARA PROBLEMAS DE EQUILÍBRIO

Consideremos a bifunção f e o conjunto C dados como em (1). Para definir os métodos de Lagrangeano aumentado vamos estabelecer algumas considerações iniciais. Ao longo do capítulo, vamos considerar duas funções ψ, sendo estas, $\psi^1(x) = \frac{1}{2}\|x\|^2$, $\psi^2(x) = \sum_{i=1}^{n}(x_i \log(x_i) - x_i + 1)$. Note que tais funções tem dominios $\mathcal{D}$ diferentes, isto é, $\mathcal{D} = \mathbb{R}^n$ para ψ^1 e $\mathcal{D} = \mathbb{R}^n_+$ para ψ^2, assumindo que $0\log(0) = 0$. Além disso, note que estas funções verificam as hipóteses H1-H4 para qualquer conjunto $\mathcal{K} \subset int(D)$ convexo fechado e não vazio e portanto podemos utilizar estas funções nos algoritmos de ponto proximal definidos na seção anterior. No caso particular em que $int\mathcal{D} = \mathbb{R}^n$ (isto é, para ψ^1) então $\mathcal{K}$ pode ser escolhido como $\mathbb{R}^n$.

Por outro lado para definir o método de Lagrangeano aumentado, vamos estabelecer que a bifunção f de EP(f,C), pode ser estendida ao domínio $\mathcal{K}\times\mathcal{K}$ preservando as propriedades básicas P1−P4 e P4$''$. Além disso, neste capítulo o conjunto C em EP(f,C) será considerado como

$$C = \{x \in \mathcal{K} : g_i(x) \leq 0, \forall\ i \in 1,\dots,m\}, \tag{12}$$

em que para cada $1 \leq i \leq m$, $g_i : \mathcal{D} \to \mathbb{R}$ é uma função convexa.

Agora, a função lagrangeana definida como em [3] para o problema de equilíbrio EP(f,C) é uma bifunção, $\mathcal{L} : \mathcal{K}\times\mathbb{R}^m\times\mathcal{K}\times\mathbb{R}^m \to \mathbb{R}$ dada por

$$\mathcal{L}\left((x,\lambda),(y,\mu)\right) = f(x,y) + \sum_{i=1}^{m}\lambda_i g_i(y) - \sum_{i=1}^{m}\mu_i g_i(x). \tag{13}$$

No que segue, utilizaremos duas funções de penalidade para definir métodos que resolvem problemas de equilíbrio. Para cada função ψ^j, $j = 1,2$, vamos considerar a penalidade $P_i : \mathcal{K}\times\mathcal{K}\times\mathbb{R}^m_+\times\mathbb{R}_{++} \to \mathbb{R}\cup\{+\infty\}$ dada por

$$P_i(x,y,\lambda_i,\gamma) = \gamma\left[\mathcal{P}^s\left(g_i(y),\gamma,\lambda_i\right) - \mathcal{P}^s\left(g_i(x),\gamma,\lambda_i\right)\right] \tag{14}$$

em que $1 \leq i \leq m$, $\mathcal{P}^s$ denota a função de penalidade associada à função ψ^s que estivermos considerando. Assim, definimos a função Lagrangeano aumentado para resolver problemas de equilíbrio, proposta em [3], como $\widetilde{\mathcal{L}} : \mathcal{K}\times\mathcal{K}\times\mathbb{R}^m_+\times\mathcal{K}\times\mathbb{R}_{++} \to \mathbb{R}$ dada por:

$$\widetilde{\mathcal{L}}(x,y,\lambda,z,\gamma) = f(x,y) + \gamma\langle\nabla\psi(x) - \nabla\psi(z), y - x\rangle + \sum_{i=1}^{m}P_i(x,y,\lambda_i,\gamma). \tag{15}$$

Note que a função P_i depende da escolha da função ψ. Com isto, apresentamos modificações das versões de algoritmos de Lagrangeano aumentado definida em [3], as quais dependem da escolha da função ψ.

Algoritmo IALPM (Método de Lagrangeano aumentado inexato com Projeção)

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

Passo 2: Dado o ponto (x^j,λ^j), encontre o par $(\tilde{x}^j, e^j) \in \mathcal{K}\times\mathcal{K}$ tal que $\tilde{x}^j$ resolve o problema EP($\widetilde{\mathcal{L}}^e_j,\mathcal{K}$), onde $\widetilde{\mathcal{L}}^e_j$ é definida como

$$\begin{aligned}\widetilde{\mathcal{L}}^e_j(x,y) &= f(x,y) + \gamma_j\langle\nabla\psi(x) - \nabla\psi(x^j), y - x\rangle \\ &\quad + \sum_{i=1}^{m}P_i(x,y,\lambda_i,\gamma_j) - \langle e^j, y - x\rangle,\end{aligned} \tag{16}$$

e a função P_i é dada como em (14), e e^j satisfaz

$$\|e^j\| \leq \begin{cases}\sigma\,\gamma_j D_\psi(\tilde{x}^j,x^j) & se \quad \|x^j - \tilde{x}^j\| < 1\\ \sigma\gamma_j\nu_\psi(x^j,1) & se \quad \|x^j - \tilde{x}^j\| \geq 1.\end{cases} \tag{17}$$

Passo 3: Defina λ^{j+1} como sendo

$$\lambda_i^{j+1} = \gamma_j(\mathcal{P}^s)'\left(g_i(\tilde{x}^j),\gamma_j,\lambda_i^j\right), \tag{18}$$

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

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

$$x^{j+1} = \operatorname*{argmin}_{x\in H_j} D_\psi(x,x^j) \tag{19}$$

Algoritmo IALEM (Metodo de Lagrangeano aumentado inexato extragradiente)

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

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

$$\begin{aligned}\widetilde{\mathcal{L}}^e_j(x,y) &= f(x,y) + \gamma_j\langle\nabla\psi(x) - \nabla\psi(x^j), y - x\rangle \\ &\quad + \sum_{i=1}^{m}P_i(x,y,\lambda_i,\gamma_j) - \langle e^j, y - x\rangle,\end{aligned} \tag{20}$$

e a função P_i é dada como em (14) e e^j satisfaz

$$D_\psi(\tilde{x}^j,(\nabla\psi)^{-1}[\nabla\psi(\tilde{x}^j) - \gamma_j^{-1}e^j]) \leq \sigma D_\psi(\tilde{x}^j,x^j). \tag{21}$$

Passo 3: Defina λ^{j+1} como sendo

$$\lambda_i^{j+1} = \gamma_j(\mathcal{P}^s)'\left(g_i(\tilde{x}^j),\gamma_j,\lambda_i^j\right) \tag{22}$$

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

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

$$x^{j+1} = (\nabla\psi)^{-1}[\nabla\psi(\tilde{x}^j) - \gamma_j^{-1}e^j]. \tag{23}$$

A versão exata destes algoritmos também pode ser feita, considerando

$e^j = 0$ para todo $j \in \mathbb{N}$ e como no EPPM o passo 2 deste algoritmo é simplificado ao encontrar $\tilde{x}^j$ como solução de (20) com $e^j = 0$. A atualização no passo 4 é determinada pela solução $\tilde{x}^j$ obtida no passo 2.

Observação 10. *No caso particular em que $\psi = \psi^1$, tomando $\mathcal{P}^1(t,\gamma,\lambda) = \frac{1}{2}(\max\{0, \lambda + \frac{t}{\gamma}\}^2 - \lambda^2)$, $\mathcal{K} = \mathbb{R}^n$ e realizando alguns cálculos nos algoritmos de Lagrangeano aumentado descritos acima, obtemos exatamente os métodos de Lagrangeano aumentado para problemas de equilíbrio definidos em [3], no qual também podemos encontrar a análise de convergência para cada algoritmo neste caso. Portanto, no que segue basearemos a nossa análise para o caso $\psi = \psi^2$ o qual associaremos com a penalidade exponencial, $\mathcal{P}^2(s,\gamma,\lambda) = \lambda\exp(\frac{s}{\gamma}) - 1$.*

Logo, desta observação obtemos que:

1) $\nabla\psi(x) = (\ln(x_1), \ln(x_2), \ldots, \ln(x_n))^T = \ln(x) \in \mathbb{R}^n$.
2) $P_i(x,y,\lambda_i,\gamma) = \gamma\lambda_i\left[\exp\left(\frac{g_i(y)}{\gamma}\right) - \exp\left(\frac{g_i(x)}{\gamma}\right)\right]$.
3) $(\mathcal{P}^2)'\left(g_i(\tilde{x}^j), \gamma_j, \lambda_i^j\right) = \frac{\lambda_i^j}{\gamma_j}\exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right)$.

Assim, a regularização no passo 2 e a atualização das variáveis duais no Passo 3 de IALEM e IALPM, são dados pelas equações:

$$\begin{aligned}\widetilde{\mathcal{L}}_j^e(x,y) = f(x,y) &+ \gamma_j\langle \ln(x) - \ln(x^j), y - x\rangle \\ &+ \sum_{i=1}^m P_i(x,y,\lambda_i,\gamma_j) - \langle e^j, y - x\rangle. \end{aligned} \tag{24}$$

$$\lambda_i^{j+1} = \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right), \qquad (1 \le i \le m). \tag{25}$$

A. Análise de convergência

O objetivo desta análise de convergência é estabelecer a relação entre os algoritmos de ponto proximal descrito na seção anterior e os algoritmos de Lagrangeano aumentado com penalidade exponencial que acabamos de descrever. Para conseguir isto, é preciso exibir as relação que existe entre a função lagrangeana descrita em (13) e a função f de EP(f,C). O seguinte resultado, mostrado em [3], estabelece que as propriedades P1$-$P4 e P4$''$ necessárias para a convergência dos algoritmos de ponto proximal podem ser transferidas de forma natural para esta função lagrangeana dada em (13).

Proposição 11. *Seja EP(f,C), com C dado como em (12). Se f satisfaz P1$-$P4 e P4$''$ em $\mathcal{K}\times\mathcal{K}$, então a função lagrangeana $\mathcal{L}$ definida em (13), satisfaz P1$-$P4 e P4$''$ em $(\mathcal{K}\times\mathbb{R}^m_+)\times(\mathcal{K}\times\mathbb{R}^m_+)$.*

Como consequência deste resultado, podemos pensar em resolver um problema de equilíbrio dado por EP$(\mathcal{L}, \mathcal{K}\times R^m_+)$. Assim, a ideia será como em [3], utilizar a convergência dos algoritmos de ponto proximal IPPBPM e IPPEM aplicados ao problema EP$(\mathcal{L}, \mathcal{K}\times R^m_+)$ para mostrar que a sequência gerada por estes algoritmos é equivalente à sequência gerada pelos algoritmos de Lagrangeano aumentado. Assim, começamos definindo a regularização para $\mathcal{L}$.

Dado $(\bar{x},\bar{\lambda}) \in \mathcal{K}\times\mathbb{R}^m_{++}$, defina a função de regularização como sendo $\widehat{\mathcal{L}}^e : \mathcal{K}\times\mathbb{R}^m_{++}\times\mathcal{K}\times\mathbb{R}^m_{++} \to \mathbb{R}$ dada por

$$\begin{aligned}\widehat{\mathcal{L}}^e((x,\lambda),(y,\mu)) = \mathcal{L}\left((x,\lambda)(y,\mu)\right) &+ \gamma\langle\nabla\psi(x) - \nabla\psi(\bar{x}), y - x\rangle \\ &+ \gamma\langle\nabla\psi(\lambda) - \nabla\psi(\bar{\lambda}), \mu - \lambda\rangle - \langle e, y - x\rangle, \end{aligned} \tag{26}$$

em que $\nabla\psi(x) = \ln(x)$. Note que como os Algoritmos IPPBPM e IPPEM resolvem problemas para conjuntos convexos, fechados e não vazios e a função $\widehat{\mathcal{L}}^e$ tem domínio não vazio, convexo mas não fechado então para resolver este problema, vamos definir o conjunto poliedral $\mathcal{K}_\varepsilon = \{x \in \mathbb{R}^m_+ \mid x_i \ge \varepsilon\}$ com $\varepsilon > 0$. Assim a ideia é resolver o problema aproximado EP$(\mathcal{L}, \mathcal{K}\times\mathcal{K}_\varepsilon)$ para ε suficientemente pequeno. Agora, note que com esta regularização, as propriedades mostradas nas Proposições 6 e 7 são válidas para $\widehat{\mathcal{L}}^e$. Como consequência, podemos reescrever o algoritmos IPPBMP e IPPEM para resolver o problema EP$(\mathcal{L}, \mathcal{K}\times\mathcal{K}_\varepsilon)$, isto é, se assumirmos que as condições iniciais definidas no passo 1 de IPPBMP são válidas então podemos reescrever os passos 2 e 3 como:

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

$$\begin{aligned}\widehat{\mathcal{L}}_j^e\left((x,\lambda),(y,\mu)\right) = \mathcal{L}\left((x,\lambda)(y,\mu)\right) &+ \gamma_j\langle\nabla\psi(x) - \nabla\psi(x^j), y - x\rangle \\ &+ \gamma_j\langle\nabla\psi(\lambda) - \nabla\psi(\lambda^j), \mu - \lambda\rangle - \langle e^j, y - x\rangle \end{aligned} \tag{27}$$

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

$$\|e^j\| \le \begin{cases} \sigma\,\gamma_j D_\psi((\hat{x}^j,\hat{\lambda}^j),(x^j,\lambda^j)) & se \quad \|(x^j,\lambda^j) - (\hat{x}^j,\hat{\lambda}^j)\| < 1 \\ \sigma\gamma_j\nu_\psi((x^j,\lambda^j),1) & se \quad \|(x^j,\lambda^j) - (\hat{x}^j,\hat{\lambda}^j)\| \ge 1. \end{cases} \tag{28}$$

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

$$x^{j+1} = \operatorname*{argmin}_{x\in H_j} D_\psi(x, x^j). \tag{29}$$

$$\lambda^{j+1} = \hat{\lambda}^j. \tag{30}$$

Analogamente, podemos considerar o Algoritmo IPPEM para resolver o problema EP$(\mathcal{L}, \mathcal{K}\times\mathcal{K}_\varepsilon)$. Assim, os passos 2 e 3 são dados da seguinte forma:

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

$$\begin{aligned}\widehat{\mathcal{L}}_j^e\left((x,\lambda),(y,\mu)\right) = \mathcal{L}\left((x,\lambda)(y,\mu)\right) &+ \gamma_j\langle\nabla\psi(x) - \nabla\psi(x^j), y - x\rangle \\ &+ \gamma_j\langle\nabla\psi(\lambda) - \nabla\psi(\lambda^j), \mu - \lambda\rangle - \langle e^j, y - x\rangle \end{aligned} \tag{31}$$

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

$$\begin{aligned}&D_\psi\left((\hat{x}^j,\hat{\lambda}^j), (\nabla\psi)^{-1}[\nabla\psi((\hat{x}^j,\hat{\lambda}^j)) - \gamma_j^{-1}(e^j,0)]\right) \\ &\le \sigma D_\psi((\hat{x}^j,\hat{\lambda}^j),(x^j,\lambda^j)). \end{aligned} \tag{32}$$

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

$$x^{j+1} = (\nabla\psi)^{-1}[\nabla\psi(\hat{x}^j) - \gamma_j^{-1}e^j]. \quad (33)$$

$$\lambda^{j+1} = \hat{\lambda}^j. \quad (34)$$

Note que o Teorema 9 garante que a sequência gerada por quaisquer dos algoritmos IPPBMP e IPPEM convergem para uma solução do problema EP($\mathcal{L}, \mathcal{K} \times \mathcal{K}_\varepsilon$), evidentemente, fazendo as mudanças respectivas nesta proposição. Portanto, o objetivo a seguir será, como em [3], relacionar os algoritmos IPPBMP e IPPEM com os algoritmos IALPM e IALEM, respectivamente, para posteriormente mostrar que sobre certas hipóteses ambos algoritmos geram a mesma sequência obtendo, portanto, a solução do problema inicial EP(f, C). Antes de mostrar o teorema de equivalência, vamos precisar de alguns resultados auxiliares.

Considere a seguinte notação. Para cada $x \in \mathcal{K}$, $F_x : \mathcal{K} \to \mathbb{R}$ é dado como

$$F_x(y) = f(x, y) \quad (35)$$

Note que para cada $x \in \mathcal{K}$, a função F_x é convexa devido à propriedade P2. Usaremos esta notação no seguinte resultado o qual garante que um elemento x^* é solução de um problema de equilíbrio se e somente se x^* é o minimizador da função F_{x^*} no conjunto $\mathcal{K}$.

Proposição 12. *Considere o problema EP(f,$\mathcal{K}$). As seguintes afirmações são equivalentes:*

a) $x^* \in S(f, \mathcal{K})$
b) *x^* minimiza a função F_{x^*} no conjunto $\mathcal{K}$, com F_x dada como em (35).*

Demonstração. Veja a Proposição 3.4 de [3]. □

Agora, observe que a função ψ cumpre a seguinte propriedade:

H5: Se $(y^k)_{k\in\mathbb{N}} \subset int\mathcal{D}$ e $\lim_{k\to\infty} y^k = y \in \partial\mathcal{D}$ então $\lim_{k\to\infty} \langle \nabla h(y^k), (x - y^k)\rangle = -\infty$ para todo $x \in int\mathcal{D}$.

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

Proposição 13. *Considere EP($f, \mathcal{D}$) em que $\mathcal{D}$ é um domínio convexo. Sejam f satisfazendo P1$-$P4 e P6 e também alguma das hipóteses P4*, P4$'$ ou P4$''$, $\psi : \mathcal{D} \to \mathbb{R}$ satisfazendo H1$-$H2 e H5 e uma sequência exógena $(\gamma_j)_{j\in\mathbb{N}} \subset (\theta, \bar{\gamma})$. Se $(\hat{x}^j)_{j\in\mathbb{N}}$ é uma sequência gerada pelo passo 2 do algoritmo IPPEM (ou IPPBMP) então $(\hat{x}^j)_{j\in\mathbb{N}} \subset int\mathcal{D}$.*

Demonstração. Veja [12] pag. 24. □

Com este resultado, podemos ver que esta propriedade pode ser transferida para o problema EP($\mathcal{L}, \mathcal{K} \times \mathcal{K}_\varepsilon$), se consideramos o conjunto $\mathcal{K} \times \mathcal{K}_\varepsilon$ suficientemente próximo de $\mathcal{K} \times \mathbb{R}^n_+$ que é o problema que queremos resolver. Assim, com uma escolha de ε suficientemente pequena, garantimos que a sequência gerada pelos algoritmos de ponto proximal estarão sempre no interior do conjunto $\mathcal{K} \times \mathcal{K}_\varepsilon$ e aproximando o suficiente da solução no conjunto $\mathcal{K} \times \mathbb{R}^n_+$. Tendo isto, mostramos teorema de equivalência dos algoritmos de ponto proximal e Lagrangeano aumentado.

Teorema 14. *Suponha que EP(f, C) satisfaz P1$-$P4 em $\mathcal{K} \times \mathcal{K}$. Fixe a sequência $(\gamma_j)_{j\in\mathbb{N}} \subset \mathbb{R}_{++}$ e o erro relativo com tolerância $\sigma \in (0, 1)$. Sejam $((x^j, \lambda^j))_{j\in\mathbb{N}}$ uma sequência gerada pelo algoritmo IALPM (ou IALEM) aplicado ao EP(f,C), com erro $e^j \in \mathbb{R}^n$ dado por (17) (ou (21)) e $(\bar{x}^j, \bar{\lambda}^j)_{j\in\mathbb{N}}$ a sequência gerada pelo algoritmo IPPBPM (ou IPPEM) aplicado ao problema EP($\mathcal{L}, \mathcal{K} \times \mathcal{K}_\varepsilon$) com vetor de erro $(e^j, 0) \in \mathbb{R}^n \times \mathbb{R}^m$, e utilizando os mesmos γ_j's e σ. Se $(x^0, \lambda^0) = (\bar{x}^0, \bar{\lambda}^0)$ então $(x^j, \lambda^j) = (\bar{x}^j, \bar{\lambda}^j)$ para todo $j \in \mathbb{N}$.*

Demonstração. Procederemos por indução sobre j. Para $j = 0$ a igualdade é verdadeira por hipótese. Suponha que o teorema é válido para algum $j \in \mathbb{N}$, isto é, $(x^j, \lambda^j) = (\bar{x}^j, \bar{\lambda}^j)$. Note que pela Proposição 7 o subproblema EP($\widehat{\mathcal{L}}^e_j, \mathcal{K} \times \mathcal{K}_\varepsilon$), com $\widehat{\mathcal{L}}^e_j$ dado como em 31, gerado em cada iteração no Passo 2 do Algoritmo IPPBPM tem solução única, digamos $(\hat{x}^j, \hat{\lambda}^j)$. Agora, pela Proposição 12, temos que $(\hat{x}^j, \hat{\lambda}^j)$ resolve o problema convexo definido por

$$\min \quad \widehat{\mathcal{F}}_{(\hat{x}^j,\hat{\lambda}^j)}(x, \lambda) \quad (36)$$

$$s.a \quad (x, \lambda) \in \mathcal{K} \times \mathcal{K}_\varepsilon, \quad (37)$$

com $\widehat{\mathcal{F}}_{(\hat{x}^j,\hat{\lambda}^j)}(x, \lambda) = \widehat{\mathcal{L}}^e_j((\hat{x}^j, \hat{\lambda}^j), (x, \lambda))$. Como o problema é convexo então existem $\hat{\mu}^j \in N_{\mathcal{K}}(\hat{x}^j)$ e $\eta^j \in \mathbb{R}^m$, em que $N_{\mathcal{K}}(\hat{x}^j)$ é o cone normal no ponto $\hat{x}^j$, tais que:

$$\gamma_j[\log(\bar{x}^j) - \log(\hat{x}^j)] + e^j \in \partial F_{\hat{x}^j}(\hat{x}^j) + \sum_{i=1}^{m} \hat{\lambda}^j_i \partial g_i(\hat{x}^j) + \hat{\mu}^j, \quad (38)$$

$$-g_i(\hat{x}^j) + \gamma_j[\log(\hat{\lambda}^j_i) - \log(\bar{\lambda}^j_i)] = \eta^j_i, \; (1 \le i \le m), \quad (39)$$

$$\hat{\lambda}^j \ge \varepsilon$$

$$\eta^j \ge 0, \quad (40)$$

$$(\hat{\lambda}^j_i - \varepsilon)\eta^j_i = 0, \; (1 \le i \le m) \quad (41)$$

$$\hat{\mu}^j \in N_{\mathcal{K}}(\hat{x}^j). \quad (42)$$

Logo, combinando a igualdade $\eta^j_i = -g_i(\hat{x}^j) + \gamma_j[\log(\hat{\lambda}^j_i) - \log(\bar{\lambda}^j_i)]$, $(1 \le i \le m)$, obtida de (39), com as equações (40) e (41) podemos eliminar a variável η^j e assim, o sistema acima é equivalente a

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

$$-g_i(\hat{x}^j) + \gamma_j[\log(\hat{\lambda}^j_i) - \log(\bar{\lambda}^j_i)] \ge 0, \; (1 \le i \le m), \quad (43)$$

$$\hat{\lambda}^j_i \ge \varepsilon, \; (1 \le i \le m), \quad (44)$$

$$(\hat{\lambda}^j_i - \varepsilon)\left(-g_i(\hat{x}^j) + \gamma_j[\log(\hat{\lambda}^j_i) - \log(\bar{\lambda}^j_i)]\right) = 0, \; (1 \le i \le m), \quad (45)$$

$$\hat{\mu}^j \in N_{\mathcal{K}}(\hat{x}^j). \quad (46)$$

Novamente combinando (43), (44) e (45) obtemos

$$\gamma_j[\log(\bar{x}^j) - \log(\hat{x}^j)] + e^j \in \partial F_{\hat{x}^j}(\hat{x}^j) + \sum_{i=1}^{m} \hat{\lambda}_i^j \partial g_i(\hat{x}^j) + \hat{\mu}^j, \quad (47)$$

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

$$\hat{\mu}^j \in N_{\mathcal{K}}(\hat{x}^j). \quad (48)$$

Assim, como os pontos gerados por esta sequência sempre estão no interior de $\mathcal{K} \times \mathcal{K}_\varepsilon$, então vale que

$$\hat{\lambda}_i^j = \bar{\lambda}_i^j \exp\left(\frac{g_i(\hat{x}^j)}{\gamma_j}\right), \ (1 \le i \le m) \quad (49)$$

Finalmente, substituindo (49) em (47) concluímos que

$$\gamma_j[\log(\bar{x}^j) - \log(\hat{x}^j)] + e^j \in \partial F_{\hat{x}^j}(\hat{x}^j) + \sum_{i=1}^{m} \bar{\lambda}_i^j \exp\left(\frac{g_i(\hat{x}^j)}{\gamma_j}\right) \partial g_i(\hat{x}^j) + \hat{\mu}^j, \quad (50)$$

$$\hat{\mu}^j \in N_{\mathcal{K}}(\hat{x}^j).$$

Por outro lado, observe que o Passo 2 do Algoritmo IALPM exige uma solução $\tilde{x}^j$ do subproblema EP($\widetilde{\mathcal{L}}_j^e, \mathcal{K}$), que pela Proposição 12 sabemos que equivale a que o sistema não linear e não diferenciável

$$\gamma_j[\log(x^j) - \log(\tilde{x}^j)] + e^j \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^{m} \lambda_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right) \partial g_i(\tilde{x}^j) + \widetilde{\mu}^j, \quad (51)$$

$$\widetilde{\mu}^j \in N_{\mathcal{K}}(\tilde{x}^j). \quad (52)$$

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

$$\gamma_j[\log(\bar{x}^j) - \log(\tilde{x}^j)] + e^j \in \partial F_{\tilde{x}^j}(\tilde{x}^j) + \sum_{i=1}^{m} \bar{\lambda}_i^j \exp\left(\frac{g_i(\tilde{x}^j)}{\gamma_j}\right) \partial g_i(\tilde{x}^j) + \tilde{\mu}^j, \quad (53)$$

$$\tilde{\mu}^j \in N_{\mathcal{K}}(\tilde{x}^j). \quad (54)$$

Agora, note que se substituímos $\tilde{x}^j$ por $\hat{x}^j$ em (53) e (54) obtemos exatamente o sistema (50), e portanto, $\hat{x}^j \in S(\widetilde{\mathcal{L}}_j^e, \mathbb{R}^n)$. Como pela Proposição 7 as soluções do sistema (50) são determinadas de forma única então concluímos que

$$\hat{x}^j = \tilde{x}^j. \quad (55)$$

Por outro lado, se combinamos a equação (19) no Passo 4 de IALPM (ou (23) no algoritmo IALEM) para resolver EP(f, C), a equação (29) no Passo 3 do algoritmo IPPBPM (ou (33) em IPPEM) para EP($\mathcal{L}, \mathcal{K} \times \mathcal{K}_\varepsilon$), a hipótese indutiva e a igualdade (55) concluímos que $x^{j+1} = \bar{x}^{j+1}$. Agora, para a atualização das variáveis duais, basta olhar para as equações (34) e (49), e assim obter que:

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

Logo, comparando a última igualdade com (22), levando em conta (55) e $\bar{\lambda}^j = \lambda^j$ concluímos que $\bar{\lambda}^{j+1} = \lambda^{j+1}$. $\square$

Tendo mostrado a equivalência destes algoritmos e sabendo que um deles converge então podemos concluír que o outro também converge. Concluindo, desta forma, a convergência do método de Lagrangeano aumentado exponencial.

Destacamos que, atualmente, estamos realizando implementações numéricas de cada um dos algoritmos apresentados aqui e esperamos estar com estes resultados prontos em breve.

AGRADECIMENTOS

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

REFERÊNCIAS

[1] G. Bigi, G. Castellani, M. Pappalardo and M. Passacantando, *Existence and solution methods for equilibria*. European Journal of Operational Research, vol. 227, 1–11, 2013.
[2] E. Blum and W. Oettli, *From Optimization and Variational Inequality to Equilibrium Problems*. The Mathematics Student, vol. 63, 123–145, 1994.
[3] A. N. Iusem and M. Nasri. *Augmented Lagrangian methods for equilibrium problems*. RAIRO, Recherche Opérationnelle, Vol. 44, 5-26, 2010.
[4] A. N. Iusem and M. Nasri, *Inexact Proximal Point Methods for Equilibrium Problems in Banach Spaces*, Numerical Functional Analysis and Optimization, vol. 28, 1279-1308, 2007.
[5] G. Kassay and R. Burachik, *On a generalized proximal point method for solving equilibrium problems in Banach spaces*, Nonlinear Analysis: Theory, Methods and Applications, vol. 75, 2012.
[6] K. Fan, *A generalization of Tychonoff's fixed point theorem*, Mathematische Annalen, vol. 142, 305-310, 1961.
[7] K. Fan, *A minimax inequality and applications*, *Inequality III* (O. Shisha, editor), 1972.
[8] A. N. Iusem and W. Sosa, *A proximal point method for equilibrium problem in Hilbert spaces*, Optimization, vol. 59, 1259-1274, 2010.
[9] A. N. Iusem and W. Sosa, *Iterative Algorithms for Equilibrium Problems*, Optimization, vol. 52, 301-316, 2003.
[10] A. N. Iusem and G. Kassay and W. Sosa, *On certain conditions for the existence of solutions of equilibrium problems*, Mathematical Programming, vol. 116, 259-273, 2009.
[11] D. Butnariu and A. N. Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization*, Springer Science+Business Media Dordrecht, vol. 40, 2000.
[12] A. N. Iusem, *Métodos de Ponto Proximal em Otimização*, 20 Coloquio Brasileiro de matemática (IMPA), 1995.